SCHUR-HORN THEOREM

XIAOXIANG ZHOU

Abstract. In this article, I will use the Atiyah-Guillemin-Sternberg Convexity theorem to prove the Schur-Horn theorem, which is a beautiful theorem in linear algebra, with deep symplectic geometry theory behind it. To introduce the AGM theorem, we first grasp the tools: the Lie bracket and the Exponential map; then we will focus on the vector field induced by the group action $\mathbb{T}^n \circlearrowright \mathcal{H}_\lambda$, and use the symplectic structure on $\mathcal{H}_\lambda$ to convert the vector field to an exact 1-form, and then natually introduce the moment map on $\mathcal{H}_\lambda$. After that, we will state the AGM theorem and prove the Schur-Horn theorem.

1. Introduction

Given a Hermitian matrix $A = (a_{ij}) \in \mathbb{C}^n$ with eigenvalues

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)^T \in \mathbb{R}^n$$

We want to see:

Question:What do the diagonal elements

$$(a_{11}, a_{22}, \ldots, a_{nn})$$

look like?

Facts. *(Obvious)*

- $A^H = A \Rightarrow a_{11}, a_{22}, \ldots, a_{nn} \in \mathbb{R}$
- A *is* ***unitary similar*** *to* $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$
  $\Rightarrow \sum_{i=1}^n a_{ii} = \mathrm{tr}\, A = tr(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)) = \sum_{i=1}^n \lambda_i$
- $\forall\, \tau \in S_n, \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ *is unitary similar to* $\mathrm{diag}(\lambda_{\tau(1)}, \ldots, \lambda_{\tau(n)})$
  $\Rightarrow$ *WLOG, we can rearrange* $(\lambda_1, \ldots, \lambda_n)$ *s.t.*
  $$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$
  NOTICE:After that we will assume $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$.

Facts. *(Not Obvious)*

- $\forall\, i \in \{1, \ldots, n\}, \lambda_n \leqslant a_{ii} \leqslant \lambda_1$
- $\forall\, k \in \{1, \ldots, n\}, \sum_{i=1}^k a_{ii} \leqslant \sum_{i=1}^k \lambda_i$ *

Denote

$$\begin{aligned}
&\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)^T \in \mathbb{R}^n \\
&\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)^T \in \mathbb{R}^{n\times n} \\
&\mathcal{H}(n) = \{A \in \mathbb{C}^{n\times n} \mid A^H = A\} \\
&\mathcal{H}_\lambda = \{A \in \mathcal{H}(n) \mid A \text{ is unitary similar to } \Lambda\}
\end{aligned}$$

*Issai Schur (Russian,1875-1941) proved the above-mentioned inequalities in 1923.

$$\begin{array}{rccc} \pi: & \mathcal{H}(n) & \longrightarrow & \mathbb{R}^n \\ & A = (a_{ij})_{i,j=1}^n & \mapsto & (a_{11}, a_{22}, \ldots, a_{nn})^T \end{array}$$

Theorem 1.1. *(Schur-Horn) $\pi(\mathcal{H}_\lambda)$ is a* ***convex polyhedron*** *in $\mathbb{R}^n$ whose vertices are*

$$(\lambda_{\tau(1)}, \ldots, \lambda_{\tau(n)})^T \in \mathbb{R}^n$$

where $\tau \in S_n$.†

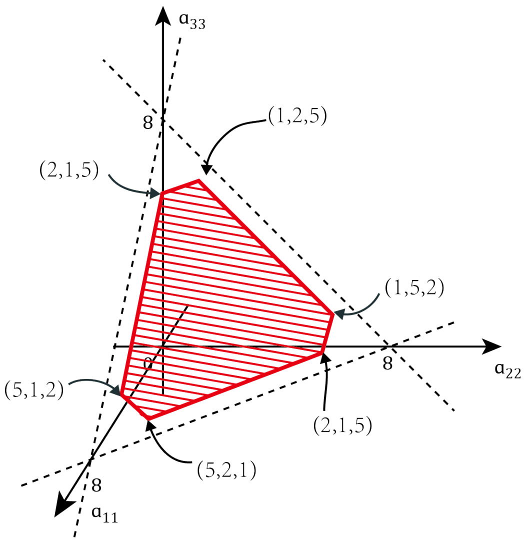
With these facts in mind, we will first discuss some examples.

Example 1.2. *(trivial) when $\boldsymbol{\lambda} = (\lambda_0, \lambda_0, \ldots, \lambda_0)^T$, we have*

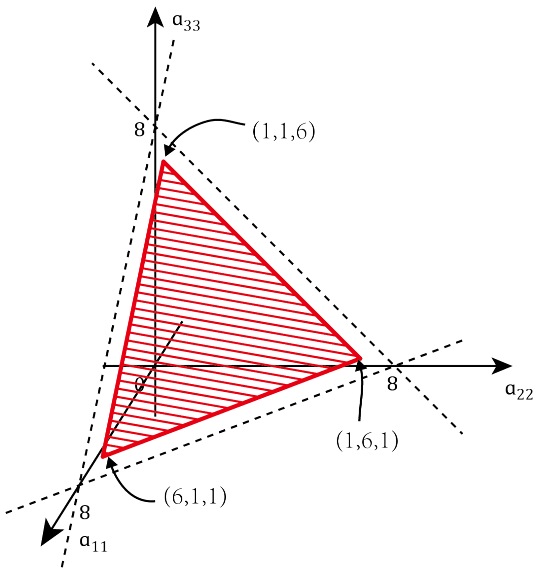
$$\begin{aligned} \Lambda &= \lambda_0 I \\ \mathcal{H}_\lambda &= \{A \in \mathbb{C}^{n\times n} \mid \exists\, U \in U(n), A = U(\lambda_0 I)U^H = \lambda_0 I\} \\ &= \{\lambda_0 I\} \qquad\qquad \textit{has only one element!} \end{aligned}$$

We leave 2-dimension example at last because it's computable.

Example 1.3. *(3-dimension condition)*
when $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)^T$, it's almost impossible to calculate, so we only draw out the final result:



$(\lambda_1, \lambda_2, \lambda_3)^T = (5, 2, 1)^T$



$(\lambda_1, \lambda_2, \lambda_3)^T = (6, 1, 1)^T$
"degenerate"

Example 1.4. *(2-dimension condition) we have*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{H}_\lambda \Leftrightarrow \exists\, U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in U(2),$$

†Alfred Horn (Amerian, UCLA) proved it in 1954.

$$\begin{aligned}
A &= U\Lambda U^H \\
&= \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} \overline{u_{11}} & \overline{u_{21}} \\ \overline{u_{12}} & \overline{u_{22}} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1|u_{11}|^2 + \lambda_2|u_{12}|^2 & \lambda_1 u_{11}\overline{u_{21}} + \lambda_2 u_{12}\overline{u_{21}} \\ \lambda_1 u_{21}\overline{u_{11}} + \lambda_2 u_{22}\overline{u_{12}} & \lambda_1|u_{21}|^2 + \lambda_2|u_{22}|^2 \end{pmatrix} \\
&= \lambda_2 I + (\lambda_1 - \lambda_2) \begin{pmatrix} |u_{11}|^2 & u_{11}\overline{u_{21}} \\ \lambda_1 u_{21}\overline{u_{11}} & \lambda_1|u_{21}|^2 \end{pmatrix}
\end{aligned}$$

Denote the line segment drawed in the figure 1 as Γ*, then*

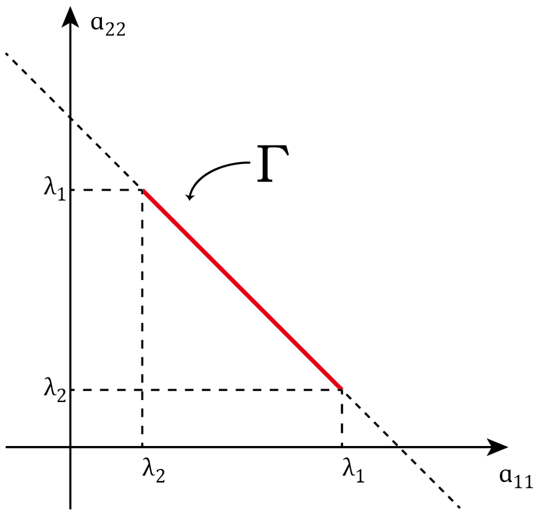


FIGURE 1

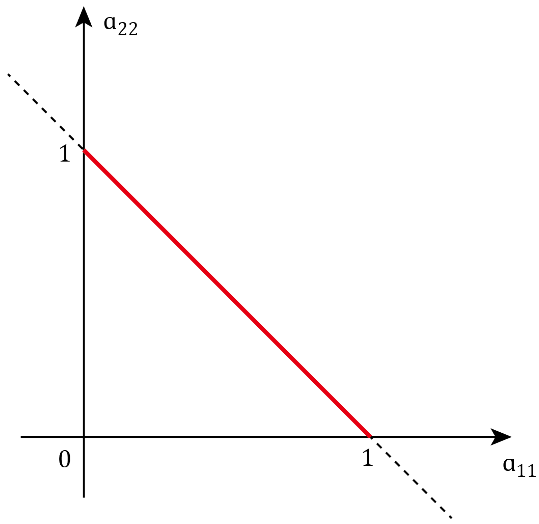


FIGURE 2

- $\pi(\mathcal{H}_\lambda) \subseteq \Gamma$ *because* $\lambda_1|u_{11}|^2 + \lambda_2|u_{12}|^2$ *is the convex combination of* λ_1, λ_2.
- $\Gamma \subseteq \pi(\mathcal{H}_\lambda)$ *because we can take*

$$\begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

.

Actually one can compute more:
WLOG(or take the coordinate trasformation), we only consider the condition when

- $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^T = (1, 0)^T$
- $A = \begin{pmatrix} |u_{11}|^2 & u_{11}\overline{u_{21}} \\ \lambda_1 u_{21}\overline{u_{11}} & \lambda_1|u_{21}|^2 \end{pmatrix}$.

Now we can calculate out

$$\mathcal{H}_\lambda = \left\{ \begin{pmatrix} a & e^{i\varphi}\sqrt{a(1-a)} \\ e^{-i\varphi}\sqrt{a(1-a)} & 1-a \end{pmatrix} \middle| a \in [0,1], 0 \leqslant \varphi < 2\pi \right\}$$

Now we know explicitly

$$\pi(\mathcal{H}_\lambda) = \{(a, 1-a) \mid 0 \leqslant a \leqslant 1\}$$

Moreover, $\pi(\mathcal{H}_\lambda)$ is a manifold diffeomorphic to S^2:

$$\begin{array}{cccc} \Phi: & \mathcal{H}_\lambda & \longrightarrow & S^2 \\ & \begin{pmatrix} a & e^{i\varphi}\sqrt{a(1-a)} \\ e^{-i\varphi}\sqrt{a(1-a)} & 1-a \end{pmatrix} & \mapsto & (\varphi, a) \end{array}$$

Remark 1.5. What is a manifold? Manifold is a VERY GOOD geometric object which always look like R^n.

We will find out more information through this isomorphism.

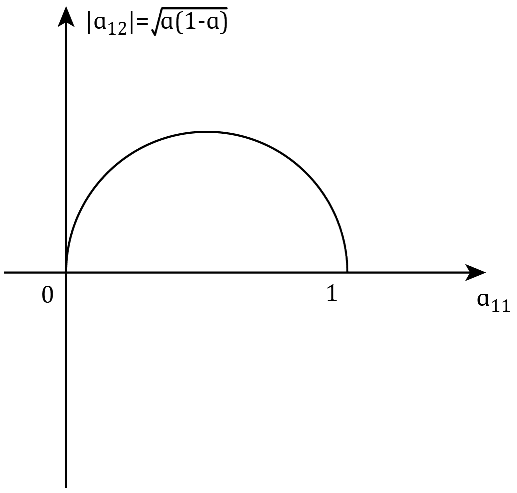


FIGURE 3

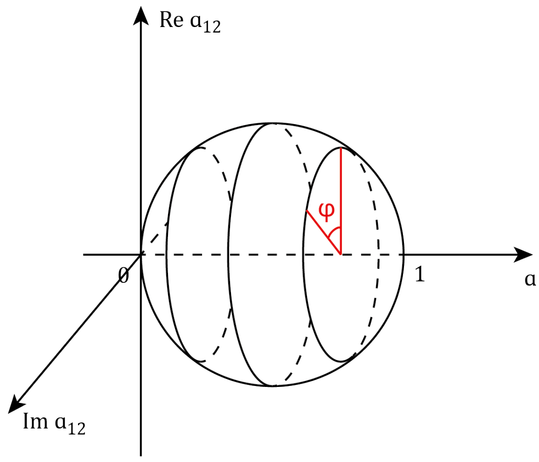


FIGURE 4

2. SIMPLE TOOLS

Let us deriate from the phenomenon for a while to obtain the most basic tools:the **Lie bracket** and the **Exponential map**.

Lie bracket.

Definition 2.1. *the Lie bracket of $M_n(\mathbb{C})$ is*

$$\begin{array}{rccc} [\ \ ,\ \ ]: & M_n(\mathbb{C}) \times M_n(\mathbb{C}) & \longrightarrow & \mathbb{C} \\ & (A, B) & \mapsto & [A, B] := AB - BA \end{array}$$

Proposition 2.2. *For any $c_1, c_2 \in \mathbb{C}, A, A_1, A_2, B, C \in M_n(\mathbb{C})$, we have the following properties:*

- *(Skew-Symmetric)* $[A, B] = -[B, A]$*;*
- *(Linear)* $[c_1A_1 + c_2A_2, B] = c_1[A_1, B] + c_2[A_2, B]$*;*
- *(Jacobi-Identity)* $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$*;*
- $[A, B]^H = [B^H A^H]$*;*
- $\mathrm{tr}(A[B, C]) = \mathrm{tr}([A, B]C)$

Proof: Exercise.

Exponential map.

Definition 2.3 (The Exponentinal map for Matrix)**.** *Suppose* $A \in M_n(\mathbb{C})$*, then we define*

$$P_n(A) := \sum_{i=0}^{n} \frac{A^i}{i!}$$

$$\exp(A) := e^A := \lim_{n\to\infty} P_n(A)$$

Remark 2.4. about the definition

- By defining the norm on $M_n(\mathbb{C})$, one is easy to find out the existence and uniqueness of the definition.
- Generally $e^A e^B \neq e^{A+B}$. But we still have

$$AB = BA \Rightarrow e^A e^B = e^{A+B}$$

.

- Like polynomials, some properties are easily derived from the definition:
  - $\forall\, U \in U(n), Ue^X U^H = e^{UXU^H}$
  - $(e^X)^H = e^{X^H}$
  - $\frac{d}{dt}e^{tX} = Xe^{tX}$; especially $\frac{d}{dt}\big|_{t=0} e^{tX} = X$
- Sometimes we denote $\exp(X) = e^X$ to enlarge superscript.
- Someone may think the Exponential map as "walking along the vector field Xe^{tX} (in $GL_n(\mathbb{C})$) for t times". You can easily check (if you've learned about the Differential Manifold) that $\exp(tX)$ is just an integral curve $\gamma_X(t)$ in $GL_n(\mathbb{C})$.

3. Group Actions

3.1. **Group action on $\mathcal{H}(n)$.** We have **VERY NICE** group actron on $\mathcal{H}(n)$:

$$U(n) \circlearrowright \mathcal{H}(n)$$

$$U \cdot H = UHU^H$$

Remark 3.1. One can easily check that this is really the group action:

- $UHU^H \in \mathcal{H}(n)$
- $I \cdot H = H$
- $(U_1U_2) \cdot H = U_1 \cdot (U_2 \cdot H)$

Question: What is the orbit of this action?
Answer: From the linear algebra theory,

$$A \in \mathcal{H}_\lambda \Leftrightarrow \exists\, U \in U(n), A = U\Lambda U^H$$

As a result,

Proposition 3.2. *The orbit of the group action is*

$$\mathcal{H}_\lambda = \{A \in \mathcal{H}(n) \mid A \textit{ has eigenvalues } \boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^T\}$$

Moreover

- $\mathcal{H}(n)$ is a $\mathbb{R}$-linear space, thus naturally a manifold

- $U(n)$ is a Lie group

So from the Lie group's theory we can obtain

Proposition 3.3. *$\mathcal{H}_\lambda$ is a manifold.*

This is not so surprising because we have calculated the $\mathcal{H}_{(1,0)^T}$ and "verified" that this is a manifold diffeomorphic to S^2. Later we will see more structures on $\mathcal{H}_\lambda$, and these structures in all will help us to find out more informations about $\pi(\mathcal{H}_\lambda)$.

3.2. **Subgroup actions.** We have found

$$\begin{aligned} S^1 &= \left\{ \begin{pmatrix} e^{i\theta} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} : \theta \in \mathbb{R} \right\} \subseteq U(n) \\ \mathbb{T}^n &= S^1 \times S^1 \times \cdots \times S^1 \\ &= \left\{ \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} : \theta_1, \ldots, \theta_n \in \mathbb{R} \right\} \subseteq U(n) \end{aligned}$$

Then $S^1 \subseteq \mathbb{T}^n \subseteq U(n)$.

We have the induced subgroup actions:

$$\begin{gathered} S^1 \circlearrowleft \mathcal{H}_\lambda \\ A \cdot H = AHA^H \\ \theta \cdot H = \begin{pmatrix} e^{i\theta} & \\ & I_{n-1} \end{pmatrix} H \begin{pmatrix} e^{-i\theta} & \\ & I_{n-1} \end{pmatrix} \end{gathered}$$

$$\begin{gathered} \mathbb{T}^n \circlearrowleft \mathcal{H}_\lambda \\ A \cdot H = AHA^H \\ \theta \cdot H = \begin{pmatrix} e^{i\theta_1} & & \\ & \ddots & \\ & & e^{i\theta_n} \end{pmatrix} H \begin{pmatrix} e^{-i\theta_1} & & \\ & \ddots & \\ & & e^{-i\theta_n} \end{pmatrix} \end{gathered}$$

We may split the matrix H into 4 different parts:

$$H = \left(\begin{array}{c|c} H_{11} & H_{12} \\ \hline H_{21} & H_{22} \end{array} \right)$$

Then

$$\begin{aligned} \theta \cdot H &= \begin{pmatrix} e^{i\theta} & \\ & I_{n-1} \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} e^{-i\theta} & \\ & I_{n-1} \end{pmatrix} = \begin{pmatrix} H_{11} & e^{i\theta} H_{12} \\ e^{-i\theta} H_{21} & H_{22} \end{pmatrix} \\ \frac{d}{d\theta}(\theta \cdot H) &= \begin{pmatrix} 0 & ie^{i\theta} H_{12} \\ -ie^{-i\theta} H_{21} & 0 \end{pmatrix} \end{aligned}$$

Remark 3.4. Notice that the group action $S^1 \circlearrowleft \mathcal{H}_\lambda$ doesn't change the diagonal components.Similarly, one can easily verify that the group action $\mathbb{T}^n \circlearrowleft \mathcal{H}_\lambda$ also keeps the diagonal components. Thus we may think "the group actions decrease the other unrelated degree of freedom", and thus "gives the invariance" of $\mathcal{H}_\lambda$.

3.3. **The induced vector field of group action.**

Definition 3.5. *Suppose $j \in \{1, \dots, n\}$ the group $\mathbb{T}^n$ acts on $\mathcal{H}_\lambda$, then the induced vector field X_j at $H \in \mathcal{H}_\lambda$ is the matrix*

$$X_j(H) = \frac{d}{dt}\bigg|_{t=0} ((0, \dots, t, \dots, 0) \cdot H)$$

Example 3.6. *We have computed*

$$X_1(H) = \frac{d}{dt}\bigg|_{t=0} ((t, 0, \dots, 0) \cdot H) = \begin{pmatrix} & iH_{12} \\ -iH_{21} & \end{pmatrix}$$

Similarly, if $H = (h_{ij})_{i,j=1}^n$, then

$$X_j(H) = \begin{pmatrix} & & ih_{1j} & & \\ & & \vdots & & \\ -ih_{j1} & \cdots & 0 & \cdots & -h_{jn} \\ & & \vdots & & \\ & & ih_{nj} & & \end{pmatrix}$$

Example 3.7. *When $n = 2$, $H = \begin{pmatrix} a & e^{i\varphi}\sqrt{a(1-a)} \\ e^{-i\varphi}\sqrt{a(1-a)} & 1-a \end{pmatrix}$,*

$$\begin{aligned} X_1(H) &= \begin{pmatrix} 0 & ih_{12} \\ -ih_{21} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & ie^{i\varphi}\sqrt{a(1-a)} \\ -ie^{-i\varphi}\sqrt{a(1-a)} & 0 \end{pmatrix} \end{aligned}$$

Notice that

$$\frac{\partial H}{\partial \varphi} = \begin{pmatrix} 0 & ie^{i\varphi}\sqrt{a(1-a)} \\ -ie^{-i\varphi}\sqrt{a(1-a)} & 0 \end{pmatrix}$$

4. New Tools

4.1. **Symplectic manifold.**

Roughly speaking, the symplectic manifold is the manifold with a 2-form which locally looks like $\sum_{i=1}^n dx^i \wedge dy^i$.

Now suppose M is a manifold of dimension 2n.

Definition 4.1. *A symplectic form on M is a 2-form $w \in \Lambda^2 T^* M$ on M such that*

- *w is closed: $dw = 0$.*
- *w is non-degenerate: $w \wedge w \wedge \cdots \wedge w \neq 0$ is a volume form on M.*

The pair (M, w) is called a ***symplectic manifold****.*

Remark 4.2. Compared with Riemann metric g:

- g can be defined on any manifold, while w can't (dimension =2n, orientable, and so on).
- g is symmetric while w is skew-symmetric.
- By Darboux theorem, w looks like $\sum_{i=1}^n dx^i \wedge dy^i$ near any $p \in M$, while g has plenty of local geometric structrues (such as curvature and connection)

- g_p gives an isomorphism
$$\begin{aligned} g_p^{\#}: T_pM &\longrightarrow T_p^*M \\ X_p &\mapsto g_p(X_p,-) \end{aligned}$$
While w also gives an isomorphism
$$\begin{aligned} w_p^{\#}: T_pM &\longrightarrow T_p^*M \\ X_p &\mapsto w_p(X_p,-) \end{aligned}$$
We will use this isomorphism to convert a vector field (which I have mentioned, induced by group action) to an exact 1-form.

Example 4.3. *$(\mathbb{R}^{2n}, w)$ is a symplectic manifold with chart coordinate $(x_1,\dots,x_n,y_1,\dots,y_n)$*
$$w=\sum_{i=1}^n dx^i\wedge dy^i$$
Verify:
- $w\in\Lambda^2T^*M$
- $dw=\sum_{i=1}^n d1\wedge dx^i\wedge dy^i=0$
- $w\wedge w\wedge\cdots\wedge w=n!\,dx^1\wedge dy^1\wedge\cdots dx^n\wedge dy^n\neq 0$

Example 4.4. *(S^2,w) is a symplectic manifold where w is the canonical volume form of S^2. in $S^2\setminus\{North, South\}$, $d\theta\wedge dh$ is the local representation of w.*

Verify:
- $w\in\Lambda^2T^*M$
- $dw=0$ *because w is a top form.*
- w *is no-degenerate since it is already a volume form.*

Example 4.5. *From the diffeomorphism*
$$\begin{array}{cccc} \Phi: & \mathcal{H}_{(1,0)^T} & \longrightarrow & S^2 \\ & H(a,\varphi)\hat{=}\begin{pmatrix} a & e^{i\varphi}\sqrt{a(1-a)} \\ e^{-i\varphi}\sqrt{a(1-a)} & 1-a \end{pmatrix} & \mapsto & (\varphi,a) \end{array}$$
one can obtain a natural symplectic form on $\mathcal{H}_{(1,0)^T}$:
$$\begin{array}{cccc} \Phi^*: & \Omega^2(S^2) & \longrightarrow & \Omega^2(\mathcal{H}_{(1,0)^T}) \\ & w=d\theta\wedge dh & \mapsto & w_{can} \end{array}$$
We can calculate $(a\neq 0,1)$
$$(d\Phi)^{-1}(\frac{\partial}{\partial\theta})=\begin{pmatrix} 0 & ie^{i\varphi}\sqrt{a(1-a)} \\ -ie^{-i\varphi}\sqrt{a(1-a)} & 0 \end{pmatrix}\hat{=}\frac{\partial}{\partial\phi}=X_1(H(a,\phi))$$
$$(d\Phi)^{-1}(\frac{\partial}{\partial h})=\begin{pmatrix} 1 & e^{i\varphi}\dfrac{1-2a}{2\sqrt{a(1-a)}} \\ e^{-i\varphi}\dfrac{1-2a}{2\sqrt{a(1-a)}} & -1 \end{pmatrix}\hat{=}\frac{\partial}{\partial a}$$
$$T_{(e^{i\varphi},a)}S^2=\left\langle \left.\frac{\partial}{\partial\theta}\right|_{(e^{i\varphi},a)}, \left.\frac{\partial}{\partial h}\right|_{(e^{i\varphi},a)}\right\rangle_{span}\Rightarrow T_{H(a,\varphi)}\mathcal{H}_{(1,0)^T}=\left\langle \left.\frac{\partial}{\partial\phi}\right|_{H(a,\varphi)}, \left.\frac{\partial}{\partial a}\right|_{H(a,\varphi)}\right\rangle_{span}$$

$$1 = w(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial h}) = w_{can}(\frac{\partial}{\partial\phi}, \frac{\partial}{\partial a}) = w^{\#}_{can}(\frac{\partial}{\partial\phi})(\frac{\partial}{\partial a}) = w^{\#}_{can}(X_1)(\frac{\partial}{\partial a}) \Rightarrow w^{\#}_{can}(X_1) = da$$

Remark 4.6. In general $\mathcal{H}_\lambda$ is also a symplectic manifold whose symplectic form can be written as (if $H = U\Lambda U^H, X = A\Lambda U^H + U\Lambda A^H, Y = B\Lambda U^H + U\Lambda B^H$)

$$w_\lambda\big|_H(X, Y) = i\,\mathrm{tr}(\Lambda[U^H A, U^H B])$$

Moreover, $w^{\#}_\lambda(X_i)$ is exact, i,e

$$\exists\, f \in C^\infty(\mathcal{H}_\lambda) \text{ such that } w^{\#}_\lambda(X_i) = df$$

This function f will bw denoted "**the moment map**".

We will verify that when $\boldsymbol{\lambda} = (1, 0)^T$, this symplectic structure defined coincide with w_{can} we've encountered. As follows:

$$\begin{aligned}
H(a,\varphi) &= \begin{pmatrix} a & e^{i\varphi}\sqrt{a(1-a)} \\ e^{-i\varphi}\sqrt{a(1-a)} & 1-a \end{pmatrix} \overset{a=\cos^2\theta}{\underset{0<\theta<\pi/2}{=\!=\!=}} \begin{pmatrix} \cos^2\theta & e^{i\varphi}\sin\theta\cos\theta \\ e^{-i\varphi}\sin\theta\cos\theta & \sin^2\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & -e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \begin{pmatrix} \cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix} \\
&\Rightarrow U = \begin{pmatrix} \cos\theta & -e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix} \qquad U^H = \begin{pmatrix} \cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial H(a,\varphi)}{\partial\varphi} = \frac{\partial(U\Lambda U^H)}{\partial\varphi} = \frac{\partial U}{\partial\varphi}\Lambda U^H + U\Lambda\left(\frac{\partial U}{\partial\varphi}\right)^H \\
\Rightarrow & A = \frac{\partial U}{\partial\varphi} = \begin{pmatrix} 0 & -ie^{i\varphi}\sin\theta \\ -ie^{-i\varphi}\sin\theta & 0 \end{pmatrix} \\
\Rightarrow & U^H A = \begin{pmatrix} \cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 & -ie^{i\varphi}\sin\theta \\ -ie^{-i\varphi}\sin\theta & 0 \end{pmatrix} \\
&\qquad = -i\sin\theta \begin{pmatrix} \sin\theta & -e^{i\varphi}\cos\theta \\ e^{-i\varphi}\cos\theta & -\sin\theta \end{pmatrix} \\
&\frac{\partial H(a,\varphi)}{\partial a} = \frac{\partial(U\Lambda U^H)}{\partial a} = \frac{\partial U}{\partial a}\Lambda U^H + U\Lambda\left(\frac{\partial U}{\partial a}\right)^H \\
\Rightarrow & B = \frac{\partial U}{\partial a} = \frac{1}{2\cos\theta\sin\theta}\frac{\partial U}{\partial\theta} = \frac{1}{2\cos\theta\sin\theta}\begin{pmatrix} -\sin\theta & -e^{i\varphi}\cos\theta \\ e^{-i\varphi}\cos\theta & -\sin\theta \end{pmatrix} \\
\Rightarrow & U^H B = \frac{1}{2\cos\theta\sin\theta}\begin{pmatrix} \cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} -\sin\theta & -e^{i\varphi}\cos\theta \\ e^{-i\varphi}\cos\theta & -\sin\theta \end{pmatrix} \\
&\qquad = \frac{1}{2\cos\theta\sin\theta}\begin{pmatrix} & -e^{i\varphi} \\ -e^{-i\varphi} & \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
[U^H A, U^H B] =& -\frac{i}{2\cos\theta}\left[\begin{pmatrix} \sin\theta & -e^{i\varphi}\cos\theta \\ e^{-i\varphi}\cos\theta & -\sin\theta \end{pmatrix}, \begin{pmatrix} & -e^{i\varphi} \\ -e^{-i\varphi} & \end{pmatrix}\right] \\
=& -\frac{i}{2\cos\theta}\left\{\begin{pmatrix} \cos\theta & -e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix} - \begin{pmatrix} -\cos\theta & e^{i\varphi}\sin\theta \\ -e^{-i\varphi}\sin\theta & -\cos\theta \end{pmatrix}\right\} \\
=& -\frac{i}{\cos\theta}\begin{pmatrix} \cos\theta & -e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
w_\lambda\big|_H(X,Y) =& i\,\mathrm{tr}(\Lambda[U^H A, U^H B]) \\
=& \frac{1}{\cos\theta}\,\mathrm{tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \cos\theta & -e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & \cos\theta \end{pmatrix}\right) \\
=& 1
\end{aligned}$$

4.2. **Moment Map.**

Definition 4.7. *Suppose $S^1 \circlearrowleft \mathcal{H}_\lambda$,then the moment map is a map*

$$\mu : \mathcal{H}_\lambda \longrightarrow \mathbb{R}$$

such that $w_{can}^{\#}(X_1) = d\mu$.

From Example 4.5 we can see, the moment map of $S^1 \circlearrowleft \mathcal{H}_{(1,0)^T}$ is

$$\begin{aligned}
\mu : \quad & \mathcal{H}_\lambda \longrightarrow \mathbb{R} \\
& \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto a_{11}
\end{aligned}$$

Definition 4.8. *Suppose $\mathbb{T}^n \circlearrowleft \mathcal{H}_\lambda$,then the moment map is a map*

$$\begin{aligned}
\mu : \mathcal{H}_\lambda & \longrightarrow \mathbb{R}^n \\
A & \mapsto (\mu_1(A), \dots, \mu_n(A))^T
\end{aligned}$$

such that for any $i \in \{1, \dots, n\}, w_{can}^{\#}(X_i) = d\mu_i$.

Remark 4.9. Like the examples we have seen, in general, if $\mathbb{T}^n \circlearrowleft \mathcal{H}_\lambda$ in a canonical way, then

$$\begin{aligned}
\mu = \pi : \quad \mathcal{H}_\lambda & \longrightarrow \mathbb{R}^n \\
A = (a_{ij})_{i,j=1}^n & \mapsto (a_{11}, \dots, a_{nn})^T
\end{aligned}$$

is just the projection to its diagonal components! Its proof require the knowledge of coadjoint orbit, so I regret that I'll skip it.

Definition 4.10. *We will call $(\mathcal{H}_\lambda, w_\lambda, \mathbb{T}^r, \mu)$ as the **Hamiltonian** $\mathbb{T}^r$**-manifold**.*

5. Proof of the Schur-Horn Theorem

After we've introduced all conceptions, we state the last theorem which is ingenious formally but its proof need deep symplectic geometry knowledge.

Theorem 5.1. *(Atiyah-Guillemin-Sternberg Convexity theorem) Suppose $(\mathcal{H}_\lambda, w_\lambda, \mathbb{T}^r, \mu)$ be a Hamiltonian $\mathbb{T}^r$-manifold.If M is compact and connected, then*

$\boxed{\mathcal{H}_\lambda \textit{ is a convex polyhedron in } \mathbb{R}^n \textit{ whose vertices are the images of the } \mathbb{T}^r\textit{-fixed points.}}$

Proof of Schur-Horn theorem:

- $(\mathcal{H}_\lambda, w_\lambda, \mathbb{T}^n, \mu)$ be a Hamiltonian $\mathbb{T}^n$-manifold.
- $\mathcal{H}_\lambda$ is compact:
  - $\mathcal{H}_\lambda$ is bounded by λ_1;
  - $\mathcal{H}_\lambda$ is closed. You can see $\mathcal{H}_\lambda$ as the zero set of some algebraic functions on $\mathcal{H}(n)$, or you can realize it as the orbit of the compact Lie groups $U(n)$,thus by the theory of Lie group's theory a closed set in $\mathcal{H}(n)$.
- $\mathcal{H}_\lambda$ is connected:for any $A \in \mathcal{H}_\lambda$, there exists $U \in U(n)$ such that $A = U\Lambda U^H$.

$$\begin{aligned}
&U(n) \text{ is connected}\\
\Rightarrow\ &\text{there exists } U_t : [0,1] \to U(n) \text{ such that } U_0 = I, U_1 = U\\
\Rightarrow\ &\text{there exists } A_t := U_t\Lambda U_t^H : [0,1] \to \mathcal{H}_\lambda \text{ such that } A_0 = \Lambda, A_1 = A\\
\Rightarrow\ &\mathcal{H}_\lambda \text{ is connected}
\end{aligned}$$

$\rightsquigarrow \pi(\mathcal{H}_\lambda)$ is a convex polyhedron in $\mathbb{R}^n$.

- For the $\mathbb{T}^n$-fixed points, we will find that they're just

$$\operatorname{diag}(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)}) \in \mathbb{R}^{n\times n} \qquad \text{where } \tau \in S_n$$

Now suppose $A = (a_{ij})_{i,j=1}^n \in \mathcal{H}_\lambda$.

- If $(\theta_1, \dots, \theta_n) \cdot A = A$ for any $(\theta_1, \dots, \theta_n) \in \mathbb{R}$, then

$$\Rightarrow \begin{pmatrix} e^{i\theta_1} & & 0\\ & \ddots & \\ 0 & & e^{i\theta_n}\end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix} e^{i\theta_1} & & 0\\ & \ddots & \\ 0 & & e^{i\theta_n}\end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} e^{i\theta_1}a_{11} & \cdots & e^{i\theta_1}a_{1n}\\ \vdots & \ddots & \vdots\\ e^{i\theta_n}a_{n1} & \cdots & e^{i\theta_n}a_{nn}\end{pmatrix}\begin{pmatrix} e^{i\theta_1}a_{11} & \cdots & e^{i\theta_n}a_{1n}\\ \vdots & \ddots & \vdots\\ e^{i\theta_1}a_{n1} & \cdots & e^{i\theta_n}a_{nn}\end{pmatrix}$$

$$\begin{aligned}
&\Rightarrow a_{ij} = 0 \text{ for any } i \neq j, \qquad A = \operatorname{diag}(a_{11}, \dots, a_{nn})\\
&\Rightarrow \operatorname{diag}(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)}) \qquad \text{where } \tau \in S_n
\end{aligned}$$

- On the other hand, if $A = \operatorname{diag}(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)})$ where $\tau \in S_n$, then

$$(\theta_1, \dots, \theta_n) \cdot A = A \qquad \text{for any } (\theta_1, \dots, \theta_n) \in \mathbb{R}$$

- In a word, all the $\mathbb{T}^n$-fixed points are

$$\begin{aligned}
\mathbb{T}^n_{fix} &= \{\operatorname{diag}(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)}) \in \mathcal{H}_\lambda \mid \tau \in S_n\}\\
\Rightarrow \pi(\mathbb{T}^n_{fix}) &= \{(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)})^T \in \mathbb{R}^n \mid \tau \in S_n\}
\end{aligned}$$

Thus by the AGM-convexity theorem,
$\pi(\mathcal{H}_\lambda)$ is a **convex polyhedron** in $\mathbb{R}^n$ whose **vertices** are

$$(\lambda_{\tau(1)}, \dots, \lambda_{\tau(n)})^T \in \mathbb{R}^n$$

where $\tau \in S_n$.

6. Miscellaneous

Using deeper results in symplectic geometry, one is able to prove more results in linear algebra. Take one for example:(see V. Guillemin, R. Sjamaar [3])

Denote the principal $k \times k$ minor of a matrix $A \in \mathcal{H}(n+1)$, denote the eigenvalues of A_k by $\mu_{1k}, \mu_{2k}, \dots, \mu_{kk}$, and assume that they are arranged in decreasing order: $\mu_{1k} \geqslant \mu_{2k} \geqslant \cdots \geqslant \mu_{kk}$.

Theorem 6.1. *(Gelfand-Cetlin) The μ_{ik}'s satisfy the interlacing conditions. Moreover, for every sequence of μ_{ik}'s satisfying these interlacing conditions*

$$
\begin{array}{ccccccccc}
\lambda_1 & & \lambda_2 & & \lambda_3 & \cdots\cdots\cdots & \lambda_{n+1} & & \\
& \geqslant & & \leqslant \quad \geqslant & & \leqslant \cdots\cdots\cdots & \leqslant & & \\
& & \mu_{1n} & & \mu_{2n} & \cdots\cdots\cdots & \mu_{nn} & & \\
& & & \geqslant & & \leqslant \cdots\cdots\cdots \leqslant & & & \\
& & & & \mu_{1(n-1)} & \cdots\cdots\cdots & \mu_{(n-1)(n-1)} & & \\
& & & & & \cdots\cdots\cdots & & & \\
& & & & & \cdots\cdots & & & \\
& & & & & \mu_{1k} \cdots\cdots \mu_{kk} & & &
\end{array}
$$

there exists a matrix $A \in \mathcal{H}_\lambda$, for which the eigenvalues of its k-th principal minor are $\mu_{1k}, \mu_{2k}, \dots, \mu_{kk}$.

At last, I want to thank for the Prof.Wang who gives me plenty of material and help me along the way to finish this task.

References

[1] Zuoqin Wang. "Introduction to Symplectic Geometry". 2015. url:http://staff.ustc.edu.cn/∼wangzuoq/Courses/15S-Symp/SympGeom.html

[2] Zuoqin Wang: Geometry of Moment Maps. 2016.

[3] V. Guillemin, R. Sjamaar: Convexity properties of Hamiltonian group actions. 2005. url:http://www.math.ornell.edu/∼sjamaar

[4] V. Guillemin and S. Sternberg: Convexity Properties of the Moment Mapping. Inventiones Mathematicae 67, pp. 491-513 (1982)

[5] Ana Rita Pires: Convexity in Symplectic Geometry: The Atiyah-Guillemin-Sternberg Theorem. 2007. url:http://faculty.fordham.edu/apissarrapires/convexitytalk.pdf

[6] A. Cannas da Silva, Lecture Notes on Symplectic Geometry. Lecture Notes in Mathematics 1764, Springer Verlag, Berlin (2001)

School of Mathematical Sciences, University of Science and Technology of China, Hefei, 230026, P.R. China,

Email address: `email:xx352229@mail.ustc.edu.cn`