arXiv:2302.05957v1 [math.MG] 12 Feb 2023

弱不变范数：斜厄米矩阵空间中球面的几何

GABRIEL LAROTONDA 和 IVAN REY

摘要。设 N 是斜厄米矩阵空间 $\mathfrak{u}_n(\mathbb{C})$ 上的弱酉不变范数（即在酉群的余伴随作用下不变）。本文研究此类范数的单位球面的几何，并说明其几何性质如何由矩阵特征值的主序性质所刻画。我们详细刻画给定范数下元素的范数化泛函，继而证明一个精确判据，用以判断交换子 $[X,[X,V]]$ 是否属于支撑单位球面上 V 的超平面。我们证明，$\mathfrak{u}_n(\mathbb{C})$ 在自身上的伴随作用 $V \mapsto V + [X,V]$ 会将向量推离单位球面。作为前述结果的应用，对于严格凸范数，我们证明当且仅当 X 与 V 对易时，此范数在这一伴随作用下保持不变。对于任意弱 Ad-不变范数，我们还给出了更详细的描述。

目录

1. 引言 1
2. Ad-不变芬斯勒范数 2
2.1. 凸体与范数 5
2.2. 球面的面、极点与对偶范数 10
2.3. 范数化泛函：细节 14
3. 的伴随作用于自身 22
3.1. 侧向导数 22
3.2. 耗散算子与不等式 24
3.3. 等号成立的情形 25
致谢 27
参考文献 27

1. 引言

复矩阵空间 $M_n(\mathbb{C})$ 中的 $n\times n$ 酉不变范数理论已相当成熟，围绕这一主题已有大量理论和数值研究。它在矩阵分析之外也有许多应用，例如量子信息论（参见 [17, 19, 26] 及其中的参考文献），以及通过框架和凸优化应用于信号处理 [1, 7, 12, 28]，以上列举的是最相关的应用。

日期：2023年2月14日。2020年数学主题分类。主要分类 15A60，52A21；次要分类 15B57，46B20。关键词与短语。伴随作用；凸集；芬斯勒范数；主序；范数化泛函；多胞体；斜厄米矩阵；支撑超平面；酉不变范数；弱不变范数。

然而，弱不变范数理论较少受到关注。由于这类范数是通过对酉群的余伴随作用保持不变来定义的，$\mathbf{U}_n = \mathbf{U}_n(\mathbb{C})$，我们认为其自然的研究背景应是该群的李代数，即斜伴随矩阵空间 $\mathfrak{u}_n = \mathfrak{u}_n(\mathbb{C})$。此外，鉴于范数与单位球之间存在对应关系，而单位球是凸集，我们认为应当先验地放宽范数必须完全齐次（等价地，单位球为平衡集）这一条件。如此便得到一个更大、更宽泛的正齐次且次可加范数族，我们在 $\mathfrak{u}_n$ 中称之为 Ad-不变芬斯勒范数。这一范数族与 $\mathbb{R}^n$ 的对称凸体有很好的对应关系，而单位球面的面则可借助主序和凸分析工具加以研究。

值得注意的是，对于酉矩阵U、V和 $X \in M_n$（$\mathbb{C}$），（强）酉不变性条件 $\|UXV\|_\phi = \|X\|_\phi$确保了 $\mathbb{R}^n$中（强）酉不变范数与对称规范函数之间的对应关系。此外，存在一族有限的范数（Ky-Fan 范数），能够控制两个给定矩阵的大小：对于任意酉不变范数，$\|X\|_\phi \leq \|Y\|_\phi$当且仅当 Ky-Fan 范数满足不等式 $\|X\|_{(k)} \leq \|Y\|_{(k)}$, $k = 1, \ldots, n$。另一方面，众所周知，不存在这样一族有限范数，能够针对任意给定的弱不变范数控制 X、Y 的大小。我们这里以轨道范数作为这种控制的良好替代；它们是 Ad-不变芬斯勒范数，其对偶单位球恰为 $\mathrm{conv}(\mathcal{O}_C)$，即零迹矩阵 $C \in \mathfrak{u}_n$的酉轨道的凸包。这些轨道范数是 [24] 中引入的 Hofer 范数的简化版本（并适用于矩阵情形）。由此可以完整描述给定范数的范数化泛函，单位球面的面（在若干重要情形下是一个多胞体）及其极值点也同样如此。随后，我们将这些描述推广，以获得任意芬斯勒范数的范数化泛函的性质（命题 2.31），并将其应用于本文的主要结果（定理 2.35）：我们刻画了满足如下条件的 X： $V + [X, [X, V]]$属于范数球面上支撑 V 的仿射超平面（推广了一个对弗罗贝尼乌斯范数显而易见的结果）。通过类似的证明，我们注意到，对于任意 $X \in \mathfrak{u}_n$和任意 Ad-不变芬斯勒范数，都有 $\|V + [X, V]\| \geq \|V\|$（这可以用 Birkhoff正交来理解，见注 3.10）：特别地，如果 $[X, V] \neq 0$，那么对于任意严格凸范数，不等式都严格成立。值得注意的是，这种 Birkhoff正交的概念（尽管是在厄米矩阵的情形下）可以为量子信息论中的 NPPT有界纠缠问题提供不同的视角（例如参见 [16, 第 4.4] 节及其中的参考文献）。

本文的组织结构如下：在第2.1 节中，我们介绍范数与 $\mathbb{R}^n$中的凸体之间的对应关系、与主序的关系，以及轨道范数族。在第2.2节中，通过考察对偶范数 N'，利用迹泛函和极性对偶，建立两种范数的单位球面之间的对应关系。因此，单位球面的面可以用单位范数泛函来描述，或者等价地，用对偶范数单位球面中的元素来描述 $\mathfrak{u}_n$（特别值得关注的是那些作为多面体的球面，例如迹范数的球面、谱范数的球面或轨道范数的球面）。在第2.3 节中，我们将精确刻画这些范数化泛函、球面的极点和光滑点。随后，我们证明结果（定理2.35），给出了不等式 $\varphi([X, [X, V]]) \leq 0$成立为等式的充要条件（这里 φ 是 V的范数化泛函）。作为所建立工具的应用，在第3

我们进一步详细研究 $V \mapsto V + [X, V]$ 在自身上的伴随作用：我们证明，该作用总是使向量远离单位球面，并且只有在严格条件下才保持范数不变。这些条件特别意味着，对于严格凸范数，我们有 $\|V + [X, V]\| \geq \|V\|$，且等号成立当且仅当 $[X, V] = 0$。

借助这些工具，酉群 $\mathbf{U}_n$ 配备由芬斯勒范数诱导的左不变度量后，其几何性质 N 将在后续论文中研究（目前正在准备中），相关思路源自[3, 4, 5, 6, 10, 23, 24]。

2. $\mathfrak{u}_n$中的 Ad-不变芬斯勒范数

在本文中，$\mathbf{U}_n$ 表示酉矩阵群 $\mathbf{U}_n = \mathbf{U}(n, \mathbb{C}) = \{U \in M_n(\mathbb{C}) : U^{-1} = U^*\}$，而 $\mathfrak{u}_n = \mathfrak{u}_n(\mathbb{C})$ 是李群 $\mathbf{U}_n$ 的李代数，即斜厄米矩阵构成的实线性空间，其实维数为 $dim_{\mathbb{R}}(\mathfrak{u}_n) = n^2$。我们还用 $\mathfrak{su}_n = \mathfrak{su}_n(\mathbb{C})$ 表示无迹斜厄米矩阵。我们将研究对群的余伴随作用不变的范数，事实上，我们还将讨论更广泛的一类芬斯勒范数：

定义 2.1（$\mathfrak{u}_n$上的 Ad-不变芬斯勒范数）。它是一个函数 $\|\cdot\| : \mathfrak{u}_n \to \mathbb{R}_{\geq 0}$，满足对任意$X, Y \in \mathfrak{u}_n$

(1) $\|X\| = 0$当且仅当 $X = 0$（非退化） (2) $\|X + Y\| \leq \|X\| + \|Y\|$（次可加） (3) $\|\lambda X\| = \lambda \|X\|$ 对任意 $\lambda \in \mathbb{R}_{\geq 0}$（正齐次） (4) $\|UXU^*\| = \|X\|$ 对任意 $U \in \mathbf{U}_n$（Ad-不变）。

注意，我们不要求光滑性，也不要求完全齐次。尽管芬斯勒范数不一定是范数，它会在 $\mathfrak{u}_n$ 上诱导一种拓扑，该拓扑由收敛的概念给出：$x_\alpha \longrightarrow x \Leftrightarrow \|x_\alpha - x\| \longrightarrow 0$。

定义 2.2. 在标准矩阵理论教材中，定义在 $M_n(\mathbb{C})$ 上的范数 $||\cdot||$，若对每个 $X \in M_n(\mathbb{C})$ 和 $U \in \mathbf{U}_n$，均满足 $||UXU^*|| = ||X||$，则称为弱不变；若对每个 $X \in M_n(\mathbb{C})$ 和 $U,\ V \in \mathbf{U}_n$，均满足 $||UXV|| = ||X||$，则称为强不变。

注 2.3（闵可夫斯基泛函）。设 $\|\cdot\|$ 是 $\mathfrak{u}_n$ 中的 Ad-不变芬斯勒范数，则对于 $r > 0$，集合 $B_r = \{V \in \mathfrak{u}_n : \|V\| < r\}$ 是开、吸收、凸且 Ad-不变的。反过来，若 B 是 $0 \in \mathfrak{u}_n$ 附近的一个吸收、凸且 Ad-不变集合，则其闵可夫斯基泛函 ρ_B 定义了 $\mathfrak{u}_n$ 中的 Ad-不变芬斯勒半范数（对于 $X \neq 0$，$\rho_B(X)$ 可能为零）。此外，若 B不包含从零出发的射线，则该闵可夫斯基泛函定义了一个 Ad-不变芬斯勒范数（$\rho(X) = 0$蕴含$X = 0$）。

注 2.4. 设 $ip \in \mathfrak{u}_n$ 为 p 一维正交投影。那么对于每个一维正交投影 q，都有 $q = Upu^*$，其中某个 $U \in \mathbf{U}_n$，因此 $\|iq\| = \|UipU^*\| = \|ip\|$；我们选择归一化 $\|ip\| = 1$。$\mathfrak{u}_n$中的所有 Ad-不变芬斯勒范数都是等价的，而且它们还与任意其他范数等价；特别地，它们都诱导相同的拓扑。$M_n(\mathbb{C})$中的每个弱不变范数通过限制，在 $\mathfrak{u}_n$中诱导出一个 Ad-不变范数。反过来，对于 $\mathfrak{u}_n$ 中的任意 Ad-不变范数，我们可以在 $M_n(\mathbb{C})$中构造一个（事实上是许多个）弱不变范数，使其在 $\mathfrak{u}_n$ 上的限制就是原范数，参见注 3.6。

例 2.5。 下面列出一些相关的 Ad-不变范数：

(1) $\|X\|_1 = \mathrm{Tr}\,|X|$（迹范数）。(2) $\|X\|_\infty = \max\{|\lambda| : \lambda \in \sigma(X)\}$（谱范数）。(3) 对于 $1 < p < \infty$，$\|X\|_p = (\mathrm{Tr}\,|X|^p)^{1/p}$（$p$-范数）。(4) $\|X\|^{(k)} = \sum_{1 \le i \le k} |\lambda_i|$，其中 λ_i 是 X 的特征值，按绝对值递减排列（Ky-Fan 范数）。(5) $\|X\|_C = \max\{|\mathrm{Tr}(CU^*XU)| : U \in \mathbf{U}_n\}$，对于 $C \in M_n(\mathbb{C})$，若其为非标量矩阵且 $\mathrm{Tr}(C) \neq 0$（C-半径范数）。

我们有 $\|\cdot\|_{KF(1)} = \|\cdot\|_\infty$、$\|\cdot\|_{KF(n)} = \|\cdot\|_1$，以及 $\|X\|_2 = \|X\|_F$（弗罗贝尼乌斯范数）。这些例子事实上全都是范数（完全齐次），并且除最后一个外，其余都是强酉不变的；关于这一主题，一个很好的参考资料是 Bhatia 的著作 [8, 第四章]。事实上，最后一个是范数的证明可见 [8, 命题 IV.4。4]。此外，众所周知，若 C 是任意一个一维投影，则 C-范数就是数值半径，因此它等于 $\mathfrak{u}_n$ 中的谱范数。

注 2.6. 对于 $A,\ B \in M_n(\mathbb{R})$，由下式给出的双线性映射

$$(A|B) := \mathrm{Tr}(AB^*)$$

定义了一个实内积，称为 Hilbert-Schmidt 内积。弗罗贝尼乌斯范数 $\|\cdot\|_F$ 是由该内积诱导的范数。事实上，弗罗贝尼乌斯范数及其正倍数，是定义在 $\mathfrak{u}_n$ 上的唯一由内积诱导的 Ad-不变范数。

注 2.7（迹对偶性，范数化泛函）。 我们将利用由 $M_n(\mathbb{C})$ 的迹所诱导的对偶性，将 $(\mathfrak{u}_n, \|\cdot\|)$ 的对偶空间与 $\mathfrak{u}_n$ 等同如下：对于 $\varphi \in \mathfrak{u}_n'$（$\mathfrak{u}_n$ 的对偶空间），考虑

$$\|\varphi\| = \max\{\varphi(X) : \|X\| \le 1\}$$

（不取模）；容易验证，这在对偶空间中定义了一个芬斯勒范数。对于 $V \in \mathfrak{u}_n$，定义

$$\varphi_V(X) = (V|X) = \mathrm{Tr}(VX^*) = -\mathrm{Tr}(VX), \tag{1}$$

则为 $\varphi_V \in \mathfrak{u}_n'$。由于 $(\cdot|\cdot)$ 是非退化的，每个非零的 $\varphi \in \mathfrak{u}_n'$ 都恰好以这种方式来自唯一的非零 $V \in \mathfrak{u}_n$。如果 $\varphi \in \mathfrak{u}_n'$ 是 $V \in \mathfrak{u}_n$ 的一个单位范数支撑泛函，即 $\varphi(V) = \|V\|$ 且 $\|\varphi\| = 1$，我们便称 $\varphi \in \mathfrak{u}_n'$ 为 $V \in \mathfrak{u}_n$ 的单位范数支撑泛函。由 Hahn-Banach 定理，每个非零向量 $V \in \mathfrak{u}_n$ 至少存在一个这样的泛函；为简便起见，我们称 φ 对于 V 是范数化的。注意，

$$\begin{aligned}(V|[X,Y]) &= -\mathrm{Tr}(V[X,Y]) = -\mathrm{Tr}(VXY - VYX) = -\mathrm{Tr}(XYV - XVY)\\ &= -\mathrm{Tr}(X(YV - VY)) = -\mathrm{Tr}(YVX - YXV) = -\mathrm{Tr}(Y(VX - XV))\end{aligned}$$

根据迹的循环性。因此，我们得到循环恒等式

$$(V|[X,Y]) = (Y|[V,X]) = (X|[Y,V]) \tag{2}$$

对任意 $X,\ Y,\ V \in \mathfrak{u}_n$。这些可以改写为 $(V|\mathrm{ad}\ X(Y)) = -(\mathrm{ad}\ X(V)|Y) = (\mathrm{ad}\ Y(V)\,|\,X)$。特别地，对于任意 $X \in \mathfrak{u}_n$，$\mathrm{ad}\,X : \mathfrak{u}_n \to \mathfrak{u}_n$ 在 Hilbert-Schmidt 内积下是斜伴随的

(这意味着 $-\mathrm{ad}^2(X) \geq 0$因此 $\|V + [X,[X,V]]\|_F \geq \|V\|_F$，关于一般化形式，请参见下文的定理 3.8。

2.1. **凸体与范数**。 接下来，我们陈述 Horn 和 Uhlmann 关于自伴矩阵 [20] 的轨道的凸包的一个定理版本；完全齐次范数的等价关系可见 Bhatia 的著作 [8, 定理 IV.4。7]首先，我们需要一些记号：

注 2.8 （记号）。 任意矩阵 $X \in \mathfrak{u}_n$ 都可写成 $X = UDU^*$，其中 $U \in \mathbf{U}_n$，而 D 是一个对角矩阵，其元素均为纯虚数。因此， X 的特征值与奇异值的绝对值相同，即 $|\lambda_i(X)| = |x_i| = \lambda_i(|X|)$。我们将 $\vec{x} = (x_1, x_2, \ldots\ldots, x_n)$记作这样一个实数列，其中 ix_j 是 X的特征值。对于 $B \subset \mathfrak{u}_n$，我们用 $\mathcal{O}(B)$表示B 在 $\mathfrak{u}_n$ 中的余伴随轨道（在酉群 $\mathbf{U}_n$的作用下），即

$$\mathcal{O}(B) = \{UXU^* : X \in B, U \in \mathbf{U}_n\}.$$

若 $C \in \mathfrak{u}_n$，我们简记 $\mathcal{O}_C = \mathcal{O}(\{C\})$。我们用 $\mathrm{conv}(\Sigma)$ 表示集合 Σ 的凸包。

采用这些记号后，现在可以陈述这些等价关系；关于球面的面，定义和讨论见第 2.2 节：

命题 2.9 （主序与芬斯勒范数）。 设 Z， $W \in \mathfrak{u}_n$。以下各条件等价： :

(1) $Z \in \mathrm{conv}(\mathcal{O}_W)$，更确切地说，存在至多 $n+1$ 个矩阵 $U_i \in \mathbf{U}_n$ 以及 $n+1$ 个实数 $\lambda_i \geq 0$，满足 $\sum_i \lambda_i = 1$，使得 $Z = \sum_{i=1}^{n+1} \lambda_i U_i W U_i^*$。(2) $\vec{z} \prec \vec{w}$（强主序）。(3) 对 $\mathfrak{u}_n$ 中所有 Ad-不变芬斯勒范数，$\|Z\| \leq \|W\|$。(4) 对所有 $C \in \mathfrak{u}_n$，$\max_{U \in \mathbf{U}_n}(UCU^*|Z) \leq \max_{U \in \mathbf{U}_n}(UCU^*|W)$。

此外，如果某个芬斯勒范数下等号成立，那么 Z 和所有$U_i W U_i^*$都位于该范数的球面的同一个面上（事实上，它们位于所有包含 Z 的面之交中）。如果该范数严格凸，那么对于某个 $U \in \mathbf{U}_n$， $Z = UWU^*$。

证明。 证明与 [24, 命题 5。8], 类似，只需作显而易见的修改，因此略去。

□

备注 2.10。 要得到这些等价关系，只需对上述条件 (3) 检验严格凸范数 即可。这是因为任意一个 Ad-不变芬斯勒范数 $\|\cdot\|$ 都可以通过 $\|x\|_\varepsilon = \|x\| + \varepsilon\|x\|_F$ 显式地逼近为严格凸的（Ad-不变芬斯勒）范数。另一方面，由于正则 $C \in \mathfrak{u}_n$ 构成稠密集，因此只需对其检验 (4)（也就是说， X 的所有特征值都不同）；此外，如果我们先验地知道 $\mathrm{Tr}(Z) = \mathrm{Tr}(W)$（这是一个必要条件），那么只需对正则 $C \in \mathfrak{su}_n$检验 (4)。

备注 2.11 （ $\mathfrak{su}_n$中的轨道范数）。 若 $X \in \mathfrak{su}_n$，则 $\sum_j x_j = -i\mathrm{Tr}(X) = 0$。这很容易推出，所有部分和 $\sum_{k=1}^m x_k$（其中 x_k 按递减顺序重新排列）都必须非负。因此，向量 $\vec{x}$ 在 $\mathbb{R}^n$中强支配零向量，即 $\vec{0} \prec \vec{x}$；根据前一个命题，对于任意 $X \in \mathfrak{su}_n$，都有 $0 \in \mathrm{conv}(\mathcal{O}_X)$。然后，对于固定的 $C \in \mathfrak{u}_n$

$$\max\{(UCU^*|X) : U \in \mathbf{U}_n\} = \max\{-\mathrm{Tr}(CU^*XU) : U \in \mathbf{U}_n\} \geq 0$$

对于任意 $X \in \mathfrak{su}_n$。对于满足 $C \neq \lambda 1$的固定 $C \in \mathfrak{u}_n$，我们将 $\mathfrak{su}_n$ 中的轨道范数 定义为

$$\|X\|_{\mathcal{O}_C} = \max\{(UCU^*|X) : U \in \mathbf{U}_n\}.$$

注意，若 $C = \lambda 1$，则对所有 $X \in \mathfrak{su}_n$，$\|X\|_{\mathcal{O}_C} = \mathrm{Tr}(X) = 0$，因此该条件是得到芬斯勒范数的必要条件。接下来我们证明它也是充分的。由于 $X \in \mathfrak{su}_n$，我们可以用 $C - \mathrm{Tr}(C/n)1$ 替换 C，并得到相同的量，因此我们将假设 $C \in \mathfrak{su}_n$。

注意，如果 C 是酉群在 $\mathfrak{u}_n$ 上的作用的不动点，那么 $-C$ 也是该作用的不动点。此外，不动点要么不存在，要么恰有两个，即 C 和 $-C$。

引理 2.12. 设 $0 \neq C \in \mathfrak{su}_n$。那么，轨道范数是 $\mathfrak{su}_n$ 上的 Ad-不变芬斯勒范数。此外：

(1) X 的范数在 UCU^* 处取得，该点与 X 对易，且二者的特征值均按递减顺序排列。(2) C 属于该范数的单位球面，当且仅当 $\|C\|_F = 1$。(3) 在这种情况下，当 $X \in \mathcal{O}_C$ 时，$\|X\|_{\mathcal{O}_C} = 1$；对于 $X \in \mathrm{conv}(\mathcal{O}_C)$，有 $\|X\|_{\mathcal{O}_C} \leq 1$，且等号成立当且仅当 $X \in \mathcal{O}_C$。(4) 该范数完全齐次，当且仅当 $\sigma(C) = -\sigma(C)$。

证明。根据前面的讨论，轨道范数非负，且显然是正齐次的。三角不等式成立且它是 Ad-不变的，这也显而易见。因此，我们需要检验它是否非退化。令 $\vec{c}$ 表示 $-iC$的特征值序列。由于 C 不是恒等映射的倍数，且 $\mathrm{Tr}(C) = 0$，众所周知，存在 $n-1$ 个置换 $\sigma_k \in S_n$，使得$\{\sigma_k(\vec{c})\}_{k=1,\ \ldots\ldots,\ n-1}$ 在 $\mathbb{R}^n$ 中的$\{v \in \mathbb{R}^n : \sum_{i=1}^n v_i = 0\}$内张成一个空间（例如参见 [18]）。因此，$C$ 的轨道张成 $\mathfrak{su}_n$。我们知道 $0 \in$ 属于 $\mathrm{conv}(\mathcal{O}_C)$，我们断言它是一个内部点。如果并非如此，我们使用 [9, 中的论证：根据哈恩-巴拿赫分离定理，存在一个单位范数泛函 φ，使得 $\varphi(0) = 0$ 且 $\varphi(\mathrm{conv}(\mathcal{O}_C)) \geq 0$。于是，对所有 $U \in \mathrm{U}_n$，都有 $\varphi(UCU^*) \geq 0$。令 dU为 U_n上的归一化哈尔测度，并令$W = \int_{\mathrm{U}_n} UCU^* dU$。那么，对所有$U \in \mathrm{U}_n$，$W \in \mathfrak{su}_n$ 且 $\mathrm{Ad}_U W = W$，而这只有在 $W = \lambda 1$时才可能，因此 $W = 0$。现在

$$0 = \varphi(W) = \int_{\mathbf{U}_n} \varphi(UCU^*) dU \geq 0$$

而这只有在对所有 U 都有 $\varphi(UCU^*) = 0$ 时才可能。但这意味着 $\mathrm{conv}(\mathcal{O}_C) \subset \ker \varphi$，矛盾。现在令 $X \neq 0$ 属于 $\mathfrak{su}_n$，缩小 X 后，我们得到 C 的轨道的凸包络内部中的一个元素，因此存在 $\lambda > 0$ 使得 $\lambda X = \sum_i \lambda_i U_i C U_i^*$。假设 $\|X\|_{\mathcal{O}_C} = 0$，那么对所有 $U \in \mathbf{U}_n$，都有 $(UCU^*|X) \leq 0$。但另一方面

$$0 < \lambda \|X\|_F^2 = \sum_i \lambda_i (U_i C U_i^*|X) \leq 0,$$

矛盾。因此必有 $\|X\|_{\mathcal{O}_C} \neq 0$，这就完成了证明： $\mathfrak{su}_n$中的轨道范数是一个真正的芬斯勒范数。令 $Z \in \mathfrak{su}_n$，取满足 $(UCU^*|X) = \|X\|_{\mathcal{O}_C}$ 的 U，则 $f(t) = (e^{tZ} UCU^* e^{-tZ}|X)$ 在 $t = 0$ 处取得最大值，因此

$$0 = f'(0) = ([Z, UCU^*]|X) = (Z|[UCU*, X])$$

根据（2），因此 $[UCU^*, X] = 0$，且我们可以同时将二者对角化，因此 $(UCU^*|X) = \sum_i c_{\sigma(i)} x_i$ 对某个置换 $\sigma \in S_n$ 成立。但任何这样的和都不超过将两组系数均按递减顺序排列后所得的和（例如参见 [8, 推论 II.44]）。这就证明了断言 1。对于归一化条件 2，注意到

$$-\operatorname{Tr}(UCU^*C) = \operatorname{Tr}(UCU^*C^*) \le \|UCU^*\|_F \|C\|_F = \|C\|_F^2 = -\operatorname{Tr}(C^2)$$

根据迹内积的柯西-施瓦茨不等式。因此 $\|C\|_{\mathcal{O}_C} \le -\operatorname{Tr}(C^2)$；另一方面，取 $U = 1$，我们也有 $\|C\|_{\mathcal{O}_C} \ge -\operatorname{Tr}(C^2)$。这表明 $\|C\|_{\mathcal{O}_C} = -\operatorname{Tr}(C^2) = \|C\|_F^2$。因此，令 $C \in \mathfrak{su}_n$ 固定，使得 $\|C\|_F = 1$。若 $X \in \mathcal{O}_C$，则根据范数的定义，我们有 $\|X\|_{\mathcal{O}_C} = \|C\|_{\mathcal{O}_C} = 1$；若 $X \in \operatorname{conv}(\mathcal{O}_C)$，显然 $\|X\|_{\mathcal{O}_C} \le 1$。假设 X 属于该轨道的凸包，且 X 的单位范数为 1，即 $X = \sum_i \lambda_i U_i C U_i^*$，并且存在 U_0，使得 $(U_0 C U_0^*|X) = \|X\|_{\mathcal{O}} = 1$。那么 $-CU_0^* X U_0 = \sum_i \lambda_i (-C)\, \tilde{U}_i C \tilde{U}_i^*$；对两边取迹并使用迹的柯西-施瓦茨不等式，可得

$$1 = (C|U_0^* X U_0) = \sum_i \lambda_i |(C|\tilde{U}_i C \tilde{U}_i^*) \le \sum_i \lambda_i \|C\|_F^2 = 1.$$

因此，对每个 i，必有 $(C|\tilde{U}_i C \tilde{U}_i^*) = \|C\|_F \|\tilde{U}_i C \tilde{U}_i^*\|_F$，且迹的柯西-施瓦茨不等式取等号，当且仅当存在 $t_i \ge 0$ 使得 $C = t_i \tilde{U}_i C \tilde{U}_i^*$。由对每个 i 都成立的等式，以及 $\|\tilde{U}_i C \tilde{U}_i^*\|_F = \|C\|_F = 1$，可知对每个 i 必有 $t_i = 1$。因此，C 与每个 i 的 $\tilde{U}_i$ 对易，且 $U_0^* X U_0 = \sum_i \lambda_i \tilde{U}_i C \tilde{U}_i^* = C$，这说明 $X = U_0 C U_0^* \in \mathcal{O}_C$，从而完成证明 3。现在假设 C 的谱是平衡的，那么若 $X \in \mathfrak{su}_n$，取 U 使得 $(UCU^*|X) = \|X\|_{\mathcal{O}_C}$，并取 U_0 使得 $U_0 C U_0^* = -C$。于是

$$\|-X\|_{\mathcal{O}_C} \ge (U U_0 C U_0^* U^*| - X) = (UCU^*|X) = \|X\|_{\mathcal{O}_C}.$$

结合 $-X$，我们得到相反的不等式。最后，如果 C 的谱不平衡，那么 $-C \notin \mathcal{O}_C$（否则 $-C = UCU^*$ 会导致矛盾）；我们假设 $\|-X\|_{\mathcal{O}_C} = \|X\|_{\mathcal{O}_C}$ 对任意 $X \in \mathfrak{su}_n$ 都成立，那么若 p 是任意一个一维投影，我们有 $c_1 = \|ip\|_{\mathcal{O}_C} = \|-ip\|_{\mathcal{O}_C} = -c_n$。现在设 p_1，p_2 是正交的一维投影，则根据本引理的断言 1，我们有

$$c_1 + c_2 = \|i(p_1 + p_2)\|_{\mathcal{O}_C} = \|-i(p_1 + p_2)\|_{\mathcal{O}_C} = -c_{n-1} - c_n,$$

因此 $c_2 = -c_{n-1}$。我们以这种方式继续，直到得到 $c_{k+1} = 0$（当 $n = 2k+1$ 为奇数时），或直到得到 $c_k = -c_{k+1}$（当 $n = 2k$ 为偶数时）。在这两种情况下，$\sigma(C) = -\sigma(C)$。 □

尽管轨道范数的自然定义域是 $\mathfrak{su}_n$，我们仍可以用若干种 Ad-不变的方式将它们扩展到 $\mathfrak{u}_n$，例如：

推论 2.13（$\mathfrak{u}_n$ 中的轨道范数）。设 $0 \neq C \in \mathfrak{su}_n$，则

$$\|X\|_{\mathcal{O}_C} = \max\{(UCU^*|X) : U \in \mathbf{U}_n\} + |\operatorname{Tr} X|$$

是 $\mathfrak{u}_n$ 中的 Ad-不变芬斯勒范数，当且仅当 C 的谱是平衡的时，它是完全齐次的。

证明。 显然，它是非负且正齐次的；三角不等式也很容易证明，不变性亦然。现在，若结果为零，由于两项均非负，必有 $X - \mathrm{Tr}(X/n)1 = 0$，并且根据前述引理，$\mathrm{Tr}\ X = 0$，因此 $X = 0$。最后一个断言可利用前述引理轻易证明，故略去。□

注 2.14. 令 $\vec{z}, \vec{w}$ 表示 $-iZ, -iW$ 的特征值序列，其中 $Z, W \in \mathfrak{u}_n$。结合命题 2.9 和引理 2.11，可得 $\|Z\| \leq \|W\|$ 对所有芬斯勒 Ad-不变范数成立，当且仅当 $\mathrm{Tr}(Z) = \mathrm{Tr}(W)$ 且

$$\sum_i c_i^{\downarrow} z_i^{\downarrow} \leq \sum_i c_i^{\downarrow} w_i^{\downarrow}$$

对于任意满足 $\sum_i c_i = 0$ 且所有 c_i 均不相同的序列 $\vec{c}$（向下箭头表示特征值按降序排列）。

现在我们讨论如何将芬斯勒范数表示为对称集合的余伴随轨道的闵可夫斯基范数：

注 2.15. 设 $\mathfrak{M} \subset \mathfrak{u}_n$ 是一个极大交换结合子代数（m.a.s.a.），则存在一个标准正交基 B，使得 $\mathfrak{M} = \{A \in \mathfrak{u}_n : A \text{ 在基 B 下为对角矩阵}\}$。所有这类代数都可通过酉矩阵共轭得到。对于任意给定的 $\mathfrak{M}$，考虑生成 $\mathfrak{M}$ 的 n 个一维投影 $(p_k)_{k=1}^n$。若将 ip_k 与 e_k（$\mathbb{R}^n$ 中的标准基向量）相对应，则与之前一样，可将任意对角矩阵 $A = i\sum a_k p_k \in \mathfrak{u}_n$ 与向量空间 $\mathbb{R}^n$ 中的向量 $\vec{a} = (a_1, \cdots, a_n)$ 相对应。赋予 $\mathfrak{u}_n$ 弗罗贝尼乌斯范数后，若在 $\mathbb{R}^n$ 上采用标准内积，则此对应为等距映射。

定义 2.16。 集合 B 围绕 $0 \in \mathbb{R}^n$，若对于每个置换 $\sigma \in S_n$ 以及 $(x_1, x_2, \cdots, x_n) \in B$，都有 $(x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(n)}) \in B$，则称其为对称的。若 $F\colon \mathbb{R}^n \to \mathbb{R}$，则当 $F \circ \sigma = F$ 时，称 F 为对称的，其中 σ 应理解为对变量的一个置换。注意，当且仅当 F 的等值集是对称的，F 才是对称的。

命题 2.17。 设 $\mathrm{k}\cdot\mathrm{k}$ 是 $\mathfrak{u}_n$ 上的一个 Ad-不变芬斯勒范数，且 B_1 是其开单位球。那么

(1) 对每个极大阿贝尔子代数 $\mathfrak{M} \subseteq \mathfrak{u}_n$，设 $B_{\mathfrak{M}} = B_1 \cap \mathfrak{M}$。那么 $B_{\mathfrak{M}}$ 是一个围绕 $0 \in \mathbb{R}^n$ 的对称开凸集（按照前一条注释中的识别），且与所选代数无关。(2) 设 $B \subset \mathbb{R}^n$ 是一个围绕 0 的开凸对称集，并且

$$B_1 = \mathrm{conv}(\mathcal{O}(B)) = \{\sum_i \lambda_i U_i X U_i^* : X \in B, U_i \in \mathbf{U}_n, \lambda_i \geq 0, \sum_i \lambda_i = 1\}.$$

那么，B_1 是一个包含 $0 \in \mathfrak{u}_n$ 的开 Ad-不变凸集，并且对于每个极大阿贝尔自伴子代数 $\mathfrak{M}$，都有 $B_1 \cap \mathfrak{M} = B$（这里使用了将 $\mathfrak{M}$ 与 $\mathbb{R}^n$ 相识别的方式）。

证明。 由于 B_1 是开集且为凸集，因此 $B_{\mathfrak{M}} = B_1 \cap \mathfrak{M} \subset \mathfrak{M} \simeq \mathbb{R}^n$ 是一个开凸集。设 $\mathfrak{M}_1$ 是另一个极大阿贝尔自伴子代数，则存在一个酉矩阵 U 使得 $U\mathfrak{M}U^* = \mathfrak{M}_1$。由于

$$B_{\mathfrak{M}_1} = B_1 \cap \mathfrak{M}_1 = UB_1U^* \cap U\mathfrak{M}U^* = U(B_1 \cap \mathfrak{M})U^* = UB_{\mathfrak{M}}U^*,$$

这表明，$B_{\mathfrak{M}}$ 和 $B_{\mathfrak{M}_1}$（视为 $\mathbb{R}^n$ 的子集时）是同一个集合。该集合显然是对称的，因为我们可以通过置换作用于 $\mathbf{U}_n$。这就证明了第一个断言。现在令 $B \subset \mathbb{R}^n$

为包含 0 的开凸对称集，并将其表示为某个极大阿贝尔自伴子代数的子集。根据 $\mathfrak{M}$的定义，我们看到它是开集、凸集、Ad-不变的，并且包含 0 矩阵。我们只需证明 B_1，因为这样对于任意给定的极大阿贝尔自伴子代数$B_1 \cap \mathfrak{M}_1 = B$，我们将有

$$B_1 \cap \mathfrak{M}_1 = U(B_1 \cap \mathfrak{M})U^* = UBU^*,$$

因此，借助我们与 $\mathbb{R}^n$ 的对应关系， B 和 UBU^* 实际上是同一个集合。令 $X \in B$，则对于所有 $U \in \mathrm{U_n}$ 都有 $UXU^* \in B_1$；若我们选择 $U = 1$，则可推出 $X \in B_1$。另一方面，我们选择 $\mathfrak{M}$ 使得 $B \subset \mathfrak{M}$，于是 $B \subset B_1 \cap \mathfrak{M}_1$。由于 B 是凸集，为证明 $B_1 \cap \mathfrak{M} \subset B$，我们需要证明：若 $Z = \sum \lambda_i U_i X U_i^* \in B_1 \cap \mathfrak{M}$，其中 $X \in B$、$U_i \in \mathrm{U_n}$、$\lambda_i \geq 0$ 且 $\sum \lambda_i = 1$，则 $Z \in B$。我们有 $Z = i\sum z_k p_k$。令 $X^i = diag(U_i X U_i^*) = \sum_k p_k U_i X U_i^* p_k$ 表示相对于给定的极大阿贝尔自伴子代数，$U_i X U_i^*$ 的对角部分。根据舒尔-霍恩定理，每个 X^i 都被 X 主化（见 [2, 定理 1。3]），或者等价地，存在一个双随机矩阵，使得 $\mathrm{diag}X^i = A^i \mathrm{diag}X$。现在，由于 Z 是对角的，

$$Z = i\sum z_k p_k = \sum_k p_k Z p_k = \sum_i \lambda_i \sum_k p_k U_i X U_i^* p_k = \sum_i \lambda_i X^i.$$

设 $A = \sum_i \lambda_i A^i$，则容易验证 A 是双随机矩阵。又由于 $\mathrm{diag}Z = \sum_i \lambda_i \mathrm{diag}X^i = \sum_i \lambda_i A^i \mathrm{diag}X = A\,\mathrm{diag}X$，我们得出 $\mathrm{diag}Z$ 被 $\mathrm{diag}X$ 主化。但这样一来，$\mathrm{diag}Z$ 就是 X 的 diag 坐标置换的凸组合，这表明 $Z \in B$。□

注 2.18。 若从 $B \subset \mathbb{R}^n$ 一个以 0 为中心的开对称集出发，则 B_1 是 $\mathfrak{u}_n$ 的一个开凸 Ad-不变集。因此，根据注 2.3，它在 $\mathfrak{u}_n$上定义了一个芬斯勒范数。由此事实和上一个命题可知， $\mathfrak{u}_n$ 上的每个 Ad-不变芬斯勒范数都由这种方式得到。显然，对 $\mathfrak{su}_n$也有相同结论，只需将 $\mathbb{R}^n$ 替换为超平面 $\sum_{i=1}^n x_i = 0$。

例 2.19。 这里给出几个简单的例子，稍后将用它们来说明光滑性与对偶性：

(1) 扭曲椭圆：考虑由下式给出的（扭曲）平面椭圆

$$\frac{1}{a^2}(x+y)^2 + \frac{1}{b^2}(x-y)^2 \leq 1.$$

该椭圆是一个对称凸体，其边界光滑（且严格凸）。根据命题2.17，该集合的酉轨道的凸包是 $\mathfrak{u}_n$中一个 Ad-不变芬斯勒范数的单位球，其中 $n = 2$。该凸体的闵可夫斯基泛函为

$$F(x,y) = \sqrt{\frac{1}{a^2}(x+y)^2 + \frac{1}{b^2}(x-y)^2},$$

它是 $\mathbb{R}^2$ 上的一个芬斯勒范数（因此对于 $n = 2$，它也是 $\mathfrak{u}_n$ 上的芬斯勒范数）：若 ix_k 是 X的特征值，则 $\|X\| = F(x_1, x_2)$。

(2) 吐司：考虑连接点 $(-1, 0)$ 和 $(0, 1)$ 的线段 $y = x + 1$，然后是连接点 $(0, -1)$ 和 $(1, 0)$ 的线段 $y = x - 1$，以及线段 $y = -x - 1$，连接

点 $(-1,\ 0)$ 和 $(0,\ -1)$。用圆周封闭此框

$$(x-\frac{1}{2})^2+(y-\frac{1}{2})^2=\frac{1}{2}, \tag{3}$$

连接点 $(1,\ 0)$ 和 $(0,\ 1)$。由这些曲线围成的区域是 $\mathbb{R}^2$ 中的一个对称凸体，即一个吐司。该凸体的范数公式（闵可夫斯基泛函）显而易见：如果 $(x,\ y)$ 属于第二、第三和第四象限，则它由 1-范数给出：$\|(x,\ y)\|=|x|+|y|$。另一方面，如果 $x,\ y\geq 0$，我们可以将方程（3）改写为 $x^2+y^2\leq x+y$。因此，在第一象限中，芬斯勒范数为

$$\|(x,y)\|=\frac{x^2+y^2}{x+y}.$$

圆周与线段（光滑地）连接处的点 $(1,\ 0)$ 是球面的一个极值点，但不存在只支撑该点的线性泛函。

2.2. **球面的面、极点和对偶范数**。设 B_1 为该范数的闭单位球。根据克莱因–米尔曼定理，它作为一个紧凸集，是其极点的凸包。

定义 2.20（极值点、面和由面生成的锥）。回顾一下，$0\neq X\in \mathfrak{u}_n$ 称为极值，当且仅当 $X/\mathrm{k}X\mathrm{k}$ 是 B_1 的极值点。注意，范数严格凸，当且仅当所有非零向量都是极值点。赋范空间 $(\mathfrak{u}_n,\mathrm{k}\cdot\mathrm{k})$ 的单位球 B_r 的一个面 F，是闭球 B_r 与由一个单位范数泛函 $\varphi\in\mathfrak{u}_n',\mathrm{k}\varphi\mathrm{k}=1$ 所确定的超平面的交集，即

$$F_\varphi(r)=\overline{B_r}\cap\{X\in\mathfrak{u}_n:\varphi(X)=r\}.$$

我们通常省略编号 r，并用 F_φ 指代包含某个向量 V 的面，即$r=\mathrm{k}V\mathrm{k}$。若 φ 是对偶空间中的极值元素，则称该面为极大 面。每个面都包含于某个极大面中：由于对偶范数的闭球是紧凸集，根据克莱因–米尔曼定理，存在单位范数的极值泛函 $\{\varphi_i\}_{i=1,\dots,n}$，使得 φ 是 φ_i 的凸组合

$$\varphi=\sum_i\lambda_i\varphi_i,\quad \lambda_i\geq 0,\quad \sum_i\lambda_i=1.$$

于是很容易验证，若 $\varphi(X)=\mathrm{k}X\mathrm{k}$，则对所有 i，都有 $\varphi_i(X)=\mathrm{k}X\mathrm{k}$。因此，若 $X\in F_\varphi$，则对所有 i，都有$X\in F_{\varphi_i}$；事实上，F_φ 是所有包含它的极大面的交集。由面 F_φ 生成的锥是 $C_\varphi=\mathbb{R}_+F_\varphi$。此锥恰好由所有满足以下条件的 $X\in\mathfrak{u}_n$ 组成：对于这个给定的单位范数 φ，$\varphi(X)=\mathrm{k}X\mathrm{k}$。

注意，对每个 $\lambda\in\mathbb{R}_{>0}$，若 $\varphi(V)=\mathrm{k}V\mathrm{k}$ 且 $\mathrm{k}\varphi\mathrm{k}=1$，则有 $\varphi(\lambda V)=\lambda\mathrm{k}V\mathrm{k}=\mathrm{k}\lambda V\mathrm{k}$，因此 φ 范数化了整条射线 $\lambda V,\lambda>0$。若 $\mathrm{k}\varphi\mathrm{k}=1$ 且 $\varphi(X)=\mathrm{k}X\mathrm{k}$，我们称 φ 为 X 的范数化泛函。

定义 2.21 (对偶范数)。引理 2.7 中的对偶性还允许我们定义 $V\in\mathfrak{u}_n$ 的

$$\|V\|'=\max\{(V|X):\|X\|\leq 1\},$$

容易验证，这个范数也是 Ad-不变的，并且如果 $V\in\mathfrak{u}_n$ 且 φ 是由下式定义的泛函 V 如（1）所示，那么 $\|\varphi\|=\|V\|'$。

注 2.22. 如果 $\|V\| = 1$，对于每个 V的范数化 φ，我们都有 $N \in \mathfrak{u}_n$具有单位对偶范数，即 $\|N\|' = 1$，使得 $\varphi(V) = (N|V) = 1$。但如果 $\varphi' = (V|\cdot)$，容易验证，由于 $\|V\| = 1$，这个泛函在双对偶空间中具有单位范数。另一方面，$\varphi'(N) = (V|N) = \|V\| = 1 = \|N\|'$，因此$\varphi'$ 对偶空间中的范数 N。令 $F_{\varphi'} \subset \mathfrak{u}_n$ 表示对偶范数的球面的面，它由泛函 φ'给出。显然，当且仅当 $F_{\varphi'} = \{N\}$ 时，$\varphi = (N|\cdot)$ 是支撑 V的唯一泛函；并且，当且仅当 $\varphi' = (V|\cdot)$ 是对偶空间中支撑 N的唯一泛函时，$F_\varphi = \{V\}$。

注 2.23（凸性与光滑性）。 根据定义，二重对偶范数等于原范数，这是显然的。另一方面，以下性质将很有用：

（1）范数 $\|\cdot\|$ 在 $x \neq 0$ 处盖托可微，当且仅当它是弗雷歇可微的（这由 Šmulian 引理推出，参见 [11, 引理 8。4] 例如）。因此，当对任意 $x \neq 0$都成立时，我们简单地称该范数是光滑的；在这种情况下，范数函数实际上在 C^1之外 $x = 0$ [11,推论 8。5]。（2）范数是光滑的，当且仅当对偶范数是严格凸的，当且仅当存在唯一的泛函支撑每个 $x \neq 0$ [11, 引理 8.4 和事实 8。12]。（3）范数是严格凸的，当且仅当它是一致凸的（这可由上述性质以及二重对偶范数等于原范数这一事实推出，也可以直接由单位球面的紧性证明）。（4）如果范数在 $X \neq 0$处光滑，则范数 F 在 X 处（相对于迹内积）的梯度返回 $N \in \mathfrak{u}_n$，使得 $\varphi = (N|\cdot)$）是 X的唯一范数化泛函。这一点可由下式看出：

$$t^{-1}(\|X + tV\| - \|X\|) \leq \|V\| \qquad \forall\, t > 0,$$

这意味着对于所有 V，$(\nabla F(X)\,|V) \leq \|V\|$，且 $(\nabla F(X)\,|X) = \|X\|$。特别地，当范数光滑时，$X \mapsto \nabla F(X)$ 是从 ∂B_1 到 $\partial B_1'$ 的双射。

(5) 无论如何，由于范数是凸函数，$X \neq 0$ 处范数的次微分完全刻画了 X 的范数化泛函（见 [30, 第 2] 节和下文定理 3.3）。

另请参见 [15]，了解巴拿赫空间中球体的更多一般性质。

注释 2.24（极性对偶）。 设 E 是内积空间 $(V, (\cdot|\cdot))$ 中的一个凸子集。E 的极对偶是集合

$$E^\circ = \{Y \in V : (Y|X) \leq 1 \quad \forall\, X \in E\}.$$

显然，$E \subseteq (E^\circ)^\circ$；事实上，我们有（$E^\circ$）$^\circ$ 等于 $E \cup \{0\}$的凸包的闭包。如果 E 是闭的且 $0 \in E$，由于我们假设 E 是凸的，根据 Hahn-Banach 定理，我们得到 $E = (E^\circ)^\circ$。进一步讨论请参见 [14] 及其中的参考文献。设 $E = B_1$ 是 $\mathfrak{u}_n$中某个芬斯勒范数 $\|\cdot\|$ 的单位球，设 $B = B_1 \cap \mathfrak{M} \subset \mathbb{R}^n$ 是生成 B_1 的 n维凸集（命题 2.17）。显然，当且仅当边界集 ∂B 在 $\mathbb{R}^n$ 中严格凸时，范数 $\|\cdot\|$ 才是严格凸的（详情参见例如 [31, 定理 3.1]）。然后，根据对偶性并利用命题 2.17，我们有 $\mathfrak{u}_n$ 中的范数光滑，当且仅当 ∂B 在 $\mathbb{R}^n$中光滑。

定义 2.25 （自对偶范数）。 根据 $\mathfrak{u}_n$中对偶范数的定义，如果 B_1' 是对偶范数的单位球，那么关于 $\mathfrak{u}_n$中的迹内积，我们有 $B_1' = B_1^\circ \subset \mathfrak{u}_n$。此外，根据同一命题， $B_1' = \mathrm{conv}(B^\circ)$，其中我们取 $B \subset \mathbb{R}^n$ 关于该处标准内积的极对偶。不难验证，围绕 0 的唯一自极闭凸体是欧几里得范数的单位球。但由于我们关注的是球的几何性质，因此可以允许旋转和膨胀/收缩。因此，如果该范数的单位球的极对偶与原单位球相似于某个旋转，我们就称该范数为自对偶 。

例 2.26. 现在我们讨论一些对偶范数及其单位球面的几何性质。根据前面的注记，只需研究将芬斯勒范数的单位球 B_1 与任意一个极大阿贝尔自伴子代数相交后得到的 n 维凸集 $B \subset \mathbb{R}^n$ 的极对偶。作为凸集，它们在取内部和闭包时表现良好，因此我们稍微滥用一下记号，用 B, B_1 也表示这些球的闭包。

(1) 众所周知， p-范数的对偶是满足 $q = 1 - 1/p$ 的 q 范数，其中包括 $q = \infty$ 的情形，此时 $p = 1$ ，反之亦然。对于 $1 < p < \infty$，唯一的 $\varphi = (N|\cdot)$ 范数化$X = i\sum_k x_k p_k$ 由下式给出

$$N = \frac{1}{\|X\|_p^{p-1}} \sum_k sg(x_k) x_k^{p-1}$$

(2) 唯一的自极对偶范数是弗罗贝尼乌斯范数（$p = 2$）。1-范数和 ∞-范数仅在 $n = 2$ 时自对偶，而 p-范数在 $p \neq 1, 2$ 时并非自对偶。
(3) 根据备注 2.23.4，直接计算可知，例 2.19 中扭曲椭圆的对偶范数的单位球面是另一个扭曲椭圆，其半径现在为 $2/a$, $2/b$。这些范数是光滑且一致凸的。

(4) 例 2.19 中吐司的对偶单位球也可用类似方法得到：对直对角线段取对偶后，我们得到三个点 $(-1, 1)$、$(-1, -1)$、$(1, -1)$；第一个点与第二个点之间以垂直线段连接，第二个点与第三个点之间以水平线段连接。接着，$(-1, 1)$ 必须通过圆周的极对偶与 $(1, -1)$ 相连；计算其梯度可知，这是一条由 $\alpha(t) = 2(\cos t + \sin t + \sqrt{2})^{-1}(\cos t, \sin t)$ 在区间 $t \in [-\pi/4, 3\pi/4]$ 内参数化的曲线（它不是圆周）。

(5) Ky-Fan 范数的对偶范数 k由下式给出：$\|X\|'_{(k)} = \max\left\{\frac{1}{k}\|X\|_1, \|X\|_\infty\right\}$，如[29, 定理 1]所示。Watson 在 $\|\cdot\|_{(k)}$ 中利用范数的次梯度，完整刻画了 Ky-Fan 范数的范数化泛函（备注 2.23.5）。

(6) 我们指出 Ky-Fan 范数 k 的一个特殊范数化泛函：令 $X = i\sum_j x_j p_j \in \mathfrak{u}_n$，并记 $X = \Omega|X| = |X|\Omega$为 X 的极分解。通过取 Ω 为酉算子，我们可以假设

$$\Omega = i\sum_j sg(x_j) p_j \qquad |X| = \sum_j |x_j| p_j$$

其中 $sg(x_j) = 1$当 $x_j \geq 0$且 $sg(x_j) = -1$当 $x_j < 0$。于是，令 U为一个酉矩阵，将 $|X|$的奇异值按递减顺序排列，即 $U|X|U^* = \sum_j |x_j|^{\downarrow} p_j$，并令

$P = \sum_{j=1}^{k} p_j$。考虑 $\varphi = (N|\cdot)$，其中 $N = -i\mathrm{Re}(U^*PU\Omega^*) \in \mathfrak{u}_n$。容易验证，$\varphi(X) = \|X\|_{(k)}$。另一方面，对于任意 p 和任意矩阵 Z，均有 $\|\mathrm{Re}Z\|_p \le \|Z\|_p$，因此

$$\|N\|_1 \le \|U^*PU\Omega^*\|_1 = \|P\|_1 = k \qquad \|N\|_\infty \le \|U^*PU\Omega^*\|_\infty = \|P\|_\infty = 1.$$

由前面的注记可知，$\|N\|'_{(k)} \le 1$，因此 φ 是 X 的范数化函数。

由于轨道范数具有许多有趣的性质（这些性质有助于我们直观理解关于任意范数的进一步一般性结论，见下一节），我们将其性质作为一个单独的例子来介绍。自然的背景是 $\mathfrak{su}_n$，因此我们将讨论范围限制于此：

例 2.27（轨道范数及其对偶）。 C 在 $\mathfrak{su}_n$ 中的轨道范数的对偶，是轨道的凸包的范数，即 $B_1^\circ = B_1' = \mathrm{conv}(\mathcal{O}_C)$（这从定义中显而易见）。等价地， $B_1 = \mathrm{conv}(\mathcal{O}_C)^\circ$，特别是，最大值可以在轨道或轨道的凸包上取到：

$$\|X\|_{\mathcal{O}_C} = \max_{V \in \mathcal{O}_C} (X|V) = \max_{Z \in \mathrm{conv}(\mathcal{O}_C)} (X|Z) \quad \text{for any } X \in \mathfrak{su}_n .$$

还可得到另一条性质：

（1）根据舒尔-霍恩定理， $B_1' \cap \mathfrak{M}$ 的极值点是通过对 $\vec{c}$ 进行排列置换得到的向量（与前文一样， $\mathfrak{M}$ 是任意一个极大阿贝尔自伴子代数）。

(2) 对偶范数不一定是轨道范数（等价地，$\|\cdot\|_{\mathcal{O}_C}$ 的单位球不一定是某个酉轨道的凸包）。下面给出一个简单的例子：令 $C = \mathrm{diag}(1, 0, -1)$，则 C 在 $\mathbb{R}^3$ 的超平面 $x + y + z = 0$ 中的轨道的凸包是一个正六边形；这个凸包就是 $B_1' \cap \mathfrak{M}$，即对偶范数的单位球与极大阿贝尔自伴子代数的截面。容易看出，六边形各边的法向量是向量 $(1, 1, -2)$，$(2, -1, -1)$ 及其排列。因此，轨道范数 $\|\cdot\|_{\mathcal{O}_C}$ 的单位球 B_1 是由这六个点生成的正六边形，但无法将这六个点构造成任意给定点的轨道。

(3) 如果 C 像前一个例子那样是正则且平衡的，那么轨道是正 n 边形，因此其极对偶也是正 n 边形。所以在这种情况下，该范数是自对偶的。

(4) 若 C 是正则的但不平衡，则得到非正则的 n 边形（且该范数不是自对偶的）。

(5) 若 C 不是正则的，则对于 $n!$ 的某些因数，可得到 k 边形。例如，若 $C = (1, 1, -2)$，则 C 的轨道生成一个正三角形，因此其极对偶（轨道范数的单位球）也是正三角形；故该范数是自对偶的。

(6) 采用归一化 $\|C\|_F = 1$，我们有 $\mathrm{conv}(\mathcal{O}_C) \subset B_1$（引理 2.12），因此 $B_1' \subset B_1$，于是 $\|X\|'_{\mathcal{O}_C} \ge \|X\|_{\mathcal{O}_C}$ 对任意 $X \in \mathfrak{su}_n$。

(7) 在此归一化下，有 $\mathcal{O}_C = \partial B_1 \cap \partial B_1^\circ$。为证明这一点，回想对偶范数的单位球是 $\mathrm{conv}(\mathcal{O}_C)$，因此我们也有 $\|C\|'_{\mathcal{O}_C} \le 1$，从而 $\|C\|'_{\mathcal{O}_C} = 1$。这说明轨道 $\mathcal{O}_C$ 位于单位球面与该球的极对偶的单位球面的交集中。现在，如果 $\|X\|_{\mathcal{O}_C} = \|X\|'_{\mathcal{O}_C} = 1$，特别是 $\|X\|'_{\mathcal{O}_C} \le 1$，因此 $X \in \mathrm{conv}(\mathcal{O}_C)$。由于 $\|X\|_{\mathcal{O}_C} = 1$，根据引理 2.12.3，我们有 $X \in \mathcal{O}_C$。

(8) 在上述归一化下，对于任意 $U \in \mathbf{U}_n$，我们有 $\varphi = (UCU^*|\cdot)$ 是轨道范数 $\mathrm{k}\cdot\mathrm{k}_{\mathcal{O}_C}$ 的一个单位范数泛函，因为 $\mathrm{k}UCU^*\mathrm{k}' = \mathrm{k}C\mathrm{k}' = 1$。对于任意 $X \neq 0$，存在 U_0 使得 $(U_0CU_0^*|X) = \mathrm{k}\,X\mathrm{k}_{\mathcal{O}_C}$，因此 $\varphi = (U_0CU_0^*|\cdot)$ 对 X 是范数化的。

2.3. **范数化泛函：细节**。现在我们来讨论表示固定芬斯勒范数的范数化泛函的矩阵的一些更相关的性质。

引理 2.28. 设 $\mathrm{k}\cdot\mathrm{k}$ 是 $\mathfrak{u}_n$ 上的一个 Ad-不变芬斯勒范数，设 $0 \neq X \in \mathfrak{u}_n$，设 $\varphi = (N|\cdot)$ 是 X 或 $-X$ 的一个范数化泛函。对于 $Y \in \mathfrak{u}_n$，我们有

(1) $\varphi([X,Y]) = 0$(等价地，$\varphi \circ \mathrm{ad}\ X \equiv 0$ 在 $\mathfrak{u}_n$ 上)，且 N 与 X 对易。(2) 若 φ 是 X 的范数化泛函，则 $\varphi([Y,[Y,X]]) \leq 0$。

证明。考虑在 $\mathfrak{u}_n$ 中的展开式

$$\mathrm{Ad}_{e^{sY}} X = e^{s\,\mathrm{ad}\,Y}X = X + s\ \mathrm{ad}\,Y(X) + \frac{s^2}{2}[Y,[Y,X]] + o(s^3), \tag{4}$$

这首先给出：$e^s \mathrm{ad}^Y X - X = s[Y,X] + o(s^2)$。接着，对每个 X 的范数化泛函 φ

$$\varphi(e^{s\,\mathrm{ad}\,Y}X) - \varphi(X) \leq \|e^{s\,\mathrm{ad}\,Y}X\| - \|X\| = \|e^{sY}Xe^{-sY}\| - \|X\| = 0.$$

两边除以 $s > 0$ 并令 $s \to 0^+$，得到 $\varphi([Y,X]) \leq 0$。将 Y 替换为 $-Y$，可证明 $\varphi([Y,X]) = 0$。当 $-\varphi(X) = \mathrm{k}-X\mathrm{k}$ 时，证明类似，故略去。于是，对所有 $Z \in \mathfrak{u}_n$，我们有

$$0 = \varphi([X,Z]) = -\operatorname{Tr}(N[X,Z]) = (N|[X,Z]) = (Z|[N,X])$$

此处我们使用了循环恒等式 (2)。取 $Z = [N,X]$，得到 $\|[N,X]\|_F^2 = 0$，因此 $[N,X] = 0$。现在由 (4)

$$o(s^3) + \frac{s^2}{2}\varphi([Y,[Y,X]]) + 0 + \varphi(x) = \varphi(e^{s\,\mathrm{ad}\,Y}X) \leq \|e^{s\,\mathrm{ad}\,Y}X\| = \|X\| = \varphi(X),$$

因此 $o(s^3) + \frac{s^2}{2}\varphi([Y,[Y,X]]) \leq 0$。两边除以 s^2 并令 $s \to 0$，即可证明第二个断言。 □

注 2.29。由于 $g\colon s \mapsto \varphi(e^s \mathrm{ad}^Y X)$ 在 $s = 0$ 处取得最大值，因此一种可能是所有导数在 $s = 0$ 处均为零；这只有在 $\varphi(e^s \mathrm{ad}_Y) \equiv g(0) = \mathrm{k}\,X\mathrm{k}$ 时才会发生，因为 g 是实解析函数。若非如此，则表达式 $\varphi((\mathrm{ad}\ Y)^k(X)) = 0$ 的各项直至某个有限的 偶数 次项都为零，而该项必须严格为负，即 $\varphi((\mathrm{ad}\ Y)^{2n}(X)) < 0$，且此前的各次幂均为零。

定义 2.30 （对角/余对角分解）。设 $0 \neq V \in \mathfrak{u}_n$，写作 $V = i\,\acute{\mathrm{I}}_{k\in F} v_k P_k \,\mathrm{with} v_1 > v_2 > \cdots > v_F$，其中 V 的对角化使特征值各不相同且按降序排列。每个 P_k 是投影到相应特征子空间上的正交投影，并且 $P_k P_j = 0 \,\mathrm{for}\ k \neq j$。对于任意 $X \in \mathfrak{u}_n$，我们可以将其分解为 $X = X_D + X_C$，其中

$$X_D = \sum_k P_k X P_k \quad \text{and} \quad X_C = \sum_{k\neq j} P_k X P_j.$$

X 的分块对角和分块余对角部分。需要提醒的是，很容易出现 $[X_D, X_C] \neq 0$。不过请注意，

$$X = X_D + X_C, \qquad [X,V] = [X_C,V] \quad \text{and} \quad [X_D,V] = 0.$$

设 $\varphi = (N|\cdot) = -\mathrm{Tr}(N\cdot)$ 是 V 的范数化泛函。由于 N 与 V 对易（引理2.28），我们有 $N = i\sum_{k\in F} N_k$，其中 N_k 是自伴的，且 $N_k N_j = 0$ 适用于 $k \neq j$（N 为分块对角矩阵）。接下来，我们将轨道范数情形下所述的范数化泛函刻画（引理2.12.1）推广到任意 Ad-不变芬斯勒范数：

命题 2.31（范数化泛函的刻画）。设 $0 \neq V = i\sum_{k\in F} v_k P_k$，并令 $N = i\sum_k N_k$，其中 $\varphi = (N|\cdot)$ 是 V 的一个范数化泛函，令 F_φ 为由 φ 给出的面。那么

(1) 若 $j > k$，则 N_k 的任一特征值都大于或等于 N_j 的所有特征值。(2) 若其中一对特征值相等，则置换相应的一维投影可得到另一个 $V \in F_\varphi$。(3) 若 $F_\varphi = \{V\}$，则第一条中的不等式实际上总是严格的。(4) 对于给定的 $s \in F$，交换 N_s 中对应于不同特征值的一维投影，可得到另一个使 V 范数化的 ϕ。(5) 若 φ 是使 V 范数化的唯一泛函，则 N 是对角的，即 $N_k = \sum_{k\in F} n_k P_k$，且 $n_{k+1} \geq n_k$。

特别地，如果该范数严格凸且光滑，那么它必定满足 $N = \sum_{k\in F} n_k P_k$，且 n_k 严格递减。

证明。我们有 $\|V\| = \varphi(V) = (N|V) = -\mathrm{Tr}(NV) = -\mathrm{Tr}(i\sum_s N_s i\sum_l v_l P_l) = \sum_s v_s \mathrm{Tr}(N_s)$。现在若 $U \in \mathbf{U}_n$，则

$$-\mathrm{Tr}(U^*NUV) = -\mathrm{Tr}(NUVU^*) = \varphi(UVU^*) \leq \|UVU^*\| = \|V\| = -\mathrm{Tr}(NV). \tag{5}$$

还要注意，若 $\tilde{V} = UVU^*$，则 $\varphi(\tilde{V}) \leq \|\tilde{V}\| = \|V\|$。现在选取 $\mathrm{Ran}(P_s)$ 的一组标准正交基 $\{e^s_1, \ldots, e^s_{d(s)}\}$，使每个 N_s 都对角化（这里 $d(s) = dim(P_s)$），即我们有

$$N_s e^s_l = \lambda_l(N_s) e^s_l \qquad l = 1, \ldots, d(s)$$

对每个 $s \in F$ 都成立；我们还可以进一步假设 $\lambda_l(N_s)$ 按非递增顺序排列。由于 $\sum_s P_s = 1$，若按给定顺序将这些基合在一起，就得到一组同时对角化 N 和 V 的基。令 $U \in \mathbf{U}_n$ 为交换 $e^k_{d(k)}$ 与 e^j_1 的酉矩阵（即将 N_k 的最后一个特征向量与 N_j 的第一个特征向量交换），其中 $j > k$。则

$$\|V\| = \varphi(V) = -\mathrm{Tr}(NV) = \sum_{s\in F} v_s \sum_{l=1}^{d(s)} \lambda_l(N_s),$$

而 $-\mathrm{Tr}(U^*NUV) = -\mathrm{Tr}(NUVU^*)$ 则是相同的求和，只是将所提到的两个特征值互换。利用（5），消去所有其余相同的项后，我们得到

$$v_j \lambda_{d(k)}(N_k) + v_k \lambda_1(N_j) \leq v_k \lambda_{d(k)}(N_k) + v_j \lambda_1(N_j).$$

由此我们得到 $(v_k - v_j)\,(\lambda_{d(k)}(N_k) - \lambda_1(N_j)) \geq 0$，且由于当 $j > k$ 时 $v_k > v_j$，第一个断言的结论便成立。注意，若 $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$，那么实际上我们有

$$\varphi(\tilde{V}) = -\mathrm{Tr}(NUVU^*) = -\mathrm{Tr}(U^*NUV) = -\mathrm{Tr}(NV) = \varphi(V) = \|V\| = \|UVU^*\| = \|\tilde{V}\|,$$

因此 $\tilde{V} \in F_\varphi$，第二个断言得证。若此面是一个单点集，且等式 $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$ 成立，那么 $UVU^* = \tilde{V} = V$；但这是不可能的，因为 $v_k \neq v_j$。因此，在这种情况下，我们必有严格不等式，这就证明了第三个断言。现在，若我们通过酉算子 U 置换 N_k 的两个不同特征投影，那么 $\mathrm{Ran}(UN_kU^*) = \mathrm{Ran}(N_k) = \mathrm{Ran}(P_k)$，并且 U 在其他块上是恒等算子，因此若 $\phi = (UNU^*|\cdot)$，我们有

$$\begin{aligned}\phi(V) &= \sum_{s\neq k} v_s \operatorname{Tr}(N_s) + \sum_l v_l \operatorname{Tr}(UN_kU^*P_l) = \sum_{s\neq k} v_s \operatorname{Tr}(N_s) + v_k \operatorname{Tr}(UN_kU^*P_k)\\ &= \sum_{s\neq k} v_s \operatorname{Tr}(N_s) + v_k \operatorname{Tr}(N_k) = (N|V) = \varphi(V) = \|V\|.\end{aligned}$$

另一方面，对于任意 $Z \in \mathfrak{u}_n$

$$\phi(Z) = -\operatorname{Tr}(UNU^*Z) = -\operatorname{Tr}(NU^*ZU) = (N|UZU^*) \leq \|N\|'\|UZU^*\| = \|\varphi\|\|Z\| = \|Z\|,$$

结合前一个等式可知，$\|\phi\| = 1$ 和 ϕ 是 V的定范泛函。最后，如果恰有一个泛函能为 V定范，那么每个 N_s的特征值都必须相等，因此对每个 s，$N_s = n_sP_s$。□

注2.32。由于 N与 V对易，并且原则上可以在每个块 N_k内任意重排特征向量，因此必要时我们可以假设 N的特征值均按递减顺序排列。另一方面，显然，如果 $\lambda_{d(k)}(N_k)$ 是 N_k的最小特征值，那么

$$N_k \geq \lambda_{d(k)}(N_k)P_k.$$

同样，如果 $\lambda_1(N_j)$ 是 N_j的最大特征值，那么 $N_j \leq \lambda_1(N_j)P_j$；我们最近还证明了，当 $j > k$ 时，$\lambda_{d(k)}(N_k) \geq \lambda_1(N_j)$。这里的两个不等式均相对于矩阵偏序而言，

$$X \geq Y \quad \text{iff} \quad \langle X\xi, \xi\rangle \geq \langle Y\xi, \xi\rangle \quad \forall\, \xi \in \mathbb{C}^n$$

注2.33。如果 $[X,[X,Y]] = 0$或 $[Y,[Y,X]] = 0$，那么 $[X,Y] = 0$，因为

$$0 = (Y|0) = (Y|[X,[X,Y]) = ([Y,X]|[X,Y]) = -\|[X,Y]\|_F^2,$$

于是 $[X,Y] = 0$。对于弗罗贝尼乌斯范数（唯一源自内积的 Ad-不变范数），使 $0 \neq V \in \mathfrak{u}_n$ 范数化的唯一范数化泛函为 $\varphi = (\ ^V/_{\|V\|_F}|\cdot)$，因此 $N_V = \ ^V/_{\|V\|_F}$，故由前面的计算，在弗罗贝尼乌斯范数的情形下

$$\varphi_V([X,[X,V]]) = 0$$

只有当 $[X,V] = 0$ 时才可能发生。接下来我们将探讨其他 Ad-不变范数下的这一条件。特别地，我们将证明，对任意严格凸范数，前述结论都成立。

注2.34。若 $X = X_D + X_C$ 是 X 相对于 $V = i\sum_k v_kP_k$分解为对角部分和余对角部分的结果，并且 $v_1 > v_2 > \cdots > v_F$与前文相同，则 $[X,V] = [X_C,V]$，因为 X_D对易

其中 V（V的每个块都是单位子矩阵的倍数）。那么，若 $\varphi=(N|\cdot)$范数 V，则由雅可比恒等式和（2）可得

$$\begin{aligned}\varphi([X,[X,V]]) &= \varphi([X,[X_C,V]]) = (N\,|\,[[X,X_C],V]+[X_C,[X,V]])\\ &= -(N\,|\,[V,[X,X_C]])+(N\,|\,[X_C,[X_C,V]])\\ &= -([N,V]\,|\,[X,X_C])+([N,X_C]\,|\,[X_C,V]) = ([N,X_C]\,|\,[X_C,V])\end{aligned}$$

由于 N 与 V 对易，即

$$\varphi([X,[X,V]]) = ([N,X_C]\,|\,[X_C,V]). \tag{6}$$

定理 2.35 (等号成立的情形)。设 $X,V\in\mathfrak{u}_n$，写作 $V=i\sum_k v_kP_k$，其中 $v_1>\cdots>v_F$，并令 $\varphi=(N|\cdot)$为 V 的任意范数化泛函。那么，$\varphi([X,[X,V]])=0$ 当且仅当 $[X_C,N]=0$。此外，若 N 关于 V 是对角的（特别是当 φ 是使 V 范数化的唯一泛函时），则$[X,N]=0$。另一方面，若 $F_\varphi=\{V\}$，则 $[X,V]=0$。

证明。假设 $[X_C,N]=0$，则由 (6) 可得 $\varphi([X,[X,V]])=([N,X_C]\,|\,[X_C,V])=0$。现在假设 $\varphi([X,[X,V]])=0$，先取 $Z=P_jXP_k+P_kXP_j\in\mathfrak{u}_n$，并注意到 $Z=Z_C$ 当 $k\neq j$ 时。然后由命题 2.28 以及恒等式 (6)，可得

$$\begin{aligned}\operatorname{Tr}([Z,N][Z,V]) &= -\operatorname{Tr}([N,Z][Z,V]) = ([N,Z]\,|\,[Z,V])\\ &= ([N,Z_C]\,|\,[Z_C,V]) = \varphi([Z,[Z,V]])\le 0.\end{aligned} \tag{7}$$

现在我们来计算

$$[Z,N] = i(P_jXP_k+P_kXP_j)\sum_s N_s - i\sum_s N_s(P_jXP_k+P_kXP_j) \tag{8}$$

$$= i(P_jXN_k+P_kXN_j-N_jXP_k-N_kXP_j). \tag{9}$$

同样

$$[Z,V] = i(P_jXv_kP_k+P_kXv_jP_j-v_jP_jXP_k-v_kP_kXP_j) = i(v_k-v_j)(P_jXP_k-P_kXP_j).$$

Then

$$\begin{aligned}\operatorname{Tr}([Z,N][Z,V]) &= -(v_k-v_j)\operatorname{Tr}((P_jXN_k+P_kXN_j-N_jXP_k-N_kXP_j)(P_jXP_k-P_kXP_j))\\ &= -(v_k-v_j)\operatorname{Tr}(-P_jXN_kXP_j+P_kXN_jXP_k+N_jXP_kXP_j-N_kXP_jXP_k)\\ &= -2(v_k-v_j)\operatorname{Tr}(N_jXP_kX-N_kXP_jX).\end{aligned}$$

因此，由（7）可得

$$2(v_k-v_j)\operatorname{Tr}(N_jXP_kX-N_kXP_jX)\ge 0 \qquad \forall\, k\neq j. \tag{10}$$

现在我们来计算

$$\begin{aligned}[X_C, N] &= i\sum_{r\neq s}(P_rXP_s + P_sXP_r)\sum_l N_l - i\sum_l N_l\sum_{r\neq s}(P_rXP_s + P_sXP_r)\\ &= i\sum_{r\neq s}(P_rXN_s + P_sXN_r) - i\sum_{r\neq s}(N_rXP_s + N_sXP_r)\\ &= i\sum_{r\neq s} P_rXN_s + P_sXN_r - N_rXP_s - N_sXP_r.\end{aligned}$$

同样,

$$\begin{aligned}[X_C, V] &= i(\sum_{j\neq j} P_kXv_jP_j + P_jXv_kP_k - v_kP_kXP_j - v_jP_jXP_k)\\ &= i\sum_{k\neq j}(v_k - v_j)(P_jXP_k - P_kXP_j).\end{aligned}$$

那么 $\varphi([X,[X,V]]) = \mathrm{Tr}([X_C,N][X_C,V])$ 等于

$$\begin{aligned}&= -\sum_{j\neq k}\sum_{m\neq n, r\neq s}\mathrm{Tr}((P_rXN_s + P_sXN_r - N_rXP_s - N_sXP_r)(v_k - v_j)(P_jXP_k - P_kXP_j))\\ &= -4\sum_{k\neq j}(v_k - v_j)\,\mathrm{Tr}(N_jXP_kX - N_kXP_jX).\end{aligned}$$

根据不等式 (10)，只有当 $\mathrm{Tr}(N_jXP_kX - N_kXP_jX) = 0$对所有 $k \neq j$时，它才可能为零。于是，假设该和为零，并首先考虑 $j > k$的情形。由引理 10， $\lambda_{d(k)}(N_k) \geq \lambda_1(N_j)$。我们将证明这些数实际上相等，并且

$$P_kXP_jN = P_kXN_j = \lambda_1(N_j)P_kXP_j, \qquad P_jXP_kN = P_jXN_k = \lambda_1(N_j)P_jXP_k. \tag{11}$$

取伴随后，我们还将得到 $NP_jXP_k = \lambda_1(N_j)P_jXP_k$, $NP_kXP_j = N_kXP_j = \lambda_1(N_j)P_kXP_j$，因此对于 $k > j$，我们得到

$$P_jXP_kN = P_jXN_k = \lambda_1(N_j)P_jXP_k, \quad NP_jXP_k = N_jXP_k = \lambda_1(N_j)P_jXP_k,$$

即 $[P_jXP_k, N] = 0$。根据对称性，同样的论证也适用于 $j < k$(此时用 $\lambda_1(N_k)$ 替代 $\lambda_1(N_j)$)，于是我们就证明了定理 $[X_C, N] = 0$ 所述的结论。

为简化计算中的符号变化，方便起见，我们将 $\mathrm{e}\,\mathrm{e}\,\mathrm{e}X = iX$ 写作 $X^* = X$，实际上还会省略波浪号，这样做稍微滥用了记号（注意， X_Ce 与 N 对易，当且仅当 X_C 与 N 对易）。因此，假设 $j > k$，并从数据 $\mathrm{Tr}(N_jXP_kX) = \mathrm{Tr}(N_kXP_jX)$ 开始，其中 $X^* = X$ 自伴。于是，由于 $N_k \geq \lambda_{d(k)}(N_k)P_k$ 和 $\lambda_{d(k)} \geq \lambda_1(N_j)$，我们得到

$$P_jXN_kXP_j \geq \lambda_{d(k)}(N_k)P_jXP_kXP_j \geq \lambda_1(N_j)P_jXP_kXP_j,$$

其中，在第二个不等式中，我们利用了 $P_jXP_kXP_j = \|P_kXP_j\|_F^2 \geq 0$。同样地，利用 $N_j \leq \lambda_1(N_j)P_j$我们得到

$$P_kXN_jXP_k \leq \lambda_1(N_j)P_kXP_jXP_k = \lambda_1(N_j)\|P_jXP_k\|_F^2.$$

取迹并利用假设可得

$$\lambda_1(N_j)\,\|P_kXP_j\|_F^2 \le \lambda_{d(k)}(N_k)\,\|P_kXP_j\|_F^2 \le \mathrm{Tr}(P_jXN_kX) = \mathrm{Tr}(P_kXN_jX) \le \lambda_1(N_j)\,\|P_jXP_k\|_F^2。$$

由于 $\|P_jXP_k\|_F^2 = \|P_kXP_j\|_F^2 = \mathrm{Tr}(P_kXP_jX)$，事实上我们有以下等式

$$\lambda_{d(k)}(N_k)\,\mathrm{Tr}(P_kXP_jX) = \mathrm{Tr}(P_jXN_kX) = \mathrm{Tr}(P_kXN_jX) = \lambda_1(N_j)\,\mathrm{Tr}(P_kXP_jX)$$

若 $\|P_kXP_j\|_F = 0$，则 $P_kXP_j = P_jXP_k = 0$ 且（11）显然成立。若不然，则必有 $\lambda_{d(k)}$（N_k）$=\lambda_1$（N_j）$=\lambda$。此外，注意到

$$A = P_jXN_kXP_j - \lambda P_jXP_kXP_j \ge 0$$

但我们刚刚证明了 $\mathrm{Tr}(A) = 0$。因此 $A = 0$，也就是 $P_jXN_kXP_j = \lambda P_jXP_kXP_j$。同样，由于 $P_kXN_jXP_k \le \lambda P_kXP_jXP_k$，根据得到的迹相等关系，它们必定相等，因此我们有

$$P_jXN_kXP_j = \lambda P_jXP_kXP_j \quad \text{and} \quad P_kXN_jXP_k = \lambda P_kXP_jXP_k. \tag{12}$$

首先假设 $\lambda \ge 0$，令 $N_j = N_j^+ - N_j^-$，其中 N_j^+，$N_j^- \ge 0$ 是将 N_j 分解为 N_j 的正部和负部的哈恩分解。由于 $N_j \le \lambda P_j$，必有 $0 \le N_j^+ \le \lambda P_j$，并且$0 \le (N_j^+)^2 \le \lambda^2 P_j$。现在

$$\lambda P_kXP_jXP_k = P_kXN_jXP_k = P_kXN_j^+XP_k - P_kXN_j^-XP_k \le \lambda P_kXP_jXP_k - P_kXN_j^-XP_k,$$

因此 $0 \le P_kXN_j^-XP_k \le 0$，而这只有在 $P_kXN_j^-XP_k = 0$ 时才有可能。由于 $P_kXN_j^-XP_k = |N_j^-\overline{XP_k}|^2$，可知必定有 $N_j^-XP_k = 0 = P_kXN_j^-$。现在

$$\|P_kXN_j^+\|_F^2 = \mathrm{Tr}(P_kX(N_j^+)^2XP_k) \le \lambda^2\,\mathrm{Tr}(P_kXP_jXP_k)$$

因此 $\|P_kXN_j^+\|_F \le \lambda\|P_jXP_k\|_F$。由此可知，以下柯西-施瓦茨不等式取等号：

$$\begin{aligned}\lambda\|P_jXP_k\|_F^2 &= \mathrm{Tr}(P_kXN_jXP_k) = \mathrm{Tr}(P_kXN_j^+XP_k) = (P_kXN_j^+|P_kXPj)\\ &\le \|P_kXN_j^+\|_F\|P_kXPj\|_F \le \lambda\|P_kXPj\|_F^2.\end{aligned}$$

但是，柯西-施瓦茨不等式取等号当且仅当向量共线[22]，，即$P_kXN_j^+ = \mu P_kXPj$，在右侧乘以 P_jXP_k后，我们很容易看出 $\mu = \lambda$。因此

$$P_kXN_j = P_kXN_j^+ - P_kXN_j^- = P_kXN_j^+ = \lambda P_kXP_j,$$

这就是 (11) 中所述的第一个恒等式。为得到第二个恒等式，注意到 $N_k \ge \lambda P_k \ge 0$ 于是 $\sqrt{N_k} \ge \sqrt{\lambda}P_k \ge 0$，进而 $P_jX\sqrt{\lambda}P_kXP_j \le P_jX\sqrt{N_k}XP_j$。另一方面，利用 (12)，我们有

$$\|P_jX\sqrt{N_k}\|_F^2 = \mathrm{Tr}(P_jX\sqrt{N_k}\sqrt{N_k}XP_j) = \mathrm{Tr}(P_jXN_kXP_j) = \lambda\|P_kXP_j\|_F^2 = \lambda\|P_jXP_k\|_F^2,$$

再次，我们在柯西-施瓦茨不等式中得到等式：

$$\begin{aligned}\sqrt{\lambda}\|P_kXP_j\|_F^2 &= \sqrt{\lambda}\,\mathrm{Tr}(P_jXP_kXP_j) \le \mathrm{Tr}(P_jX\sqrt{N_k}XP_j)\\ &= (P_jX\sqrt{N_k}|P_jXP_k) \le \|P_jX\sqrt{N_k}\|_F\|P_jXP_k\|_F \le \sqrt{\lambda}\|P_jXP_k\|_F^2.\end{aligned}$$

在这种情况下，我们得到 $P_jX\sqrt{N_k} = \sqrt{\lambda}P_jXP_k$，然后

$$P_jXN_k = P_jX\sqrt{N_k}\sqrt{N_k} = \sqrt{\lambda}P_jXP_k\sqrt{N_k} = \sqrt{\lambda}P_jX\sqrt{N_k} = \lambda P_jXP_k,$$

这就是（11）中所声称的第二个恒等式。由此解决了 $\lambda \geq 0$ 的情形。对于 $\lambda < 0$ 的情形，可以用类似的方法处理，但需要做一些修改：首先注意到 $N_j \leq \lambda P_j < 0$，因此 $-N_j \geq 0$，此外 $\sqrt{-N_j} \geq \sqrt{-\lambda}P_j \geq 0$。于是 $P_kX\sqrt{-N_j}XP_k \geq P_kXP_j\sqrt{-\lambda}XP_k$，并且

$$\begin{aligned}\sqrt{-\lambda}\|P_jXP_k\|_F^2 &\leq \operatorname{Tr}(P_kX\sqrt{-N_j}XP_k) = (P_kX\sqrt{-N_j}|P_kXP_j)\\ &\leq \|P_kX\sqrt{-N_j}\|_F\|P_kXP_j\|_F \leq \sqrt{\operatorname{Tr}(P_kX(-N_j)XP_k)}\|P_kXP_j\|_F\\ &= \sqrt{-\lambda\operatorname{Tr}(P_kXP_jXP_k)}\|P_kXP_j\|_F = \sqrt{-\lambda}\|P_jXP_k\|_F^2.\end{aligned}$$

再次，对于迹内积的柯西-施瓦茨不等式，我们得到一个等式；只有当 $P_kX\sqrt{-N_j} = \sqrt{-\lambda}P_kXP_j$ 时这才可能成立。像之前一样论证（在右侧乘以 $\sqrt{-N_j}$），我们得到 $P_kXN_j = \lambda P_kXP_j$，这就证明了（11）中的第一个恒等式。现在注意到，由于 $N_k \geq \lambda P_k$ 且 $\lambda < 0$，因此若 $N_k = N_k^+ - N_k^-$ 是 N_k 的哈恩分解，则必有 $-N_k^- \geq \lambda P_k$ 以及 $(N_k^-)^2 = (-N_k^-)^2 \leq \lambda^2 P_k$。按照前一种情形的论证，我们首先看到 $P_jXN_k^+XP_j = 0$，即 $P_jXN_k^+ = N_k^+XP_j = 0$。于是 $P_jXN_k = P_jXN_k^-$ 并且

$$\|P_jXN_k\|_F^2 = \operatorname{Tr}(P_jXN_k^2XP_j) = \operatorname{Tr}(P_jX(N_k^-)^2XP_j) \leq \lambda^2\operatorname{Tr}(P_jXP_kXP_j) = \lambda^2\|P_kXP_j\|_F^2.$$

最后一次讨论柯西-施瓦茨不等式。

$$\begin{aligned}-\lambda\|P_kXP_j\|_F^2 &= \operatorname{Tr}(-P_jXN_kXP_j) = (-P_jXN_k|P_jXP_k) \leq \|P_jXN_k\|_F\|P_kXP_j\|_F\\ &\leq |\lambda|\,\|P_kXP_j\|_F^2 = -\lambda\|P_kXP_j\|_F^2,\end{aligned}$$

这就得到 $P_jXN_k = \lambda P_jXP_k$，即式 (11) 中的第二个恒等式。既然现在我们知道 $[X_C, N] = 0$，回顾命题 2.31：当只有一个 φ 使 V达到范数时，N 是对角的，即 $N = i\sum_k n_kP_k$。然而，X 相对于 V的对角部分，此时也相对于 N是对角的，或者等价地，$[X_D, N] = 0$。因此在这种情况下，$[X, N] = [X_C, N] + [X_D, N] = 0+0 = 0$。另一方面，如果 $F_\varphi = \{V\}$，那么根据同一命题，等式 $\lambda_{d(k)}(N_k) = \lambda_1(N_j)$ 不可能成立，因此每一步都必有 $P_kXP_j = 0$，这表明 $X_C = 0$，因而 $X = X_D$，或者等价地，X与 V对易。□

注 2.36（严格凸范数）。如果该范数是严格凸范数，那么球面的每个面都是单点集，因此特别地，$\varphi([X,[X,Y]]) = 0$蕴含 $[X,Y] = 0$当且仅当 φ支撑 X。

注记 2.37。如果对于某个非零的 X，V，有 $\varphi([X,[X,V]]) = 0$，且 φ使 V达到范数，我们可以将 V写成半径为 $\|V\|$ 的球面的极点的凸组合，即 $V = \sum_i \lambda_iV_i$，其中 $0 < \lambda_i \leq 1$且 $\sum_i \lambda_i = 1$。由于这些 V_i必定都位于与 V相同的面上，因此泛函 φ使每个 V_i都达到范数。于是我们有

$$0 = \varphi([X,[X,V]]) = \sum_i \lambda_i\varphi([X,[X,V_i]]).$$

必有 $\varphi([X,[X,V_i]]) = 0$对于每个 V_i（根据引理 2.28，所有项均非正）。用类似的推理，设 $\varphi = (N|\cdot) = \sum_i \varphi_i = \sum_i (N^i|\cdot)$的 φ_i极值，则必有 $\varphi_i([X,[X,V]]) = 0$。

引理 2.38。 设 $V = i\acute{\mathrm{I}}_k v_k P_k \in \mathfrak{u}_n$，且 $v_1 > v_2 > \cdots > v_F$。存在一个 $\varphi_0 = (N_0|\cdot)$ 范数，使 V 满足范数条件，且 N_0为对角矩阵，即 $N_0 = i\acute{\mathrm{I}}_{k\in F} \lambda_k P_k$。设 $X \in \mathfrak{u}_n$，则 $\varphi_0(\,[X,[X,V]]) = 0$ 当且仅当$[X, N_0] = 0$。

证明。 设 $\varphi = (N|\cdot)$ 是 V 的任意一个范数化泛函，则 $N = i\acute{\mathrm{I}}_{k\in F} N_k$，其中 $N_k^* = N_k = N_k P_k = P_k$ 对所有 $k \in F$ 均成立。与前面一样，选取 $\mathrm{Ran}(P_k)$ 的一组标准正交基 $\{e_1^k, \ldots, e_{d(k)}^k\}$，使每个 N_k 对角化（此处 $d(k) = dim(P_k)$），即有

$$N_k e_l^k = \lambda_l(N_k) e_l^k \qquad l = 1, \ldots, d(k)$$

对每个 $k \in F$ 均成立；按给定顺序将这些基合在一起，得到一组同时对角化 N 和 V 的基。现在，对每个 $k \in F$ 以及每个 $l, s \in F_k = \{1, \ldots, d(k)\}$，令 $U_{k,ls}$ 为交换 e_l^k 与 e_s^k 的酉矩阵。显然

$$U_{k,ls} P_j U_{k,ls}^* = P_j \qquad \forall\, j, k \in F,$$

因此 $U_{k,ls} V U_{k,ls}^* = V$。注意，对每个 k，从每个 F_k 中选取 $l \neq s$ 恰有 $C(k) = \binom{d(k)}{2}$ 种方式，并将其基数记为 $|F|$。令

$$N_0 = \frac{1}{|F|} \sum_{k\in F} \frac{1}{C(k)} \sum_{l\neq s\in F_k} U_{k,ls} N U_{k,ls}^*.$$

我们声称，这个 N_0 具备所有所需的性质。首先注意，通过重新排列每个 N_k中的对角元素，每个新块都是一个对角块，其对角线上是 N_k 的特征值的算术平均值，即 $N_0 = i\acute{\mathrm{I}}_k \lambda_k P_k$，其中

$$\lambda_k = \frac{1}{d(k)} \sum_{l=1}^{d(k)} \lambda_l(N_k).$$

现在，注意

$$\|N_0\|' \leq \frac{1}{|F|} \sum_{k\in F} \frac{1}{C(k)} \sum_{l\neq s\in F_k} \|N\|' = \frac{1}{|F|} |F| \frac{1}{C(k)} C(k) = 1$$

并且

$$(N_0|V) = \frac{1}{|F|} \sum_{k\in F} \frac{1}{C(k)} \sum_{l\neq s\in F_k} (N|U_{k,ls}^* V U_{k,ls}) = \frac{1}{|F|} \sum_{k\in F} \frac{1}{C(k)} \sum_{l\neq s\in F_k} \|V\| = \|V\|,$$

因此 $\mathrm{k}\, N_0 \mathrm{k}' = 1$，且 $\varphi_0 = (N_0|\cdot)$ 范数 V。若 $X \in \mathfrak{u}_n$，则有

$$\varphi_0([X,[X,V]]) = (N_0|[X,[X,V]]) = ([N_0, X]|[X,V])$$

由此可清楚看出 $[X, N_0] = 0$ 意味着这个数为 0。另一方面，根据定理2.35，若这个数为 0，则 $[X_C, N] = 0$，但由于 N_0 的每个块都是对角块，我们也有$[X_D, N] = 0$。□

注 2.39. 设 N norm X；由于 N_k 的特征值大于或等于 N_{k+1}的特征值，它们的算术平均值也如此。因此，使用前述引理中的记号，我们有 $\lambda_k \geq \lambda_{k+1}$，且等号 $\lambda = \lambda_k = \lambda_{k+1}$成立当且仅当它们全都相等。也就是说，当且仅当对于任意 $l \in \{1, \ldots, d(k)\}$, $s \in \{1, \ldots, d(k+1)\}$, $\lambda_l(N_k) = \lambda_s(N_{k+1}) = \lambda$。这说明必有 $N_k + N_{k+1} = \lambda(P_k + P_{k+1})$。

另一方面，如果并非如此，那么 $\lambda_k > \lambda_{k+1}$，并且按照定理 2.35的证明过程，在这种情况下 $\varphi_0([X,[X,V]]) = 0$推出 $P_k X P_{k+1} = 0$。

3. $\mathfrak{u}_n$ 对自身的伴随作用

本节给出定理 2.35 的一个应用。首先，我们刻画（芬斯勒）范数函数的单侧方向导数。

3.1. **侧向导数**。

注 3.1。 设 $h:\ \mathbb{R} \to \mathbb{R}$ 是连续凸函数，则 $\lim_{t\to 0^+} \frac{h(x+t)-h(x)}{t}$ 存在（即它是一个实数或 $-\infty$）；此外，该商关于 $t \to 0^+$递减：由凸性可知，若 $0 < s < t$，则

$$\frac{h(x+s)-h(x)}{s} \leq \frac{h(x+t)-h(t)}{t}, \tag{13}$$

这说明了变量 $t \to 0^+$上的单调性，由此可明显看出关于极限的结论。现在，给定 X, Y，在芬斯勒赋范空间 $(E, \|\cdot\|)$ 中，且 $X \neq 0$，令 $h(t) = \|X+tY\|$，并注意到这是一个连续凸函数。因此，极限

$$\lim_{t\to 0^+} \frac{\|X+tY\|-\|X\|}{t}$$

存在；由于商 $\frac{\|X+tY\|-\|X\|}{t}$有数 $-\|Y\|$ 作为下界，因此极限是有限数。交换 t和 $-t$，我们看到 $t^{-1}(\|X+tY\| - \|X\|)$ 随着$t \to 0^-$递增，而且极限也存在；此外，对于任意 $t > 0$，都有

$$2\|X\| = \|X+tY+X-tY\| \leq \|X+tY\| + \|X-tY\|$$

这意味着 $\|X\| - \|X+tY\| \leq \|X-tY\| - \|X\|$。将两边除以 $-t$，可得

$$\lim_{t\to 0^+} \frac{\|X+tY\|-\|X\|}{t} \geq \lim_{t\to 0^+} \frac{\|X-tY\|-\|X\|}{-t} = \lim_{t\to 0^-} \frac{\|X+tY\|-\|X\|}{t}. \tag{14}$$

定义 3.2 （范数化泛函）。 对于 $0 \neq X \in \mathfrak{u}_n$，记作

$$N_X = \{\varphi \in \mathfrak{u}_n' : \|\varphi\| = 1 \quad \text{and} \quad \varphi(X) = \|X\|\}.$$

如前所述，这是对偶空间中的一个非空紧凸集。此外，如果 $\varphi \in N_X$ 那么对于任意 $\varphi \in N_{\lambda X}$，也有 $\lambda > 0$ （参见定义2.20）。

以下结果对于范数（即完全齐次的范数）是众所周知的。由于我们未能找到关于芬斯勒范数的合适参考文献，因此给出证明：

命题 3.3。 设 $(E, \|\cdot\|)$ 是一个芬斯勒赋范空间，且 $0 \neq X \in E$，那么对任意 $Y \in E$

$$\lim_{t\to 0^+} \frac{\|X+tY\|-\|X\|}{t} = \max_{\varphi\in N_X} \varphi(Y).$$

证明。 使用辅助映射 $h(t) = \|X+tY\|$ 以及记号 $h'(0^+)$ 表示我们正在计算的极限，将会很方便。我们还将用 $h'(0^-)$ 表示 $t \to 0^-$ 时的极限。首先注意，对每个 $\varphi \in N_X$，我们都有

$$\frac{h(t)-h(0)}{t} = \frac{\|X+tY\|-\|X\|}{t} \geq \frac{\varphi(X+tY)-\varphi(X)}{t} = \varphi(Y).$$

因此 $h'(0^+) \geq \max_{\varphi\in N_X} \varphi(Y)$。下面证明，至少存在一个 $\varphi \in N_X$ 使等式成立，这样便完成了证明。令 $\psi : \mathrm{span}\{x,y\} \longrightarrow \mathbb{R}$ 为由 $\psi(aX+bY) = a\|X\| + bh'(0^+)$（对 $a,b \in \mathbb{R}$）给出的线性泛函。注意，$\psi(X) = \|X\|$，且 $\psi(Y) = h'(0^+)$。我们断言 $\psi(aX+bY) \leq \|aX+bY\|$；当 $a=0$ 或 $b=0$ 时，这显然成立。现在回忆一下，根据（13），h 是递减函数。首先，假设 $a>0$ 且 $b>0$，则

$$\begin{aligned} a\|X\| + bh'(0^+) &= a\|X\| + b\lim_{t\to 0^+}\frac{\|X+tY\|-\|X\|}{t} \leq a\|X\| + b\frac{\|X+\frac{b}{a}Y\|-\|X\|}{\frac{b}{a}} \\ &= a\|X\| + a\|X+\frac{b}{a}Y\| - a\|X\| = \|aX+bY\|. \end{aligned}$$

如果 $a<0$ 且 $b<0$，请回忆右极限小于或等于左极限（14），且左极限是递增的；由于 $b<0$ 我们得到

$$\begin{aligned} a\|X\| + bh'(0^+) &\leq a\|X\| + bh'(0^-) = a\|X\| + b\lim_{t\to 0^-}\frac{\|X+tY\|-\|X\|}{t} \leq a\|X\| + b\frac{\|X-Y\|-\|X\|}{-1} \\ &= a\|X\| + b\|Y\| - b\|X-Y\| = \|bY-bX\| - \|-(a+b)X\| \\ &\leq \|bY-bX+(a+b)X\| = \|aX+bY\|. \end{aligned}$$

若 $a<0$ 且 $b>0$则

$$\begin{aligned} a\|X\| + bh'(0^+) &= a\|X\| + b\lim_{t\to 0^+}\frac{\|X+tY\|-\|X\|}{t} \leq a\|X\| + b\frac{\|X-\frac{b}{a}Y\|-\|X\|}{-\frac{b}{a}} \\ &= 2a\|X\| + \|-aX+bY\| = \|-aX+bY\| - \|-2aX\| \\ &\leq \|-aX+bY+2aX\| = \|aX+bY\|. \end{aligned}$$

若 $a>0$ 且 $b<0$则

$$\begin{aligned} a\|X\| + bh'(0^+) &\leq a\|X\| + b\frac{\|X-\frac{b}{a}Y\|-\|X\|}{-\frac{b}{a}} = 2a\|X\| - a\|X-\frac{b}{a}Y\| \\ &= \|2aX\| - \|aX-bY\| \leq \|2aX-aX+bY\| = \|aX+bY\|. \end{aligned}$$

因此，对于所有 $a,b\in\mathbb{R}$，$\psi(aX+bY) \leq \|aX+bY\|$，故 $\|\psi\| = 1$。根据 Hahn-Banach 定理，存在 $\varphi \in E'$，它以相同的范数延拓 ψ；特别地，$\varphi(X) = \|X\|$，因此 $\varphi \in N_X$。但还有 $\varphi(Y) = h'(0^+)$，证明完毕。□

注3.4。 根据前面的讨论，$h: t \mapsto t^{-1}(\|X+tY\| - \|X\|)$ 关于 $t>0$ 单调不增；交换 Y 与 $-Y$ 后可知，对于 $t<0$，函数 h 单调不减；对于 $s<0<t$，我们有

$$\frac{\|X+sY\|-\|X\|}{s} \leq \frac{\|X+tY\|-\|X\|}{t},$$

并且对于每个 $X \neq 0$，我们有

$$\lim_{t\to 0^-} \frac{\|X+tY\|-\|X\|}{t} = \min_{\varphi\in N_X} \varphi(Y) \leq \max_{\varphi\in N_X} \varphi(Y) = \lim_{t\to 0^+} \frac{\|X+tY\|-\|X\|}{t}. \tag{15}$$

3.2. **耗散算子与不等式**。 现在回顾一下，若一个算子 T 作用于复巴拿赫空间 $(E, \|\cdot\|_E)$ 中，则称其为 耗散的，如果对每个 $\xi \in E$ 以及向量 ξ的每个范数化泛函 $\phi \in E^*$，都有 $\mathrm{Re}\phi(T\xi) \leq 0$。由 Hille-Yosida 定理和耗散算子理论（见 [27,第 1] 章及 [21]）可知，对任意 $T \in \mathcal{B}(E)$

定理 3.5. 以下命题等价

(1) 对每个 $\xi \in E$，都存在某个 ξ 的范数化泛函 $\phi \in E^*$，使得 $\mathrm{Re}\phi(T\xi) \leq 0$。*(2)* 对所有 $s \geq 0$，$\|e^{sT}\| \leq 1$。*(3)* 对每个 $s \geq 0$，$1 - sT$ 都是可逆且扩张的，即 $\|\xi - sT\xi\|_E \geq \|\xi\|_E$。

注 3.6（复化与 Taylor 范数）。 设 A，$B \in \mathfrak{u}_n$，对于 $\mathfrak{u}_n$ 上给定的 Ad-不变范数，令

$$\|A+iB\|_T = \sup_{t\in[0,2\pi]} \|A\cos t - B\sin t\|$$

为 $A+iB$ 在 $\mathfrak{u}_n$ 的复化 $M_n(\mathbb{C}) = \mathfrak{u}_n \oplus i\,\mathfrak{u}_n$ 上的 Taylor 范数。这是复化上的一个范数，它将 $\mathfrak{u}_n$ 上的范数扩展到复化上；容易验证 $\mathrm{k}A+iB\mathrm{k}_T = \mathrm{k}A-iB\mathrm{k}_T$，并且对于任意酉算子 U，都有 $\mathrm{k}U(A+iB)U^*\mathrm{k}_T = \mathrm{k}A+iB\mathrm{k}_T$（事实上，可能有许多种复化，见 [25]）。另一个问题：我们能否利用 [25]，中列出的任一种标准复化方法，将其应用于 $\mathfrak{u}_n$ 上的 p-范数，从而恢复整个 $M_n(\mathbb{C})$ 上的 p-范数?

定理 3.7。 对于任意 $X \in \mathfrak{u}_n$，算子 $\mathrm{ad}\ X$ 和 $\mathrm{ad}^2 X$ 均为耗散的。特别地，对于所有 $s \in \mathbb{R}$，$1 + s\mathrm{ad}\ X$ 都是可逆且扩张的；对于所有 $s \geq 0$，$1 - s\mathrm{ad}^2 X$ 都是可逆且扩张的。

证明。 如果将 $\mathrm{ad}\ X$复化，即 $\mathrm{ad}\ X(A+iB) = \mathrm{ad}\ X(A) + i\,\mathrm{ad}\ X(B)$，则可推出 $\mathrm{k}e^{s\,\mathrm{ad}\ X}(A+iB)\mathrm{k}_T = \mathrm{k}e^{sX}(A+iB)e^{-sX}\mathrm{k}_T = \mathrm{k}A+iB\mathrm{k}_T$。因此，$\mathrm{ad}\ X$的复化是耗散的，这意味着 $\mathrm{ad}\ X$ 也是耗散的。将 X 替换为 sX 和 $-sX$后，对 $\pm s\mathrm{ad}\ X$同样成立。现在可以写成 $1 - s^2\,\mathrm{ad}^2 X = (1 - s\,\mathrm{ad}\ X)(1 + s\,\mathrm{ad}\ X)$，由于两个因子都是可逆且扩张的，故其乘积也是如此，因此 $\mathrm{ad}^2 X$ 也是耗散的。□

定理 3.8。 设 X、$V \in \mathfrak{u}_n$，并设 $\|\cdot\|$ 是 $\mathfrak{u}_n$上的一个 Ad-不变芬斯勒范数。那么，如果 $0 \leq s \leq s'$，则有

$$\|V\| \leq \|V - s[X,[X,V]]\| \leq \|V - s'[X,[X,V]]\|$$

并且如果 $t'/t \geq 1$，则有

$$\|V\| \leq \|V + t[X,V]\| \leq \|V + t'[X,V]\|.$$

证明。 为简便起见，我们将 $\mathrm{ad}^2 X = (\mathrm{ad}\ X)^2$ 记作。令 $s \geq 0$，则 $1 - s\mathrm{ad}^2 X$ 是扩张的且可逆的，特别地 $\mathrm{k}(1 - s\mathrm{ad}^2 X)^{-1}\mathrm{k} \leq 1$。令 $\xi = A + iB \in (\mathfrak{u}_n \oplus i\,\mathfrak{u}_n,\ \mathrm{k}\cdot\mathrm{k}_T)$，设 ϕ是 ξ的任意范数化泛函，则

$$\mathrm{Re}\phi(((1 - s\,\mathrm{ad}^2 X)^{-1} - 1)\xi) = \mathrm{Re}\phi(((1 - s\,\mathrm{ad}^2 X)^{-1}\xi) - \|\xi\|_T \leq \|\xi\|_T - \|\xi\|_T = 0,$$

因此 $(1-s\operatorname{ad}^2 X)^{-1}-1=s\operatorname{ad}^2 X(1-s\operatorname{ad}^2 X)^{-1}$ 是耗散算子，$\operatorname{ad}^2 X(1-s\operatorname{ad}^2 X)^{-1}$也同样是耗散算子，因为耗散算子的任意正倍数仍是耗散算子。因此，若 $0\le s\le s'$，则算子

$$A=(1-s'\operatorname{ad}^2 X)(1-s\operatorname{ad}^2 X)^{-1}=1-(s'-s)\operatorname{ad}^2 X(1-s\operatorname{ad}^2 X)^{-1}$$

是扩张的且可逆，因此其逆映射是一个压缩映射。于是

$$\begin{aligned}\|(1-s\operatorname{ad}^2 X)V\| &= (1-s\operatorname{ad}^2 X)(1-s'\operatorname{ad}^2 X)^{-1}(1-s'\operatorname{ad}^2 X)V\| \\ &= \|A^{-1}(1-s'\operatorname{ad}^2 X)V\|\le\|(1-s'\operatorname{ad}^2 X)V\|,\end{aligned}$$

由此证明了第一个断言。第二个断言的证明非常类似，因此略去。□

注 3.9 (主序)。 利用命题2.9 ，我们可以用主序来重述前一个定理：如果 $\lambda(A)\subset\mathbb{R}^n$ 表示 $A=\sum^k i\lambda_k p_k\in\mathfrak{u}_n$ 的特征值序列，那么对任意V， $X\in\mathfrak{u}_n$ 以及 $s'/s\ge 1$ ，都有

$$\lambda(V)\prec\lambda(V+s[X,V])\prec\lambda(V+s'[X,V]),$$

对于 $-[X,[X,V]]$ 也有类似的结论。

注3.10 (Birkhoff正交)。 前一个结果也可以改述如下：对于任意X， $V\in\mathfrak{u}_n$ 以及 $\mathfrak{u}_n$ 上的任意 Ad-不变范数，向量 V 与由 $[X,V]$ 张成的子空间 S Birkhoff正交，即

$$\|V\|=\inf_{s\in\mathbb{R}}\|V-s[X,V]\|=\operatorname{dist}(V,S).$$

3.3. **等号成立的情形**。 对于这一特定情形，利用我们已建立的等号判据（定理 2.35），现在我们在命题 2.9 的等价关系中补充一个关于范数化泛函与 ad作用的新等价关系，对应于上文定理 3.8 中等号成立的情形:

定理 3.11. 设 X、 $V\in\mathfrak{u}_n$ ，且 $\|\cdot\|$ 是 $\mathfrak{u}_n$ 上的 Ad-不变芬斯勒范数。以下各条件等价:

(1) $\|V+[X,V]\|=\|V\|$*(2)* 存在矩阵 $V+[X,V]$ 的一个范数化泛函 $\psi=(N|\cdot)$，使得 $[N,X_C]=0$。*(3)* V 与 $V+[X,V]$ 属于球面的同一个面。*(4)* $V+[X,V]$ 属于包含 V 的球面的所有面。

如果将 $[X,V]$ 替换为 $-[X,[X,V]]$，则相同的断言仍成立。如果 V是范数的光滑点，那么在第二个条件中，我们可以将 $[N,X_C]=0$ 替换为 $[N,X]=0$。如果存在满足 $F_\varphi=\{V\}$ 的 φ 范数化V （特别是当范数严格凸时），那么只有在$[X,V]=0$的情况下等号才成立。

证明。 假设第一个条件成立。根据前面的定理，对于 $0<h<1$ ，我们有

$$\|V\|=\|V+[X,V]\|\ge\|V+(1-h)[X,V]\|\ge\|V\|.$$

因此，如果 $f(s)=\mathrm{k}V+s[X,V]\mathrm{k}$，我们有

$$0=f'(1^-)=\lim_{h\to 0^+}\frac{\|V+(1-h)[X,V]\|-\|V+[X,V]\|}{-h}=\min_{\psi\in N_{V+[X,V]}}([X,V]).$$

根据方程（15）。因此存在 $\psi = (N|\cdot)$范数化 $V + [X,V]$使得 $\psi([X,V]) = 0$。现在注意到

$$\|V\| = \|V + [X,V]\| = (N|V + [X,V]) = (N|V)$$

因此 ψ范数 V也如此。另一方面，由于 ψ范数 $V + [X,V]$，我们有 $\psi \circ \mathrm{ad}\, V + \psi \circ \mathrm{ad}\, [X,V] = 0$根据引理 2.28。因此

$$\psi([X,[X,V]]) = -\psi \circ \mathrm{ad}[X,V](X) = \psi \circ \mathrm{ad}\, V(X) = \psi([V,X]) = 0,$$

并且根据定理 2.35 必有 $[N, X_C] = 0$，这就证明了 1。 $\Rightarrow$ 2。由同一定理可知，若范数在 V处光滑，则必有 $[N,X] = 0$；并且若范数严格凸，则必有 $[X,V] = 0$。现在假设第二个断言成立，我们可知

$$\begin{aligned}\|V\| \geq \psi(V) = (N|V) = (N|V) + ([N,X_C]|V) = (N|V) + (N|[X_C,V])\\ = (N|V + [X,V]) = \psi(V + [X,V]) = \|V + [X,V]\| \geq \|V\|\end{aligned}$$

因此 ψ范数同时 V且 $V + [X,V]$，并且它们具有相同的范数，所以 2。 $\Rightarrow$ 3。。若第三个断言成立，则显然 $\|V + [X,V]\| = \|V\|$，因此若 $V \in F_\varphi$，我们有 $\varphi([X,V]) = 0$，于是

$$\|V\| = \varphi(V) = \varphi(V + [X,V]) \leq \|V + [X,V]\| = \|V\|$$

然后 $V + [X,V] \in F_\varphi$，因此 3。 $\Rightarrow$ 4。第四个断言蕴含第一个断言是显然的。关于 $V - [X,[X,V]]$ 的证明类似，因此略去。□

推论 3.12。设 X， $V \in \mathfrak{u}_n$，设 $\lambda \in \mathbb{R}$。则

(1) 若对所有 k， $\lambda_k(V) = \pm i\lambda$，则 $\|V + [X,V]\|_\infty > \|V\|$，除非 $[X,V] = 0$。(2) 若 $V = i\lambda P$，且 P 是一维投影，则 $\|V + [X,V]\|_1 > \|V\|$，除非 $[X,V] = 0$。(3) 设 $1 < k < n$，设 V 如前两个断言中的任意一个所述。则 $\|X + [X,V]\|_{(k)} > \|V\|_{(k)}$，除非 $[X,V] = 0$。

将 $[X,V]$ 替换为 $-[X,[X,V]]$ 可得到类似的陈述。

证明。在第一种情况下， V 是单位球面的极点，因此可以用某个 φ 使得 $F_\varphi = \{V\}$来暴露它，故结论由前一个定理得到。对于第二个断言，证明过程相同，因为这些正是迹范数单位球面的极点。对于第三个断言，使用类似的论证以及 [13, 定理 4]中给出的 Ky-Fan 范数极点刻画。□

问题 3.13。上述定理 3.11 的条件意味着，对于任意范数化泛函 $\varphi_V = (N_V|\cdot)$， $[N_V, X_C] = 0$ 于 V。这是因为，如果条件 4. 成立，且 $\varphi = (N|\cdot)$ 是 V 的任意范数化泛函，那么它也是 $V + [X,V]$的范数化泛函，因此重复 1. 中的论证，我们有 $\Rightarrow$ 2。即$[N, X_C] = 0$。对于所有 N 范数化 V 的 $[N, X_C] = 0$，是否意味着在某个区间 $|s| < \delta$ 内 $\|V + s[X,V]\| = \|V\|$?

注释 3.14。在采取适当的谨慎措施后，本文的结果可以在厄米矩阵的框架下表述。

致谢。 本研究得到国家科学技术研究委员会（CONICET）、国家科学技术促进局（ANPCyT）以及阿根廷布宜诺斯艾利斯大学（UBA）的资助。

参考文献

[1] P.-A. Absil, R. Mahony, R. Sepulchre：《矩阵流形上的优化算法》。普林斯顿大学出版社，普林斯顿（2008）。[2] T. Ando：主序、双随机矩阵与特征值比较。Linear Algebra Appl. 118 (1989), 163–248.[3] E. Andruchow, G. Larotonda, L. Recht：芬斯勒几何与 p-Schatten 酉群的作用。《美国数学会会刊》362 (2010) no. 1, 319–344.[4] E. Andruchow, L. Recht：自伴算子对的截面曲率与对易性。《算子理论杂志》55 (2006) no. 2, 225–238.[5] J. Antezana, G. Larotonda, A. Varela：酉群中对称作用的最优路径。《数学物理通讯》328 (2014), no. 2, 481–497.[6] C. J. Atkin：希尔伯特空间等距变换群的芬斯勒几何。《澳大利亚数学会杂志》Ser. A 42 (1987) no. 2, 196–222.[7] J. J. Benedetto, M. Fickus：有限归一化紧框架。《计算数学进展》18, 357–385 (2003).[8] R. Bhatia：矩阵分析。《数学研究生教材》169。施普林格出版社，纽约，1997。[9] L. Biliotti, A. Ghigi, P. Heinzner：余伴随轨道凸体。《大阪数学杂志》51 (2014), no. 4, 935–968.[10] C. E. Durán, L. E. Mata-Lorenzo, L. Recht：带迹的 C^*-代数的格拉斯曼流形上的自然变分问题。《数学进展》154 (2000), no. 1, 196–228.[11] M. Fabian, P. Habala, P. Hájek, V. Montesinos Santalucía, J. Pelant, V. Zizler：泛函分析与无限维几何。CMS数学丛书／数学／加拿大数学会数学著作丛书，8。施普林格出版社，纽约，2001。[12] M. Fickus, D. G. Mixon, J.C. Tremain：施泰纳等角紧框架。《线性代数及其应用》436 (2012) no.5, 1014–1027.[13] R. Grone, M. Marcus：矩阵代数的等距变换。《代数杂志》47 (1977), no. 1, 180–189[14] A. Jensen：自极多面体。多面体与离散几何，101–124，《当代数学》764，美国数学会，罗得岛州普罗维登斯，2021。[15] K. Fan, I. Glicksberg：赋范线性空间中球面的一些几何性质。《杜克数学杂志》25 (1958), 553–568.[16] N. Johnston, S. Moein, R. Pereira, S. Plosker：迹范数中的比尔霍夫–詹姆斯正交性及其在量子资源理论中的应用。《线性代数电子期刊》38 (2022) 760–776.[17] M. Hayashi：量子信息论。数学基础。第二版。《物理学研究生教材》。施普林格出版社，柏林，2017。[18] R. Hemasinha：置换基。《科学与技术数学教育国际期刊》25 (1994), no.1, 103–111.[19] A. S. Holevo：位移参数广义不确定性的界。概率论与数理统计（第比利斯，1982），243–251，《数学讲义》1021，施普林格出版社，柏林，1983。[20] A. Horn：双随机矩阵与旋转矩阵的对角线。《美国数学杂志》76 (1954), 620–630.[21] G. Larotonda：算子理想中的范数不等式。《泛函分析杂志》255 (2008), no. 11, 3208–3228.[22] G. Larotonda：矩阵与算子的赫尔德不等式中等号成立的情形。《爱尔兰皇家科学院数学学报》118A (2018), no. 1, 1–4。

[23] G. Larotonda：无限维李群及其齐性空间的度量几何。Forum Math. 31 (2019), no. 6, 1567–1605.[24] G. Larotonda, M. Miglioli：紧李群的霍费尔度量，arXiv 预印本（2020）。[25] G. Muñoz, Y. Sarantopoulos, A. Tonge：实巴拿赫空间、多项式与多重线性映射的复化。Studia Math. 134 (1999), no. 1, 1–33.[26] M. A. Nielsen, I. L. Chuang：量子计算与量子信息。剑桥大学出版社，剑桥，2000。[27] A. Pazy：线性算子半群及其在偏微分方程中的应用。应用数学科学，44。施普林格出版社，纽约，1983。[28] J. A. Tropp：放宽约束：利用凸规划方法从噪声中识别稀疏信号。IEEE 信息论汇刊，第52卷，第3期，第1030–1051页。[29] G. A. Watson：利用一类多面体范数进行线性最佳逼近。Numer. Algorithms 2 (1992), no. 3-4, 321–335.[30] G. A. Watson：关于具有 *Ky Fan* k 范数的矩阵逼近问题。逼近算法，III（牛津，1992）。Numer. Algorithms 5 (1993), no. 1-4, 263–272.[31] K. Ziętak：矩阵单位球面极点的刻画。Linear Algebra Appl. 106 (1988), 57–75。

数学系，FCEyN-UBA，以及阿根廷数学研究所，CONICET，布宜诺斯艾利斯，阿根廷电子邮箱地址：glaroton@dm.uba.ar

数学系，FCEyN-UBA，以及阿根廷数学研究所，CONICET，布宜诺斯艾利斯，阿根廷电子邮箱地址：ivanrey1988@gmail.com