arXiv:2602.21258v1 [math.DG] 23 Feb 2026

J-厄米矩阵锥与几何平均

JOSE FRANCO 和 ALLAN MERINO

摘要。我们研究与不定符号矩阵 $J = \mathrm{Id}_{p,q}$ 相关的正 J-厄米矩阵锥 $\mathscr{P}_J$。我们证明 J-指数映射是双射，并借此分析 $\mathscr{P}_J$ 的代数与几何结构。通过与正定矩阵锥的典范同一化，我们赋予 $\mathscr{P}_J$ 自然的黎曼结构。在此框架下，我们将 J-几何平均定义为测地线的中点，并证明它唯一地由一个 Riccati 型方程的解刻画。

目录

1. 引言 1 2. J-厄米矩阵锥 5 3. $\mathscr{P}_J$ 上的平方根及其性质 14 4. 锥 $\mathscr{P}_J$ 17上的 Loewner 序 5. 锥 $\mathscr{P}_J$ 19上的测地线 6. 锥 $\mathscr{P}_J$ 21上的几何平均 附录 A. 四元数情形 29 A.1. 除代数 $\mathbb{H}$ 29 A.2. 正四元数厄米矩阵锥 31 A.3. $\mathscr{P}$ 35上的黎曼型 A.4. $\mathscr{P}$ 37上的测地线 A.5. $\mathscr{P}$ 41上的几何平均 参考文献 42

1. 引言

Pusz 和 Woronowicz [19] 首次通过函数演算，令人满意地解决了如何定义两个正定矩阵的几何平均这一问题。他们将其定义为

$$A \sharp B = A^{\frac{1}{2}} \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \right)^{\frac{1}{2}} A^{\frac{1}{2}} .$$

2010数学主题分类。主要分类：15A42；次要分类：47A63。关键词和短语。Krein 空间、J-厄米矩阵、测地线、几何平均、四元数。

随后，基于 Kubo 和 Ando [15]对算子均值的系统研究，这种均值的加权版本可以自然地定义为 $t \in [0,1]$，定义如下：

$$A \sharp_t B = A^{\frac{1}{2}} \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \right)^t A^{\frac{1}{2}} .$$

从那时起，人们便从许多不同的角度研究这种平均。以 Ando-Hiai 不等式为例，[1] 该不等式指出：

$$A \sharp_t B \leq \mathrm{Id} \qquad \Longrightarrow \qquad A^p \sharp_t B^p \leq \mathrm{Id}, \qquad (p \geq 1) .$$

这一不等式有多种应用，涉及 Golden-Thompson 不等式和 Furuta 不等式。

另一方面，人们从黎曼几何的角度研究了正定矩阵的加权几何平均。在 [7], 中，Bhatia 和 Holbrook 将加权几何平均确立为连接锥中两个点的测地曲线，其中该锥由正定矩阵构成。与这一观点密切相关的是，CAT(0) 空间存在唯一的重心。从这个意义上说，几何平均可视为重心

$$A \sharp B = \arg\min_{X \geq 0} \left(\delta_2(X, A)^2 + \delta_2(X, B)^2 \right) .$$

这一视角自然引出了多元几何均值，但本文不讨论这一主题。感兴趣的读者可参阅 [16] 了解更多详情。

此外，还有一个问题：如何将几何平均的定义推广到正定矩阵以外的其他空间。例如，Hiai 和 Kosaki 将定义推广到了正 τ-可测算子以及 Haagerup 的 L^p 空间中的正定元素 [12]。在 Choi、Kim 和 Lim 最近发表的论文 [8], 中，他们为幂零矩阵的加权几何平均给出了二项式公式。此处我们还提及 [18], 中的研究，该研究通过李理论方法为非紧型对称空间建立了几何平均。函数演算是这些构造中的关键要素，但它依赖于谱定理，因此也依赖于算子的自伴性。当算子不是厄米算子（甚至不是正规算子）时，函数演算便不再适用，正算子的经典平均理论也就无法使用。然而，在许多应用中，这一限制并不构成问题，因为算子通常是自伴的。

在相对论或不定度量情境中，相关算子通常并非通常希尔伯特空间意义下的自伴算子。例如，在闵可夫斯基空间中，洛伦兹厄米算子所起的作用类似于量子理论中的厄米算子。这促使人们发展新的框架，将均值理论从正定情形推广到 Krein 空间；本文针对几何平均研究这一问题。[9]中曾研究过这个问题。然而，他们未能通过算子理论技术将该理论推广到这一情境。算子论技术在不定情形中的失效，促使我们采用一种基于李理论和表示论工具的结构性方法，从而得到显式公式和全局几何描述，而这些结果无法通过经典泛函演算获得。

在 Krein 空间上研究矩阵分析课题并非新鲜事，参见例如[2, 4, 5] 等少数文献。

设 $\mathbb{D} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$，并令 J 为 $\mathrm{Mat}(n) := \mathrm{Mat}(n \times n, \mathbb{D})$ 中的矩阵，定义为 $\mathrm{J} = \mathrm{diag}\left(\mathrm{Id}_p, -\mathrm{Id}_q\right)$，其中 $p + q = n$。利用 J，我们在 $\mathrm{Mat}(n)$ 上定义一个对合 $\sharp$，如下:

$$\mathrm{A} \mapsto \mathrm{A}^{\sharp} := \mathrm{J}\mathrm{A}^{*}\mathrm{J},$$

其中 $\mathrm{A}^{*} = \overline{\mathrm{A}}^{t}$，它在不定情形中起伴随的作用。我们记

$$\mathfrak{p}_{\mathrm{J}} := \left\{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X}^{\sharp} = \mathrm{X}\right\}$$

为 J-厄米矩阵构成的实向量空间，并令

$$\mathfrak{p} = \{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X} = \mathrm{X}^{*}\}.$$

令 $\mathscr{P}$ 为 $\mathfrak{p}$ 中由下式给出的子集

$$\mathscr{P} := \{\mathrm{X} \in \mathfrak{p}, \mathrm{X} \text{ is positive}\}.$$

众所周知，指数映射 $\exp : \mathfrak{p} \to \mathscr{P}$ 有定义且为双射。本文在 $\mathfrak{p}_{\mathrm{J}}$ 中定义一个锥，它与 $\mathscr{P}$ 类似。然而，指数映射在 $\mathfrak{p}_{\mathrm{J}}$ 上的限制并非单射，因此我们并未采用这种方法。

在 $\mathfrak{p}_{\mathrm{J}}$ 中，我们区分出开锥

$$\mathscr{P}_{\mathrm{J}} := \{\mathrm{A} \in \mathfrak{p}_{\mathrm{J}}, \mathrm{JA} \text{ is positive}\},$$

我们将其称为 J-正矩阵锥（在 $\mathfrak{p}_{\mathrm{J}}$ 中）。尽管 $\mathscr{P}_{\mathrm{J}}$ 在乘法下并不稳定，但它具有若干与正定埃尔米特矩阵的经典锥 $\mathscr{P} := \exp(\mathfrak{p})$ 类似的性质。特别是，$\mathscr{P}_{\mathrm{J}}$ 在求逆下稳定，并且群 $\mathrm{GL}(n)$ 通过合同作用在 $\mathscr{P}_{\mathrm{J}}$ 上传递。在此，我们想将我们的结果与 [18] 中得到的结果进行对比。虽然空间 $\mathscr{P}_{\mathrm{J}}$ 与非紧型对称空间同构，但从 [18], 中的研究来看，尚不清楚如何从计算角度着手。最明显的例子是四元数情形，其中黎曼度量源于约化迹而非正则迹（见附录 A）。他们的构造依赖于李代数层面的经典指数映射；当 $\mathrm{J} \neq \mathrm{Id}_n$ 时，该映射在不定情形下不再具有整体双射性。因此，我们的工作应被视为对 [18] 的补充。

本文的核心是线性映射

$$\Phi_{\mathrm{J}} : \mathfrak{p}_{\mathrm{J}} \ni \mathrm{X} \longrightarrow \mathrm{JX} \in \mathfrak{p},$$

它在 $\mathscr{P}_{\mathrm{J}}$ 与经典正锥 $\mathscr{P}$ 之间构成光滑双射。借助这一同一化，我们可以将解析和几何结构从 $\mathscr{P}$ 传递到 $\mathscr{P}_{\mathrm{J}}$。特别地，我们将 J-指数映射定义为

$$\exp_{\mathrm{J}}(\mathrm{X}) := \mathrm{J}\exp(\mathrm{JX}), \qquad (\mathrm{X} \in \mathfrak{p}_{\mathrm{J}}),$$

由此得到从 $\mathfrak{p}_{\mathrm{J}}$ 到 $\mathscr{P}_{\mathrm{J}}$的整体微分同胚。其双射性直接源于经典指数映射 exp 在 $\mathfrak{p}$ 上的双射性，以及线性

同构 Φ_J。其逆映射 $\log_J$ 提供了自然的框架，可通过从 $\mathscr{P}$的经典几何拉回，在 $\mathscr{P}_J$ 上定义测地线和黎曼结构。在集合 $\mathrm{Mat}(n)$ 上，我们用 $\bullet$ 表示如下乘法:

$$X \bullet Y = XJY\,, \qquad (X\,, Y \in \mathrm{Mat}(n))\,.$$

这种乘法满足结合律，并且 $X \bullet J = J \bullet X = J$。此外，矩阵 A 可逆当且仅当存在矩阵 $B \in \mathrm{Mat}(n)$，使得 $A \bullet B = B \bullet A = J$，且有 $B = JA^{-1}J$。这种乘法使我们能够以特别简单的形式定义测地线。

我们将 $\mathscr{P}$ 上的黎曼度量记为 $\langle \cdot\,, \cdot \rangle$，其定义为

$$\langle U\,, V\rangle_P = \widetilde{\mathrm{tr}}(P^{-1}UP^{-1}V)\,, \qquad (P \in \mathscr{P}\,, U\,, V \in T_P(\mathscr{P}) \cong \mathfrak{p})\,,$$

其中 $\widetilde{\mathrm{tr}} : \mathrm{Mat}(n) \to \mathbb{R}$ 定义为

$$\widetilde{\mathrm{tr}}(X) = \begin{cases} \mathrm{tr}(X) & \text{if } \mathbb{D} \in \{\mathbb{R}\,, \mathbb{C}\} \\ \mathrm{Re}(\mathrm{tr}(X)) & \text{if } \mathbb{D} = \mathbb{H} \end{cases}$$

并令 $\omega := \Phi_J^*\langle \cdot\,, \cdot \rangle$ 为黎曼度量在 $\mathscr{P}_J$ 上的拉回。在第5节中，我们证明以下定理。

定理1.1. 设 $A,B \in \mathscr{P}_J$。映射 $\gamma\colon [0\,, 1] \to \mathscr{P}_J$ 定义为

$$\gamma(t) := A_J^{\frac{1}{2}} \bullet \left(A_J^{-\frac{1}{2}} \bullet B \bullet A_J^{-\frac{1}{2}}\right)_J^t \bullet A_J^{\frac{1}{2}}\,, \qquad (t \in [0\,, 1])\,,$$

这是 $\mathscr{P}_J$ 中A 与 B 之间测地线的方程，其中采用度量 ω；这里，X_J^t是 $\mathscr{P}_J$ 中的矩阵，定义为

$$X_J^t = \exp_J(t\log_J(X))\,, \qquad (X \in \mathscr{P}_J)\,.$$

在此框架下，我们将 J-几何平均定义为 $\mathscr{P}_J$中测地线的中点，即

$$A \sharp^J B := A_J^{\frac{1}{2}} \bullet \left(A_J^{-\frac{1}{2}} \bullet B \bullet A_J^{-\frac{1}{2}}\right)_J^{\frac{1}{2}} \bullet A_J^{\frac{1}{2}}\,, \qquad (A\,, B \in \mathscr{P}_J)\,.$$

我们的主要结果表明，这种平均具有类似于 Pusz–Woronowicz构造的显式闭式公式，并满足合同不变性（在 $\mathrm{GL}(n$ 下）、关于 J-Loewner序的单调性，以及在同一化Φ_J下与经典几何平均的相容性等基本性质。

理论框架在第 2 节中建立，函数演算在第 3 节中引入。在第 4 节中，我们针对本情形引入 J-Loewner序。在第5节中，我们研究 J-正矩阵锥上的测地曲线，进而给出 J-正矩阵的加权几何平均的定义，并在第 6 节中通过研究其基本性质作结。在附录 A 中，我们构造正四元数厄米矩阵锥。

2. J-厄米矩阵锥

设 $\mathbb{D} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$，并设 ι 是 $\mathbb{D}$ 上由 $\iota(\mathrm{X}) = \mathrm{X}$ 给出的对合；特别地，对于 $\mathbb{D} = \mathbb{R}$，ι 是平凡的，并且对于 $\mathbb{D} = \mathbb{H}$（见附录 A.1），我们有

$$\iota(\mathrm{X} := a + ib + jc + kd) = a - ib - jc - kd\,, \qquad (\mathrm{X} \in \mathbb{H})\,.$$

设 V 是一个右-$\mathbb{D}$-模，且 $\dim_{\mathbb{D}}(\mathrm{V}) = n$，B 是 V 上一个非退化的 $(\iota, 1)$-埃尔米特形式，即对所有 x、y、$z \in \mathrm{V}$ 以及所有 α、$\beta \in \mathbb{D}$

(1) $\mathrm{B}(x, y\alpha + z\beta) = \mathrm{B}(x, y)\alpha + \mathrm{B}(x, z)\beta$，(2) $\mathrm{B}(x\alpha + y\beta, z) = \alpha\mathrm{B}(x, z) + \overline{\beta}\mathrm{B}(y, z)$，(3) $\mathrm{B}(y, x) = \mathrm{B}(x, y)$，

(4) $\mathrm{Rad}(\mathrm{B}) := \{y \in \mathrm{V},\ \mathrm{B}(x, y) = 0,\ (\forall x \in \mathrm{V})\} = \{0\}$。

我们用 $\mathrm{End}(\mathrm{V}) := \mathrm{End}_{\mathbb{D}}(\mathrm{V})$ 表示 V 上线性映射的集合，即

$$\mathrm{End}(\mathrm{V}) := \{\mathrm{T} : \mathrm{V} \to \mathrm{V}\,, \mathrm{T}(v_1 + v_2\lambda) = \mathrm{T}(v_1) + \mathrm{T}(v_2)\lambda\,, v_1\,, v_2 \in \mathrm{V}\,, \lambda \in \mathbb{D}\}\,,$$

并用 $\mathrm{GL}(\mathrm{V})$ 表示 $\mathrm{End}(\mathrm{V})$ 中可逆映射的集合。我们用 $\mathrm{Mat}(n) := \mathrm{Mat}(n \times n, \mathbb{D})$ 表示元素属于 $\mathbb{D}$ 的 n 阶方阵集合。对于所有 $\mathrm{T} \in \mathrm{End}(\mathrm{V})$ 以及 V 的任意右-$\mathbb{D}$-基 $\mathscr{B}_{\mathrm{V}} = \{v_1, \ldots, v_n\}$，我们用 $\mathrm{Mat}^{\mathscr{B}_{\mathrm{V}}}(\mathrm{T}) := \left(t_{i,j}\right)_{1 \le i,j \le n}$ 表示 $\mathrm{Mat}(n)$ 中由下式给出的对应矩阵

$$\mathrm{T}(v_i) = \sum_{j=1}^{n} t_{j,i} v_j\,, \qquad (1 \le i, j \le n)\,,$$

并满足 $[\mathrm{T}(v)] = \mathrm{A}[v]$, 其中 $[v] \in \mathbb{D}^n$ 是 v 在 $\mathscr{B}_{\mathrm{V}}$ 中的坐标。众所周知，映射

$$\Gamma : \mathrm{End}(\mathrm{V}) \ni \mathrm{T} \mapsto \mathrm{Mat}_{\mathscr{B}_{\mathrm{V}}}(\mathrm{T}) \in \mathrm{Mat}(n) \tag{1}$$

是双射，并令 $\mathrm{GL}(n) := \Gamma(\mathrm{GL}(\mathrm{V}))$。

注2.1. (1) $\mathbb{R}$-线性对合 ι 可扩展到 $\mathbb{D}^n$ 和 $\mathrm{Mat}(n)$，如下所示：对于 $v = (v_i)_{1 \le i \le n} \in \mathbb{D}^n$ 和 $\mathrm{A} = \left(a_{i,j}\right)_{1 \le i,j \le n} \in \mathrm{Mat}(n)$，定义

$$\iota(v) = (\iota(v_i))_{1 \le i \le n}\,, \qquad \iota(\mathrm{A}) = \left(\iota(a_{i,j})\right)_{1 \le i,j \le n}\,.$$

此外，我们有 $\iota(\mathrm{A}v) = \iota(\mathrm{A})\iota(v)$。

(2) 设 $\mathscr{B}_{\mathrm{V}} = \{v_1, \ldots, v_n\}$ 是 V 的右-$\mathbb{D}$-基，并设 Ω 为 $\mathrm{Mat}(n)$ 中由 $\Omega = \left(\mathrm{B}(v_i, v_j)\right)_{1 \le i,j \le n}$ 给出的矩阵。利用型 B 是厄米的这一性质，可得

$$\mathrm{B}(x, y) = \iota(x)^t \Omega y\,, \qquad (x, y \in \mathrm{V})\,.$$

记号 2.2. 对于矩阵 $\mathrm{A} \in \mathrm{Mat}(n)$ 和向量 $v \in \mathbb{D}^n$，我们用 A^* 和 v^* 分别表示 $\mathrm{Mat}(n)$ 和 $\mathbb{D}^n$ 中的元素，它们由 $\mathrm{A}^* = \iota(\mathrm{A})^t$ 和 $v^* = \iota(v)^t$ 给出。

我们现在回顾以下经典定理。

定理 2.3 (西尔维斯特惯性定律)**.** 设 V 为维数为 n 的右 $\mathbb{D}$-模，并设 B 为 V 上的非退化 $(\iota,1)$-埃尔米特形式。则存在整数 $p,q \geq 0$，满足 $p+q=n$，以及 V 的一个右 $\mathbb{D}$-基 $\mathscr{B}=\{e_1,\ldots,e_n\}$，使得 B 在 $\mathscr{B}$ 中的矩阵 $\mathrm{Mat}_{\mathscr{B}}(\mathrm{B})$ 为 $\mathrm{Id}_{p,q}=\mathrm{diag}\left(\mathrm{Id}_p,-\mathrm{Id}_q\right)$。

等价地，

$$\mathrm{B}(x,y)=\sum_{k=1}^{p}\iota(x_k)y_k-\sum_{k=p+1}^{p+q}\iota(x_k)y_k,\qquad \left(x=\sum_{k=1}^{n}e_kx_k,y=\sum_{k=1}^{n}e_ky_k\right),$$

其中 x_k、$y_k\in\mathbb{D}$。

证明。 证明见 [20]。 □

注 2.4。 设 $\mathrm{T}\in\mathrm{End}(\mathrm{V})$。利用型 B 是非退化的这一点，存在唯一的元素 $\mathrm{T}^{\diamond}\in\mathrm{End}(\mathrm{V})$，使得

$$\mathrm{B}(\mathrm{T}(v_1),v_2)=\mathrm{B}(v_1,\mathrm{T}^{\diamond}(v_2)),\qquad (v_1,v_2\in\mathrm{V}). \tag{2}$$

元素 $\mathrm{T}^{\diamond}\in\mathrm{End}(\mathrm{V})$ 称为 T 关于 B 的伴随。此外，相应的映射

$$\mathrm{End}(\mathrm{V})\ni\mathrm{T}\to\mathrm{T}^{\diamond}\in\mathrm{End}(\mathrm{V})$$

定义了 End(V) 上的一个 $\mathbb{R}$-线性对合。

设 $\mathscr{B}$ 是 V 的一组基，使得 $\mathrm{Mat}_{\mathscr{B}}(\mathrm{B})=\mathrm{Id}_{p,\,q}$。为简化记号，我们用 J 表示矩阵 $\mathrm{Id}_{p,\,q}$，即

$$\mathrm{J}=\begin{bmatrix}\mathrm{Id}_p & 0\\ 0 & -\mathrm{Id}_q\end{bmatrix}.$$

我们用 U_{J} 表示 (V,B) 的等距变换群，即由下式给出的实李群

$$\mathrm{U}_{\mathrm{J}}:=\{g\in\mathrm{GL}(\mathrm{V}),\mathrm{B}(g(v_1),g(v_2))=\mathrm{B}(v_1,v_2),(v_1,v_2\in\mathrm{V})\}.$$

我们用 $\mathfrak{u}_{\mathrm{J}}$ 表示 U_{J} 的李代数。特别地，我们得到

$$\mathfrak{u}_{\mathrm{J}}=\{\mathrm{X}\in\mathrm{End}(\mathrm{V}),\mathrm{B}(\mathrm{X}(v_1),v_2)+\mathrm{B}(v_1,\mathrm{X}(v_2))=0,(\forall v_1,v_2\in\mathrm{V})\}.$$

类似地，我们用 $\mathfrak{p}_{\mathrm{J}}$ 表示 End(V) 的如下子集

$$\mathfrak{p}_{\mathrm{J}}:=\{\mathrm{X}\in\mathrm{End}(\mathrm{V}),\mathrm{B}(\mathrm{X}(v_1),v_2)-\mathrm{B}(v_1,\mathrm{X}(v_2))=0,(\forall v_1,v_2\in\mathrm{V})\}.$$

我们有

$$\mathrm{End}(\mathrm{V})=\mathfrak{u}_{\mathrm{J}}\oplus\mathfrak{p}_{\mathrm{J}},$$

并且若用 $[\cdot,\cdot]$ 表示 End(V) 上的李括号（即对于所有 $\mathrm{X},\mathrm{Y}\in\mathrm{End}(\mathrm{V})$，$[\mathrm{X},\mathrm{Y}]=\mathrm{X}\circ\mathrm{Y}-\mathrm{Y}\circ\mathrm{X}$），则可以看出

$$[\mathfrak{u}_{\mathrm{J}},\mathfrak{u}_{\mathrm{J}}]\subseteq\mathfrak{u}_{\mathrm{J}},\qquad [\mathfrak{u}_{\mathrm{J}},\mathfrak{p}_{\mathrm{J}}]\subseteq\mathfrak{p}_{\mathrm{J}},\qquad [\mathfrak{p}_{\mathrm{J}},\mathfrak{p}_{\mathrm{J}}]\subseteq\mathfrak{u}_{\mathrm{J}}.$$

注记 2.5. (1) 我们通过将 $v\in\mathrm{V}$ 替换为其在基 $\mathscr{B}$ 下的坐标，从而将 V 与 $\mathbb{D}^n$ 等同。因此，V 上的型 B 诱导出 $\mathbb{D}^n$ 上的一个非退化 $(\iota,1)$-埃尔米特形式 B。

(2) 设 $\mathscr{B}$ 是 V 的一组基，使得 $\mathrm{Mat}_{\mathscr{B}}(\mathrm{B}) = \mathrm{J}$。利用同构 (1)，我们可以将 $\mathfrak{u}_{\mathrm{J}}$ 和 $\mathfrak{p}_{\mathrm{J}}$ 写成 Mat(n) 的子集。事实上，利用对所有u、$v \in$V 都有 $\mathrm{B}(u,v) = u^{*}\mathrm{J}v$ 这一点，我们得到

$$\begin{aligned}\mathfrak{u}_{\mathrm{J}} &= \left\{\mathrm{X} \in \mathrm{Mat}(n), (\mathrm{X}v_1)^{*}\mathrm{J}v_2 + v_1^{*}\mathrm{J}(\mathrm{X}v_2) = 0, (v_1, v_2 \in \mathbb{D}^n)\right\} \\ &= \left\{\mathrm{X} \in \mathrm{Mat}(n), v_1^{*}(\mathrm{X}^{*}\mathrm{J} + \mathrm{J}\mathrm{X})v_2 = 0, (v_1, v_2 \in \mathbb{D}^n)\right\} \\ &= \{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X}^{*}\mathrm{J} = -\mathrm{J}\mathrm{X}\}.\end{aligned}$$

类似地，我们有

$$\mathfrak{p}_{\mathrm{J}} = \{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X}^{*}\mathrm{J} = \mathrm{J}\mathrm{X}\}.$$

集合 $\mathfrak{p}_{\mathrm{J}}$ 称为 J-厄米矩阵集。此外，我们有

$$\mathrm{U}_{\mathrm{J}} = \{g \in \mathrm{GL}(n), g^{*}\mathrm{J}g = \mathrm{J}\}.$$

(3) 对所有 $\mathrm{X} \in \mathrm{Mat}(n)$，记 $\mathrm{X}^{\sharp}$ 为 Mat(n) 中由下式给出的矩阵

$$\mathrm{X}^{\sharp} = \mathrm{J}\mathrm{X}^{*}\mathrm{J}.$$

容易看出，对所有 $\mathrm{X}, \mathrm{Y} \in \mathrm{Mat}(n)$，均有 $\left(\mathrm{X}^{\sharp}\right)^{\sharp} = \mathrm{X}$ 且 $(\mathrm{XY})^{\sharp} = \mathrm{Y}^{\sharp}\mathrm{X}^{\sharp}$。

因此，映射

$$\mathrm{Mat}(n) \ni \mathrm{X} \to \mathrm{X}^{\sharp} \in \mathrm{Mat}(n)$$

是一个 $\mathbb{R}$-线性对合。此外，容易看出，对所有 $\mathrm{T} \in \mathrm{End}(\mathrm{V})$，均有

$$\mathrm{Mat}_{\mathscr{B}}(\mathrm{T}^{\diamond}) = \mathrm{Mat}_{\mathscr{B}}(\mathrm{T})^{\sharp},$$

其中 $\mathrm{T}^{\diamond}$ 定义于公式（2）中。

(4) 利用上文定义的对合 $\sharp$，我们可以重写 Mat(n) 的子集 $\mathfrak{u}_{\mathrm{J}}$、$\mathfrak{p}_{\mathrm{J}}$ 和 U_{J}。事实上，我们有

$$\mathfrak{u}_{\mathrm{J}} = \left\{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X}^{\sharp} = -\mathrm{X}\right\}, \qquad \mathfrak{p}_{\mathrm{J}} = \left\{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{X}^{\sharp} = \mathrm{X}\right\},$$

and

$$\mathrm{U}_{\mathrm{J}} = \left\{g \in \mathrm{GL}(n), g^{\sharp}g = \mathrm{Id}_n\right\}.$$

(5) 令 B_{J} 为 $\mathbb{D}^n$ 上的型，定义如下:

$$\mathrm{B}_{\mathrm{J}}(u_1, u_2) = \mathrm{B}(\mathrm{J}u_1, u_2), \qquad (u_1, u_2 \in \mathbb{D}^n).$$

型 B_{J} 是 (ι, 1)-厄米且正定的。事实上，对所有 $u \in \mathbb{D}^n$，我们有

$$\mathrm{B}_{\mathrm{J}}(u, u) = \mathrm{B}(\mathrm{J}u, u) = (\mathrm{J}u)^{*}\mathrm{J}u = u^{*}\mathrm{J}^2u = u^{*}u = \sum_{i=1}^{n}|u_i|^2,$$

也就是说，对所有非零向量 $u \in \mathbb{D}^n$，都有 $\mathrm{B}_{\mathrm{J}}(u,u) > 0$。

（6）设 $\mathfrak{u}$、$\mathfrak{p}$ 分别为 Mat(n) 的以下子集

$$\mathfrak{u} = \{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{B}_{\mathrm{J}}(\mathrm{X}u_1, u_2) + \mathrm{B}_{\mathrm{J}}(u_1, \mathrm{X}u_2) = 0, (u_1, u_2 \in \mathbb{D}^n)\}$$

and

$\mathfrak{p} = \{\mathrm{X} \in \mathrm{Mat}(n), \mathrm{B}_{\mathrm{J}}(\mathrm{X}u_1, u_2) - \mathrm{B}_{\mathrm{J}}(u_1, \mathrm{X}u_2) = 0, (u_1, u_2 \in \mathbb{D}^n)\}$。

如上文所述，我们有

$$\mathfrak{u} = \{\mathrm{H} \in \mathrm{Mat}(n), \mathrm{H} = -\mathrm{H}^*\}, \qquad \mathfrak{p} = \{\mathrm{H} \in \mathrm{Mat}(n), \mathrm{H} = \mathrm{H}^*\},$$

and

$$[\mathfrak{u}, \mathfrak{u}] \subseteq \mathfrak{u}, \qquad [\mathfrak{u}, \mathfrak{p}] \subseteq \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{u}.$$

集合 $\mathfrak{p}$ 称为厄米矩阵集合。类似地，我们用 U 表示由下式给出的实李群

$$\mathrm{U} = \{g \in \mathrm{GL}(n), \mathrm{B}_\mathrm{J}(gu_1, gu_2) = \mathrm{B}_\mathrm{J}(u_1, u_2), (u_1, u_2 \in \mathbb{D}^n)\}.$$

i.e.

$$\mathrm{U} = \{g \in \mathrm{GL}(n), g^* g = \mathrm{Id}_n\}.$$

使用 [14], 的记号，我们有

$$\mathrm{U}_\mathrm{J} = \begin{cases} \mathrm{O}(p,q) & \text{if } \mathbb{D} = \mathbb{R} \\ \mathrm{U}(p,q) & \text{if } \mathbb{D} = \mathbb{C} \\ \mathrm{Sp}(p,q) & \text{if } \mathbb{D} = \mathbb{H} \end{cases}, \qquad \mathrm{U} = \begin{cases} \mathrm{O}(n) & \text{if } \mathbb{D} = \mathbb{R} \\ \mathrm{U}(n) & \text{if } \mathbb{D} = \mathbb{C} \\ \mathrm{Sp}(n) & \text{if } \mathbb{D} = \mathbb{H} \end{cases}.$$

因此，U_J 连通当且仅当 $\mathbb{D} \in \{\mathbb{C}, \mathbb{H}\}$，而紧致当且仅当 $\mathrm{J} = \mathrm{Id}_n$。

定义 2.6。设 $\mathrm{A} \in \mathrm{Mat}(n)$。若对所有非零向量 $u \in \mathbb{D}^n$，都有 $\mathrm{B}_\mathrm{J}(\mathrm{A}u, u) > 0$，则称 A 为正定。

容易看出，若 A 为正定，则 $0 \notin \mathrm{Spec}(\mathrm{A})$。我们用 $\mathscr{P}$ 表示正定厄米矩阵的集合，即

$$\mathscr{P} := \{\mathrm{X} \in \mathfrak{p}, \mathrm{X} \text{ is positive}\}.$$

记号 2.7。若矩阵 $\mathrm{X} \in \mathrm{Mat}(n)$ 为正定，则记作 $\mathrm{X} > 0$。

我们用 exp 表示 $\mathrm{Mat}(n)$ 上的指数映射，即由下式给出的映射

$$\mathrm{Mat}(n) \ni \mathrm{A} \to \exp(\mathrm{A}) := \sum_{k=0}^{\infty} \frac{\mathrm{A}^k}{k!} \in \mathrm{GL}(n).$$

众所周知（见 [6, 第 1]章），exp 在 $\mathfrak{p}$ 上的限制是单射，且 $\exp(\mathfrak{p}) = \mathscr{P}$。换言之，

$$\exp : \mathfrak{p} \to \mathscr{P}$$

是双射。

注释 2.8。现在的目标是定义 $\mathfrak{p}_\mathrm{J}$ 中的一个锥，使其类似于 $\mathfrak{p}$ 中的锥$\mathscr{P}$。自然可以尝试仿照经典锥 $\mathscr{P} \subset \mathfrak{p}$ 来定义 $\mathscr{P}_\mathrm{J} \subset \mathfrak{p}_\mathrm{J}$，即将 $\mathfrak{p}_\mathrm{J}$ 通过指数映射得到的像。然而，这种方法行不通：指数映射在 $\mathfrak{p}_\mathrm{J}$ 上的限制不是单射，因此不可能与其像构成双射。事实上，设 X 是 Mat(2) 中的矩阵，定义如下：

$$\mathrm{X} = \begin{bmatrix} 0 & 2i\pi \\ 2i\pi & 0 \end{bmatrix}.$$

对于 $J = \mathrm{Id}_{\mathfrak{l},\mathfrak{l}}$，我们有

$$JX^*J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -2i\pi \\ -2i\pi & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2i\pi \\ 2i\pi & 0 \end{bmatrix} = X$$

即 $X \in \mathfrak{p}_J$。利用 $\exp(X) = \mathrm{Id}_2 = \exp(0)$，我们得到指数映射在 $\mathfrak{p}_J$ 上不是单射。因此，指数映射不能用于参数化 $\mathfrak{p}_J$ 内的正锥，而需要一种适应于 J-结构的不同构造。

以下引理是本文的关键：它表明 $\mathfrak{p}_J$ 可以典范地等同于厄米矩阵空间。

引理 2.9. 映射

$$\Phi_J : \mathfrak{p}_J \ni X \to JX \in \mathfrak{p}$$

是良定义的且为双射。

证明。 我们首先验证 $\Phi_J(X) = JX$ 将 $\mathfrak{p}_J$ 映入 $\mathfrak{p}$。设 $H \in \mathfrak{p}_J$，从而 $H^* = JHJ$。于是

$$(JH)^* = H^*J^* = H^*J = (JHJ)\,J = JHJ^2 = JH\,,$$

因此 $JH \in \mathfrak{p}$。由于 J 可逆，映射 Φ_J 是单射。为证明满射性，设 $P \in \mathfrak{p}$，并令 $X := JP$。则

$$X^* = (JP)^* = P^*J^* = PJ\,, \qquad JXJ = J\,(JP)\,J = PJ\,.$$

因此 $X^* = JXJ$，即 $X \in \mathfrak{p}_J$，并且

$$\Phi_J(X) = JX = J\,(JP) = P\,.$$

这证明 Φ_J 是满射，因此是双射。

□

注记 2.10。 映射 $\Phi_J : \mathrm{Mat}(n) \to \mathrm{Mat}(n)$ 定义为 $\Phi_J(X) = JX$，满足$\Phi_J(u_J) = u$ 且 $\Phi_J(\mathfrak{p}_J) = \mathfrak{p}$。然而，映射 Φ_J 不是李代数同态。

特别地，我们得到一个双射

$$\exp \circ \Phi_J : \mathfrak{p}_J \to \mathscr{P}\,. \tag{3}$$

设 $\Gamma_J : \mathscr{P} \to \mathrm{Mat}(n)$ 为由下式给出的映射

$$\Gamma_J(X) = JX\,, \qquad (X \in \mathscr{P})\,,$$

并令 $\mathscr{P}_J := \mathrm{Im}(\Gamma_J)$。由构造可知，映射 $\Gamma_J : \mathscr{P} \to \mathscr{P}_J$ 是双射，并且我们用$\exp_J$ 表示由下式给出的映射

$$\begin{array}{ccc} \mathfrak{p}_J & \xrightarrow{\Phi_J} & \mathfrak{p} \\ {\scriptstyle \exp_J}\downarrow & & \downarrow{\scriptstyle \exp} \\ \mathscr{P}_J & \xleftarrow{\Gamma_J} & \mathscr{P} \end{array}$$

特别地，我们有

$$\exp_{\mathrm{J}}(\mathrm{X}) = \mathrm{J}\exp(\mathrm{JX}), \qquad (\mathrm{X} \in \mathfrak{p}_{\mathrm{J}}) . \tag{4}$$

现在我们来描述集合 $\mathscr{P}_{\mathrm{J}} := \exp_{\mathrm{J}}(\mathfrak{p}_{\mathrm{J}})$。首先引入 J-正定厄米矩阵的定义。

定义 2.11. 若对所有非零的$x \in \mathbb{D}^n$，矩阵 $\mathrm{H} \in \mathfrak{p}_{\mathrm{J}}$ 满足 $\mathrm{B}(\mathrm{H}x, x) > 0$，则称 H 为 J-正定矩阵。

注 2.12. (1) 换言之，若对所有非零的$x \in \mathbb{D}^n$，$x^*\mathrm{H}^*\mathrm{J}x > 0$，则称 $\mathrm{H} \in \mathfrak{p}_{\mathrm{J}}$ 为 J-正定矩阵。(2) 矩阵 H 属于 $\mathscr{P}_{\mathrm{J}}$ 当且仅当 $\mathrm{JH} \in \mathscr{P}$。

下面的命题给出了 J-正矩阵的一种优美参数化。

命题 2.13. 设 $\mathrm{H} \in \mathrm{Mat}(n)$。

(1) 矩阵 H 是J-厄米矩阵（即$\mathrm{H}^\sharp = \mathrm{H}$，等价地$\mathrm{H}^* = \mathrm{JHJ}$），当且仅当相对于分解 $\mathbb{D}^n = \mathbb{D}^p \oplus \mathbb{D}^q$，它可以写成

$$\mathrm{H} = \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ -\mathrm{B}^* & \mathrm{D} \end{bmatrix}, \tag{5}$$

其中$\mathrm{A} \in \mathrm{Mat}(p)$ 和$\mathrm{D} \in \mathrm{Mat}(q)$为厄米矩阵（即$\mathrm{A}^* = \mathrm{A}$且$\mathrm{D}^* = \mathrm{D}$），而$\mathrm{B} \in \mathrm{Mat}(p \times q)$ 是任意矩阵。

(2) 此外，若H 是J-正定的（即$\mathrm{JH} > 0$），则A 是正定的，D 是负定的（即$(-\mathrm{D}) > 0$），且满足以下舒尔补条件:

$$(-\mathrm{D}) - \mathrm{B}^*\mathrm{A}^{-1}\mathrm{B} > 0 .$$

证明。(1) 令 $\mathrm{H} \in \mathfrak{p}_{\mathrm{J}}$。我们将 H 按照分解$\mathbb{D}^n = \mathbb{D}^p \oplus \mathbb{D}^q$写成分块形式，即

$$\mathrm{H} = \begin{bmatrix} \mathrm{H}_{1,1} & \mathrm{H}_{1,2} \\ \mathrm{H}_{2,1} & \mathrm{H}_{2,2} \end{bmatrix},$$

其中 $\mathrm{H1}_{,1} \in \mathrm{Mat}(p)$，$\mathrm{H1}_{,2} \in \mathrm{Mat}(p \times q)$，$\mathrm{H2}_{,1} \in \mathrm{Mat}(q \times p)$，且 $\mathrm{H2}_{,2} \in \mathrm{Mat}(q)$。于是

$$\mathrm{H}^* = \begin{bmatrix} \mathrm{H}^*_{1,1} & \mathrm{H}^*_{2,1} \\ \mathrm{H}^*_{1,2} & \mathrm{H}^*_{2,2} \end{bmatrix},$$

and

$$\mathrm{JHJ} = \begin{bmatrix} \mathrm{Id}_p & 0 \\ 0 & -\mathrm{Id}_q \end{bmatrix} \begin{bmatrix} \mathrm{H}_{1,1} & \mathrm{H}_{1,2} \\ \mathrm{H}_{2,1} & \mathrm{H}_{2,2} \end{bmatrix} \begin{bmatrix} \mathrm{Id}_p & 0 \\ 0 & -\mathrm{Id}_q \end{bmatrix} = \begin{bmatrix} \mathrm{H}_{1,1} & -\mathrm{H}_{1,2} \\ -\mathrm{H}_{2,1} & \mathrm{H}_{2,2} \end{bmatrix} .$$

因此，条件 $\mathrm{H}^* = \mathrm{JHJ}$ 等价于以下等式组

$$\mathrm{H}^*_{1,1} = \mathrm{H}_{1,1} , \qquad \mathrm{H}^*_{2,2} = \mathrm{H}_{2,2} , \qquad \mathrm{H}^*_{2,1} = -\mathrm{H}_{1,2} , \qquad \mathrm{H}^*_{1,2} = -\mathrm{H}_{2,1} .$$

因此，$\mathrm{H1}_{,1} = \mathrm{H}^*_{1,1}$、$\mathrm{H2}_{,2} = \mathrm{H}^*_{2,2}$ 以及 $\mathrm{H2}_{,1} = -\mathrm{H}^*_{1,2}$。反之，任何形如 (5) 的矩阵显然都满足 $\mathrm{H}^* = \mathrm{JHJ}$，从而证明了 (1)。

（2）现在假设 H 是 J-正定的，即 $JH > 0$。利用 (1) 中得到的分块形式，我们得到

$$JH = \begin{bmatrix} \mathrm{Id}_p & 0 \\ 0 & -\mathrm{Id}_q \end{bmatrix} \begin{bmatrix} A & B \\ -B^* & D \end{bmatrix} = \begin{bmatrix} A & B \\ B^* & -D \end{bmatrix}.$$

该矩阵是厄米矩阵。由于 $JH > 0$，其主对角块必须正定，因此 $A > 0$。同理，末尾主对角块必须正定，因此 $-D > 0$，即 $D < 0$。最后，由于 $A > 0$ 可逆，对厄米分块矩阵 JH 应用正定性的舒尔补判据（见[13]），得到

$$JH > 0 \qquad \Longleftrightarrow \qquad A > 0 \quad \text{and} \quad (-D) - B^* A^{-1} B > 0 .$$

这就证明了 (2)。

□

命题 2.14. $\mathscr{P}_J$ 是$\mathfrak{p}_J$中J-正定矩阵的集合。

证明. 设 $H \in \mathfrak{p}_J$，即 $H^*J = JH$。利用对所有 $A \in \mathrm{Mat}(n)$，均有 $\exp(A)^* = \exp(A^*)$，可得

$$\begin{aligned} \exp_J(H)^* J &= (J \exp(JH))^* J = \exp(JH)^* J^* J = \exp\left((JH)^*\right) J^2 \\ &= \exp(H^* J^*) = \exp(H^* J) = \exp(JH) \end{aligned}$$

and

$$J \exp_J(H) = J\,(J \exp(JH)) = J^2 \exp(JH) = \exp(JH),$$

即 $\exp_J(H) \in \mathfrak{p}_J$。此外，对所有非零$x \in \mathbb{D}^n$，有

$$x^* (J \exp_J(H))\, x = x^* \exp(JH) x > 0$$

因为 $JH \in \mathfrak{p}$，即 $\exp(JH) \in \mathscr{P}$。因此，$\exp_J(H)$ 是 J-正定的，即

$$\mathscr{P}_J \subseteq \{X \in \mathfrak{p}_J, X \text{ is J-positive}\} .$$

现在设 H 是 $\mathfrak{p}_J$ 中的一个 J-正定矩阵。那么 $JH \in \mathscr{P}$。因此，由等式 (3) 可知，存在 $X \in \mathfrak{p}_J$，使得 $JH = \exp(JX)$。在前一个等式两边左乘 J，可得 $H = J \exp(JX) = \exp_J(X)$。因此 $H \in \exp_J(\mathfrak{p}_J)$，命题得证。

□

注 2.15。设 $X \in \mathfrak{p}$，即 $X = X^*$。如上所述，我们有 $JX \in \mathfrak{p}_J$。容易看出，XJ 也属于 $\mathfrak{p}_J$。事实上，

$$(XJ)^\sharp = JX^\sharp = J\,(JX^* J) = X^* J = XJ,$$

即 $XJ \in \mathfrak{p}_J$。此外，若 $X > 0$，则 XJ 是 J-正定的。事实上，对所有 $x \in \mathbb{D}^n$，我们有

$$x^* J(XJ) x = (Jx)^*\, X\, (Jx) > 0,$$

即 $XJ \in \mathscr{P}_J$。因此，作为集合，我们有

$$\mathscr{P}_J = J\mathscr{P} = \mathscr{P} J .$$

锥 $\mathscr{P}_J$ 不是 GL(n) 的子群。事实上，它在乘法下并不稳定。然而，下面的引理表明， $\mathscr{P}_J$中一个元素的逆仍位于 $\mathscr{P}_J$中。

引理 2.16. （*1*）若 $H \in \mathscr{P}_J$，则 $H^{-1} \in \mathscr{P}_J$。

（*2*）若 $H = \exp_J(X)$ 且 $X \in \mathfrak{p}_J$，则一般地$H^{-1} \neq \exp_J(-X)$。更精确地说，

$$\exp_J(X)^{-1} = \exp(-JX)J\,, \tag{6}$$

且 $\exp_J(X)^{-1} = \exp_J(-X)$ 成立，当且仅当 $\exp(-JX)$ 与 J 可交换。

证明。（1）设 $H \in \mathscr{P}_J$。由定义，$H \in \mathfrak{p}_J$ 且 $JH > 0$。由于 $H^\sharp = H$，取逆可得

$$\left(H^{-1}\right)^\sharp = \left(H^\sharp\right)^{-1} = H^{-1}\,,$$

从而 $H^{-1} \in \mathfrak{p}_J$。利用映射 Φ_J，我们可以写成 $H = JP$，其中 $P \in \mathscr{P}$。特别地，

$$H^{-1} = P^{-1}J\,,$$

即 $JH^{-1} = J\left(P^{-1}J\right) = JP^{-1}J$。由于$P^{-1} > 0$，故对所有非零 $x \in \mathbb{D}^n$，我们得到

$$x^*JH^{-1}x = x^*JP^{-1}Jx = (Jx)^*\,P^{-1}(Jx) > 0\,,$$

即 $H^{-1} \in \mathscr{P}_J$.

(2) 令 $H = \exp_J(X) = J\exp(JX)$，其中 $X \in \mathfrak{p}_J$。则

$$H^{-1} = (J\exp(JX))^{-1} = \exp(JX)^{-1}J^{-1} = \exp(-JX)J\,,$$

这就证明了 (6)。另一方面，

$$\exp_J(-X) = J\exp(J(-X)) = J\exp(-JX)\,.$$

因此 $(\exp_J(X))^{-1} = \exp_J(-X)$ 当且仅当

$$\exp(-JX)J = J\exp(-JX)\,,$$

即 $\exp(-JX)$ 与 J 对易当且仅当。

□

例 2.17。我们给出一个具体例子，说明 $\exp_J(X)^{-1} \neq \exp_J(-X)$。令 $p = q = 1$，并令 X 为 Mat(2) 中由下式给出的矩阵

$$X = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

利用这一点

$$JX = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix},$$

我们得到 $JX \in \mathfrak{p}$，因此 $X \in \mathfrak{p}_J$。令 $K := JX$。可以看出 $K^2 = \mathrm{Id}_2$,因此

$$\exp(-K) = \begin{bmatrix} \cosh(1) & -i\sinh(1) \\ i\sinh(1) & \cosh(1) \end{bmatrix}.$$

利用方程 (6)，我们得到

$$\exp_{\mathrm{J}}(\mathrm{X})^{-1} = \exp(-\mathrm{K})\mathrm{J} = \begin{bmatrix} \cosh(1) & i\sinh(1) \\ i\sinh(1) & -\cosh(1) \end{bmatrix},$$

然而

$$\exp_{\mathrm{J}}(-\mathrm{X}) = \mathrm{J}\exp(-\mathrm{K}) = \begin{bmatrix} \cosh(1) & -i\sinh(1) \\ -i\sinh(1) & -\cosh(1) \end{bmatrix}.$$

这些矩阵并不相同，这表明一般而言，$\exp_{\mathrm{J}}(\mathrm{X})^{-1} \neq \exp_{\mathrm{J}}(-\mathrm{X})$。

注 2.18。若 $\mathrm{H} \in \mathfrak{p}$，则 $\mathrm{Spec}(\mathrm{H}) \subseteq \mathbb{R}$。然而，对于矩阵 $\mathrm{X} \in \mathfrak{p}_{\mathrm{J}}$，一般而言并不能得到 $\mathrm{Spec}(\mathrm{X}) \subseteq \mathbb{R}$。事实上，令 $\mathrm{X} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathrm{Mat}(2)$。我们有

$$\mathrm{X}^{*}\mathrm{J} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \mathrm{JX},$$

也就是说，$\mathrm{X} \in \mathfrak{p}_{\mathrm{J}}$。然而，$\mathrm{Spec}(\mathrm{X}) = \{\pm i\}$。

注 2.19。我们以一条注记结束本节。我们将 $\mathrm{Mat}(n)$ 上由 $\bullet$ 给出的二元运算记为

$$\mathrm{A} \bullet \mathrm{B} := \mathrm{AJB}, \qquad (\mathrm{A}, \mathrm{B} \in \mathrm{Mat}(n)). \tag{7}$$

该运算 $\bullet$ 是结合的。事实上，对于所有 A、B、$\mathrm{C} \in \mathrm{Mat}(n)$，我们有

$$(\mathrm{A} \bullet \mathrm{B}) \bullet \mathrm{C} = (\mathrm{AJB})\,\mathrm{JC} = \mathrm{AJ}\,(\mathrm{BJC}) = \mathrm{A} \bullet (\mathrm{B} \bullet \mathrm{C}).$$

$\bullet$ 的中性元恰好是 J。事实上，对所有 $\mathrm{A} \in \mathrm{Mat}(n)$，

$$\mathrm{J} \bullet \mathrm{A} = \mathrm{J}\,(\mathrm{JA}) = \mathrm{J}^2\mathrm{A} = \mathrm{A} \qquad \text{and} \qquad \mathrm{A} \bullet \mathrm{J} = (\mathrm{AJ})\,\mathrm{J} = \mathrm{A}.$$

元素 A 关于 J-积可逆，当且仅当它作为矩阵可逆，并且我们有

$$\mathrm{A} \bullet \left(\mathrm{JA}^{-1}\mathrm{J}\right) = \mathrm{AJ}\left(\mathrm{JA}^{-1}\mathrm{J}\right) = \mathrm{AJ}^2\mathrm{A}^{-1}\mathrm{J} = \mathrm{AA}^{-1}\mathrm{J} = \mathrm{J},$$

即 A 关于 $\bullet$ 的逆是 $\mathrm{JA}^{-1}\mathrm{J}$。

此外，对所有 $\mathrm{A}, \mathrm{B} \in \mathrm{Mat}(n)$，我们有

$$(\mathrm{A} \bullet \mathrm{B})^{\sharp} = \mathrm{J}\,(\mathrm{AJB})^{*}\,\mathrm{J} = \mathrm{JB}^{*}\mathrm{JA}^{*}\mathrm{J} = (\mathrm{JB}^{*}\mathrm{J})\,\mathrm{J}\,(\mathrm{JA}^{*}\mathrm{J}) = \mathrm{B}^{\sharp} \bullet \mathrm{A}^{\sharp}.$$

设 $[\cdot,\cdot]$ 是 $\mathrm{Mat}(n)$ 上由下式给出的二元运算

$$[\mathrm{X}, \mathrm{Y}]_{\mathrm{J}} = \mathrm{X} \bullet \mathrm{Y} - \mathrm{Y} \bullet \mathrm{X}, \qquad (\mathrm{X}, \mathrm{Y} \in \mathrm{Mat}(n)).$$

可以看出，该映射

$$\widetilde{\Phi_{\mathrm{J}}} : (\mathrm{Mat}(n), [\cdot,\cdot]) \ni \mathrm{X} \to \mathrm{JX} \in (\mathrm{Mat}(n), [\cdot,\cdot]_{\mathrm{J}})$$

是李代数同构。

映射

$$\exp_{\mathrm{J}} : \mathfrak{p}_{\mathrm{J}} \to \mathscr{P}_{\mathrm{J}}$$

恰好是与 J-积相关联的指数映射。事实上，对于所有 $\mathrm{X}, \mathrm{Y} \in \mathfrak{p}_{\mathrm{J}}$，我们有

$$\exp_{\mathrm{J}}(\mathrm{X}) \bullet \exp_{\mathrm{J}}(\mathrm{Y}) = \exp_{\mathrm{J}}(\mathrm{J} \cdot \mathrm{BCH}_{\mathrm{J}}(\mathrm{JX}, \mathrm{JY})),$$

其中 $\mathrm{BCH}_{\mathrm{J}}$ 是标准的 Baker-Campbell-Hausdorff公式（见 [14]），其中李括号被 $[\cdot,\cdot]_{\mathrm{J}}$ 替代。

最后可以看出，对于所有的 $g \in \mathrm{GL}(n)$，存在唯一的一对 $(k, p) \in \mathrm{U}\times\mathscr{P}_{\mathrm{J}}$，使得 $g = k \bullet p$。事实上，利用 $\mathrm{GL}(n)$ 的经典极分解（见[6]），存在唯一的一对 $(k, \widetilde{p}) \in \mathrm{U} \times \mathscr{P}$，使得 $g = k\widetilde{p}$。因此，利用 $\mathscr{P}_{\mathrm{J}} = \mathrm{J}\mathscr{P}$ 且 $\mathrm{J}^2 = \mathrm{Id}_n$，我们得到

$$g = k\widetilde{p} = k\mathrm{J}^2\widetilde{p} = k\mathrm{J}\,(\mathrm{J}\widetilde{p}) = k \bullet p\,,$$

其中 $p = \mathrm{J}\widetilde{p} \in \mathscr{P}_{\mathrm{J}}$。

3. Square root on $\mathscr{P}_{\mathrm{J}}$ and properties

令 $\exp_{\mathrm{J}} : \mathrm{p}_{\mathrm{J}} \to \mathscr{P}_{\mathrm{J}}$ 为方程（4）中定义的 J-指数映射。如第 2 节所述，映射 $\exp_{\mathrm{J}}$ 是双射。令 $\log_{\mathrm{J}} := \exp_{\mathrm{J}}^{-1}\colon\ \mathscr{P}_{\mathrm{J}} \to \mathrm{p}_{\mathrm{J}}$ 为 $\exp_{\mathrm{J}}$ 的逆。特别地，我们有

$$\log_{\mathrm{J}}(\mathrm{X}) = \mathrm{J}\log(\mathrm{JX})\,, \qquad (\mathrm{X} \in \mathscr{P}_{\mathrm{J}})\,,$$

其中 $\log : \mathscr{P} \to \mathrm{p}$ 是 $\exp : \mathrm{p} \to \mathscr{P}$ 的逆。

定义3.1。对于所有 $t \in \mathbb{R}$ 和 $\mathrm{X} \in \mathscr{P}_{\mathrm{J}}$，我们用 $\mathrm{X}_{\mathrm{J}}^t$ 表示 $\mathscr{P}_{\mathrm{J}}$ 中由下式给出的矩阵

$$\mathrm{X}_{\mathrm{J}}^t = \exp_{\mathrm{J}}(t\log_{\mathrm{J}}(\mathrm{X}))\,.$$

若 $t = \frac{1}{2}$，矩阵 $\mathrm{X}^{\frac{1}{2}_{\mathrm{J}}}$ 是 X 的 J 平方根。

注 3.2. 对于所有 $\mathrm{A} \in \mathscr{P}$ 以及 $t \in \mathbb{R}$，我们将 A^t 记作由 $\mathscr{P}$ 中给出的矩阵

$$\mathrm{A}^t = \exp(t\log(\mathrm{A}))\,.$$

我们有以下性质（见 [6]）：对于所有 A、B $\in \mathscr{P}$ 以及 s、$t \in \mathbb{R}$:

(1) $\mathrm{A}^s\mathrm{A}^t = \mathrm{A}^{s+t}$，且 $(\mathrm{A}^s)^t = \mathrm{A}^{st}$。(2) 若 A 和 B 在 $\mathscr{P}$ 中可交换，则 $\mathrm{AB} \in \mathscr{P}$，且

$$(\mathrm{AB})^s = \mathrm{A}^s\mathrm{B}^s\,.$$

(3) 若 $\mathrm{C} \in \mathrm{U}$，则

$$(\mathrm{CAC}^*)^t = \mathrm{CA}^t\mathrm{C}^*\,.$$

现在我们来证明矩阵 $\mathrm{X}_{\mathrm{J}}^t$ 的一些有趣性质。

命题 3.3。 *(1)* 对所有$\mathrm{X} \in \mathscr{P}_{\mathrm{J}}$ 和 $t \in \mathbb{R}$，有 $\mathrm{X}_{\mathrm{J}}^t = \mathrm{J}(\mathrm{JX})^t$。特别地，$\mathrm{X}^{\frac{1}{2}_{\mathrm{J}}} = \mathrm{J}(\mathrm{JX})^{\frac{1}{2}}$ 且 $(\mathrm{X}^{\frac{1}{2}_{\mathrm{J}}})_{\mathrm{J}}^2 = \mathrm{X}$。

(2) 对所有 s，$t \in \mathbb{R}$ 和 $\mathrm{X} \in \mathscr{P}_{\mathrm{J}}$，有

$$\mathrm{X}_{\mathrm{J}}^0 = \mathrm{J}\,, \qquad \mathrm{X}_{\mathrm{J}}^1 = \mathrm{X}\,, \qquad (\mathrm{X}_{\mathrm{J}}^t)_{\mathrm{J}}^s = \mathrm{X}_{\mathrm{J}}^{ts}\,, \qquad \log_{\mathrm{J}}(\mathrm{X}_{\mathrm{J}}^t) = t\log_{\mathrm{J}}(\mathrm{X})\,.$$

(3) 对于所有$\mathrm{X} \in \mathscr{P}_{\mathrm{J}}$，我们有

$$(\mathrm{X}^{-1})_{\mathrm{J}}^t = (\mathrm{X}_{\mathrm{J}}^t)^{-1}\,,$$

In particular, $(\mathrm{X}^{-1})_{\mathrm{J}}^{\frac{1}{2}} = (\mathrm{X}_{\mathrm{J}}^{\frac{1}{2}})^{-1}$.

(4) 对于所有$X \in \mathscr{P}_J$和$s, \alpha \in \mathbb{R}$，我们得到

$$X_J^s = X_J^{\alpha s} \bullet X_J^{(1-\alpha)s} \qquad \textit{and} \qquad X_J^s \bullet X_J^{-s} = J .$$

证明。（1）利用$\exp_J(H) = J\exp(JH)$和$\log_J(X) = J\log(JX)$，可得

$$X_J^t = \exp_J(t\log_J(X)) = J\exp(t\log(JX)) = J(JX)^t .$$

此外，可得

$$(X_J^{\frac{1}{2}})_J^2 = J\left(JX_J^{\frac{1}{2}}\right)^2 = J\left(J^2(JX)^{\frac{1}{2}}\right)^2 = J(JX)^{\frac{1}{2}}(JX)^{\frac{1}{2}} = J(JX) = J^2X = X .$$

（2）由（1）可得，对所有$X \in \mathscr{P}_J$，$X_J^0 = J(JX)^0 = J$且$X_J^1 = J(JX)^1 = J^2X = X$。此外，对所有$s$、$t \in \mathbb{R}$，

$$\left(X_J^t\right)_J^s = J\left(JX_J^t\right)^s = J\left(J^2(JX)^t\right)^s = J\left((JX)^t\right)^s = J(JX)^{ts} = X_J^{ts},$$

and

$$\begin{aligned}\log_J(X_J^t) = J\log(JX_J^t) &= J\log(J^2(JX)^t) = J\log(\exp(t\log(JX))) \\ &= J(t\log(JX)) = t\log_J(X) .\end{aligned}$$

（3）设$X \in \mathscr{P}_J$，且$P := JX \in \mathscr{P}$。则$X = JP$，且

$$X^{-1} = (JP)^{-1} = P^{-1}J .$$

因此$JX^{-1} = J(P^{-1}J) = JP^{-1}J$。由于$J$是酉的（即$J^* = J = J^{-1}$），由注记3.2可得

$$\left(JP^{-1}J\right)^t = J\left(P^{-1}\right)^t J = JP^{-t}J .$$

由此可得

$$\left(X^{-1}\right)_J^t = J\left(JX^{-1}\right)^t = J\left(JP^{-t}J\right) = P^{-t}J .$$

另一方面，由(1)可得$X_J^t = J(JX)^t = JP^t$，因此

$$\left(X_J^t\right)^{-1} = \left(JP^t\right)^{-1} = \left(P^t\right)^{-1}J = P^{-t}J ,$$

因此

$$\left(X^{-1}\right)_J^t = \left(X_J^t\right)^{-1} .$$

(4) 利用$X_J^s = J(JX)^s$，可得

$$\begin{aligned}X_J^{\alpha s} \bullet X_J^{(1-\alpha)s} &= J(JX)^{\alpha s}J^2(JX)^{(1-\alpha)s} = J(JX)^{\alpha s}(JX)^{(1-\alpha)s} \\ &= J(JX)^{\alpha s+(1-\alpha)s} = J(JX)^s = X_J^s ,\end{aligned}$$

and

$$X_J^s \bullet X_J^{-s} = J(JX)^s J^2(JX)^{-s} = J(JX)^{s-s} = J(JX)^0 = J .$$

□

命题 3.4. 我们有

(1) 对所有的$g \in \mathrm{GL}(n)$以及$X \in \mathscr{P}_J$，$gXg^\sharp \in \mathscr{P}_J$。*(2)* $\mathrm{GL}(n)$在$\mathscr{P}_J$上的相应作用是传递的。

(3) 若 $g \in \mathrm{K_J} := \mathrm{U_J} \cap \mathrm{U}$ （即 $g^\sharp g = \mathrm{Id}_n$ 且 $g^* g = \mathrm{Id}_n$），则对所有 $t \in \mathbb{R}$，可得

$$\left(g\mathrm{X}g^\sharp\right)_\mathrm{J}^t = g\mathrm{X}_\mathrm{J}^t g^\sharp .$$

特别地，我们得到$(g\mathrm{X}g^\sharp)^{\frac{1}{2}\mathrm{J}} = g\mathrm{X}^{\frac{1}{2}\mathrm{J}} g^\sharp$。

(4) 设$\mathrm{X}, \mathrm{Y} \in \mathscr{P}_\mathrm{J}$使得$\mathrm{X} \bullet \mathrm{Y} = \mathrm{Y} \bullet \mathrm{X}$。则对所有 $t \in \mathbb{R}$，我们得到

$$(\mathrm{X} \bullet \mathrm{Y})_\mathrm{J}^t = \mathrm{X}_\mathrm{J}^t \bullet \mathrm{Y}_\mathrm{J}^t .$$

证明。 (1) 设$g \in \mathrm{GL}(n)$ 且 $\mathrm{X} \in \mathscr{P}_\mathrm{J}$。利用 $\mathrm{X}^\sharp = \mathrm{X}$，我们得到

$$\left(g\mathrm{X}g^\sharp\right)^\sharp = \left(g^\sharp\right)^\sharp \mathrm{X}^\sharp g^\sharp = g\mathrm{X}g^\sharp ,$$

即$g\mathrm{X}g^\sharp \in \mathrm{p_J}$。此外，利用 $\mathrm{X} \in \mathscr{P}_\mathrm{J}$，对所有$x \in \mathbb{D}^n$可得

$$\mathrm{B_J}(g\mathrm{X}g^\sharp x, x) = x^* (\mathrm{J}g\mathrm{X}\mathrm{J}g^*\mathrm{J})\, x = (\mathrm{J}g^*\mathrm{J}x)^* \underbrace{\mathrm{JX}}_{>0} (\mathrm{J}g^*\mathrm{J}x) ,$$

即$g\mathrm{X}g^\sharp$ 是 J-正定的。因此$g\mathrm{X}g^\sharp \in \mathscr{P}_\mathrm{J}$。

(2) $\mathrm{GL}(n)$ 在 $\mathscr{P}$ 上的作用是传递的（关于情形 $\mathbb{D} = \mathbb{H}$，参见 [6] 与附录 A）。特别地，对所有 $\mathrm{P} \in \mathscr{P}$，存在$g \in \mathrm{GL}(n)$ 使得 $\mathrm{P} = gg^*$，即 $\mathscr{P} = \{gg^*,\ g \in \mathrm{GL}(n)\}$。利用注记 2.15，我们有$\mathscr{P}_\mathrm{J} = \mathrm{J}\mathscr{P} = \mathscr{P}\mathrm{J}$。对所有$g \in \mathrm{GL}(n)$，我们得到

$$g\mathrm{J}g^\sharp = g\mathrm{J}(\mathrm{J}g^*\mathrm{J}) = g\mathrm{J}^2 g^*\mathrm{J} = \underbrace{gg^*}_{\in\mathscr{P}} \mathrm{J} ,$$

即 $\mathscr{P}_\mathrm{J} \subseteq \left\{g\mathrm{J}g^\sharp, g \in \mathrm{GL}(n)\right\} = \mathscr{P}_\mathrm{J}$，也就是说 $\mathrm{GL}(n) \curvearrowright \mathscr{P}_\mathrm{J}$ 是传递的。

(3) 令$g \in \mathrm{U} \cap \mathrm{U_J}$。根据 X_J^t 的定义，我们得到 $(g\mathrm{X}g^\sharp)_\mathrm{J}^t = \mathrm{J}\left(\mathrm{J}(g\mathrm{X}g^\sharp)\right)^t$。此外，由$g \in \mathrm{U_J}$可知，$g^\sharp g = \mathrm{Id}_n$，即 $\left(g^{-1}\right)^* \mathrm{J} = \mathrm{J}g$。因此$g^*\mathrm{J} = \mathrm{J}g^{-1}$ 且

$$\mathrm{J}(g\mathrm{X}g^\sharp) = \mathrm{J}g\mathrm{X}\mathrm{J}g^*\mathrm{J} = \left(g^{-1}\right)^* \mathrm{J}\mathrm{X}\mathrm{J}^2 g^{-1} = \left(g^{-1}\right)^* \mathrm{J}\mathrm{X}g^{-1} .$$

利用 $\mathrm{JX} \in \mathscr{P}$ 以及$g \in \mathrm{U}$，可得

$$\left(\mathrm{J}(g\mathrm{X}g^\sharp)\right)^t = \left(g^{-1}\right)^* (\mathrm{JX})^t g^{-1} .$$

将前一个等式左乘 J，得到

$$(g\mathrm{X}g^\sharp)_\mathrm{J}^t = \mathrm{J}\left(g^{-1}\right)^* (\mathrm{JX})^t g^{-1} .$$

最后，由 $\mathrm{J}\left(g^{-1}\right)^* = g\mathrm{J}$ 和 $g^\sharp = g^{-1}$ 可得

$$(g\mathrm{X}g^\sharp)_\mathrm{J}^t = g\mathrm{J}(\mathrm{JX})^t g^{-1} = g\mathrm{X}_\mathrm{J}^t g^{-1} = g\mathrm{X}_\mathrm{J}^t g^\sharp ,$$

这就证明了所需性质。

（4）利用 $XJY = YJX$，可得 $(JX)(JY) = (JY)(JX)$。特别地，由注记 3.2 可得 $((JX)(JY))^t = (JX)^t(JY)^t$。

我们得到

$$(X \bullet Y)_J^t = J\,(J(XJY))^t = J\,((JX)(JY))^t = J\,(JX)^t\,(JY)^t$$

and

$$X_J^t \bullet Y_J^t = J\,(JX)^t\,J^2\,(JY)^t = J\,(JX)^t\,(JY)^t\ ,$$

因此结论成立。

□

注 3.5. 命题 3.4 中定义的群 K_J 是 U_J 的极大紧子群。更准确地说，我们有

$$K_J := \begin{cases} O(p)\times O(q) & \text{if } U_J = O(p,q) \\ U(p)\times U(q) & \text{if } U_J = U(p,q) \\ Sp(p)\times Sp(q) & \text{if } U_J = Sp(p,q) \end{cases}.$$

注 3.6. 本节中的计算表明，$\mathscr{P}_J$ 是.....的自然类比
正定厄米矩阵锥，其中单位矩阵
由 J 替代，通常的乘积由 $\bullet$ 替代。
容易看出，这在 $\mathrm{Mat}(n)$ 上赋予了一个 $*$-代数结构。然而，注意，与对合 $\sharp$ 相关联的自伴算子，即 $\mathfrak{p}_J$ 的元素，并不具有实谱。因此，这排除了使用函数演算的可能性，而在 C^*-代数中则可以使用函数演算。在后一种情况下，几何平均是众所周知的。

4. 锥 $\mathscr{P}_J$上的 Loewner 序

本节首先回顾 $\mathscr{P}$上 Loewner 序的定义。首先，我们用 $\mathscr{P}_0$ 表示 $\mathfrak{p}$ 中半正定厄米矩阵的集合，即

$$\mathscr{P}_0 := \{X \in \mathfrak{p}, B_J(Xu,u) \geq 0, (u \in \mathbb{D}^n)\}\ .$$

特别地，$\mathscr{P} \subseteq \mathscr{P}_0$，并且若 $X \in \mathscr{P}_0$，则记作 $X \geq 0$。
$\mathscr{P}$ 上的 Loewner 序是由下式定义的偏序

$$A \preceq B \qquad \Longleftrightarrow \qquad B - A \in \mathscr{P}_0\,,$$

也就是说，$B - A$ 是一个半正定厄米矩阵。等价地，$A \preceq B$ 当且仅当对所有非零 $v^*(B - A)v \geq 0$，都有 $v \in \mathbb{D}^n$。

注 4.1。该序关系在合同变换下保持闭合。事实上，如果 $A \preceq B$，那么对于所有 $M \in \mathrm{Mat}(n)$，我们有

$$M^*AM \preceq M^*BM\,.$$

此外，对于所有 $t \in [0,1]$，我们有

$$A^t \preceq B^t\,.$$

现在，我们利用（$\mathscr{P}$，$\preceq$）以及映射$\mathscr{P}_J \ni X \to JX \in \mathscr{P}$，在锥 $\mathscr{P}_J$ 上定义了一个偏序。

定义 4.2. 对于所有 X、$Y \in \mathscr{P}_J$，我们称 $X \preceq_J Y$，当且仅当 $JX \preceq JY$。

此序恰好是通过线性同构 $\mathscr{P}_J \ni X \mapsto JX \in \mathscr{P}$ 从 $\mathscr{P}$ 上的 Loewner 序拉回得到的。

容易看出，$\preceq_J$ 在 $\mathscr{P}_J$ 上定义了一个偏序。接下来的引理给出了备注 4.1 中结果在我们的锥 $\mathscr{P}_J$ 上的类似结论。

引理 4.3. 若 $X, Y \in \mathscr{P}_J$ 满足 $X \preceq_J Y$，则对所有 $t \in [0,1]$，有

$$X_J^t \preceq_J Y_J^t .$$

证明。设 $X \preceq_J Y$，即 $JX \preceq JY$。由于 JX 和 JY 都属于 $\mathscr{P}$，由备注 4.1 可知，对所有 $t \in [0,1]$，有

$$(JX)^t \preceq (JY)^t .$$

由 $J^2 = \mathrm{Id}$，可得 $(JX)^t = J^2(JX)^t = JX_J^t$，因此

$$JX_J^t \preceq JY_J^t ,$$

这恰好说明

$$X_J^t \preceq_J Y_J^t .$$

□

备注 4.4. 引理 4.3 中对 $t \in [0,1]$ 的限制是必要的，因为映射 $A \mapsto A^t$ 仅在此范围内是算子单调的。

引理 4.5. 若 $X, Y \in \mathscr{P}_J$ 满足 $X \preceq_J Y$，则对所有 $C \in \mathrm{Mat}(n)$，有

$$C^\sharp X C \preceq_J C^\sharp Y C .$$

证明. 假设 $X \preceq_J Y$。则在 $\mathscr{P}$ 中，$JX \preceq JY$。由备注 4.1 可知，对所有 $C \in \mathrm{Mat}(n)$，有

$$C^*(JX)C \preceq C^*(JY)C .$$

利用 $C^\sharp = J^{-1}C^*J$，可得

$$C^*(JX)C = J(C^\sharp X C), \qquad C^*(JY)C = J(C^\sharp Y C) .$$

因此，

$$J(C^\sharp X C) \preceq J(C^\sharp Y C) .$$

根据 $\preceq_J$ 的定义，这等价于

$$C^\sharp X C \preceq_J C^\sharp Y C .$$

□

以下引理由 Loewner 序的定义直接得到。我们记录下来以备后用。

引理 4.6。若 X、$Y \in \mathscr{P}_J$ 满足 $X \preceq_J Y$，则

$$X^{-1} \succeq_J Y^{-1} .$$

5. 锥 $\mathscr{P}_{\mathrm{J}}$上的测地线

我们在本节开头回顾正定厄米矩阵之锥 $\mathscr{P}$的一些经典结果（$\mathbb{D} \in \{\mathbb{R}, \mathbb{C}\}$的情形见 [6]，$\mathbb{D} = \mathbb{H}$的情形见附录 A）。对每个 $\mathrm{P} \in \mathscr{P}$，我们将切空间 $\mathrm{T}_{\mathrm{P}}\mathscr{P}$ 与 $\mathfrak{p}$ 等同。$\mathscr{P}$ 上的黎曼度量定义为

$$\langle \mathrm{U}, \mathrm{V} \rangle_{\mathrm{P}} = \widetilde{\mathrm{tr}}\left(\mathrm{P}^{-1}\mathrm{U}\mathrm{P}^{-1}\mathrm{V}\right), \qquad (\mathrm{U}, \mathrm{V} \in \mathrm{T}_{\mathrm{P}}(\mathscr{P}) \cong \mathfrak{p}) . \tag{8}$$

$\widetilde{\ }$，其中 $\mathrm{tr} : \mathbb{D} \to \mathbb{R}$ 定义为

$$\widetilde{\mathrm{tr}}(\mathrm{X}) = \begin{cases} \mathrm{tr}(\mathrm{X}) & \text{if } \mathbb{D} \in \{\mathbb{R}, \mathbb{C}\} \\ \mathrm{trd}(\mathrm{X}) & \text{if } \mathbb{D} = \mathbb{H} \end{cases}$$

其中，trd 是方程 18 中定义的约化迹。
如 [6] 和定理 A.20 所述，$\mathscr{P}$ 中连接 P 与 Q 的唯一测地线为

$$\gamma(t) = \mathrm{P}^{\frac{1}{2}}\left(\mathrm{P}^{-\frac{1}{2}}\mathrm{Q}\mathrm{P}^{-\frac{1}{2}}\right)^{t}\mathrm{P}^{\frac{1}{2}}, \qquad (t \in [0, 1]) . \tag{9}$$

对于所有 $g \in \mathrm{GL}(n)$ 和 $\mathrm{X} \in \mathscr{P}$，有$g\mathrm{X}g^{*} \in \mathscr{P}$。对于所有 $\mathrm{Q} \in \mathrm{GL}(n)$，我们用 ζ_{Q} 表示映射

$$\zeta_{\mathrm{Q}} : \mathscr{P} \ni \mathrm{P} \longmapsto \mathrm{Q}\mathrm{P}\mathrm{Q}^{*} \in \mathscr{P} .$$

对于所有 $\mathrm{P} \in \mathscr{P}$，有

$$\langle (d\zeta_{\mathrm{Q}})_{\mathrm{P}}(\mathrm{U}), (d\zeta_{\mathrm{Q}})_{\mathrm{P}}(\mathrm{V}) \rangle_{\zeta_{\mathrm{Q}}(\mathrm{P})} = \langle \mathrm{U}, \mathrm{V} \rangle_{\mathrm{P}}, \qquad (\mathrm{U}, \mathrm{V} \in \mathfrak{p} = \mathrm{T}_{\mathrm{P}}(\mathscr{P})) .$$

即黎曼形式是 $\mathrm{GL}(n)$-不变的。

注记 5.1. $\mathscr{P}$ 上测地线的显式形式可利用与方程 (8) 中定义的 $\langle \cdot, \cdot \rangle$ 相关联的列维-奇维塔联络得到。
如 [17] 中所解释的（在 $\mathbb{D} \in \{\mathbb{R}, \mathbb{C}\}$ 情形下——四元数情形的证明类似），$\mathscr{P}$ 上的列维-奇维塔联络由下式给出

$$\nabla_{\mathrm{U}}(\mathrm{V}) = \mathrm{D}_{\mathrm{U}}(\mathrm{V}) - \frac{1}{2}\left(\mathrm{U}\mathrm{P}^{-1}\mathrm{V} + \mathrm{V}\mathrm{P}^{-1}\mathrm{U}\right),$$

其中 $\mathrm{D}_{\mathrm{U}}(\mathrm{V})$ 表示向量场 V 沿方向 U 的方向导数。光滑曲线 $\gamma : \mathrm{I} \to \mathscr{P}$ 是测地线，当且仅当它满足测地线方程

$$\nabla_{\dot{\gamma}(t)}(\dot{\gamma}(t)) = 0 ,$$

由此得到二阶微分方程

$$\ddot{\gamma}(t) = \dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) . \tag{10}$$

方程 (10) 的解的存在性和唯一性可由流形上的标准常微分方程理论得出。最后一步是证明方程 (9) 中给出的曲线 γ 满足方程 (10)。令 $\mathrm{L} = \log\left(\mathrm{P}^{-\frac{1}{2}}\mathrm{Q}\mathrm{P}^{-\frac{1}{2}}\right) \in \mathfrak{p}$，即 $\gamma(t) = \mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{P}^{\frac{1}{2}}$。由于 L 是常数，且 $\frac{d}{dt}\exp(t\mathrm{L}) = \exp(t\mathrm{L})\mathrm{L}$，因此得到：

$$\dot{\gamma}(t) = \mathrm{P}^{\frac{1}{2}}\left(\frac{d}{dt}\exp(t\mathrm{L})\right)\mathrm{P}^{\frac{1}{2}} = \mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}\mathrm{P}^{\frac{1}{2}},$$

$$\ddot{\gamma}(t) = \mathrm{P}^{\frac{1}{2}}\left(\frac{d}{dt}\big(\exp(t\mathrm{L})\mathrm{L}\right)\mathrm{P}^{\frac{1}{2}} = \mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}^2\mathrm{P}^{\frac{1}{2}},$$

and

$$\gamma(t)^{-1} = \left(\mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{P}^{\frac{1}{2}}\right)^{-1} = \mathrm{P}^{-\frac{1}{2}}\exp(-t\mathrm{L})\mathrm{P}^{-\frac{1}{2}}.$$

因此

$$\begin{aligned}\dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) &= \left(\mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}\mathrm{P}^{\frac{1}{2}}\right)\left(\mathrm{P}^{-\frac{1}{2}}\exp(-t\mathrm{L})\mathrm{P}^{-\frac{1}{2}}\right)\left(\mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}\mathrm{P}^{\frac{1}{2}}\right)\\ &= \mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}\exp(-t\mathrm{L})\exp(t\mathrm{L})\mathrm{L}\mathrm{P}^{\frac{1}{2}}\\ &= \mathrm{P}^{\frac{1}{2}}\exp(t\mathrm{L})\mathrm{L}^2\mathrm{P}^{\frac{1}{2}} = \ddot{\gamma}(t),\end{aligned}$$

这里用到了 $\exp(-t\mathrm{L})\exp(t\mathrm{L}) = \mathrm{Id}_n$，以及 L 与 $\exp(t\mathrm{L})$ 可交换这一事实。

现在，我们利用 $\mathscr{P}$ 上的度量 $\langle\cdot,\cdot\rangle$，在 $\mathscr{P}_\mathrm{J}$ 上构造一个黎曼度量。令Φ_J：$\mathscr{P}_\mathrm{J} \to \mathscr{P}$为映射 $\Phi_\mathrm{J}(\mathrm{X}) = \mathrm{J}\mathrm{X}$。令 $\omega = \Phi_\mathrm{J}^*(\langle\cdot,\cdot\rangle)$ 为黎曼度量在 $\mathscr{P}_\mathrm{J}$ 上的拉回。根据定义，对于所有 $\mathrm{P} \in \mathscr{P}_\mathrm{J}$，有

$$\omega_\mathrm{P}(\mathrm{U},\mathrm{V}) = \langle \mathrm{J}\mathrm{U},\mathrm{J}\mathrm{V}\rangle_{\mathrm{J}\mathrm{P}}, \qquad (\mathrm{U},\mathrm{V} \in \mathfrak{p}_\mathrm{J}).$$

特别地，我们得到

$$\begin{aligned}\omega_\mathrm{P}(\mathrm{U},\mathrm{V}) &= \langle \mathrm{J}\mathrm{U},\mathrm{J}\mathrm{V}\rangle_{\mathrm{J}\mathrm{P}} = \widetilde{\mathrm{tr}}\left((\mathrm{J}\mathrm{P})^{-1}\mathrm{J}\mathrm{U}(\mathrm{J}\mathrm{P})^{-1}\mathrm{J}\mathrm{V}\right)\\ &= \widetilde{\mathrm{tr}}(\mathrm{P}^{-1}\mathrm{J}^{-1}\mathrm{J}\mathrm{U}\mathrm{P}^{-1}\mathrm{J}^{-1}\mathrm{J}\mathrm{V}) = \widetilde{\mathrm{tr}}(\mathrm{P}^{-1}\mathrm{U}\mathrm{P}^{-1}\mathrm{V}),\end{aligned}$$

i.e.

$$\omega_\mathrm{P}(\mathrm{U},\mathrm{V}) = \widetilde{\mathrm{tr}}(\mathrm{P}^{-1}\mathrm{U}\mathrm{P}^{-1}\mathrm{V}), \qquad (\mathrm{P} \in \mathscr{P}_\mathrm{J}, \mathrm{U},\mathrm{V} \in \mathfrak{p}_\mathrm{J}).$$

群 $\mathrm{GL}(n)$ 通过 $g \cdot \mathrm{P} = g\mathrm{P}g$ 作用于 $\mathscr{P}_\mathrm{J}$ 上（见㮾题 3.4）。型 ω 是 $\mathrm{GL}(n)$-不变的，即

$$\omega_{g\mathrm{X}g^\sharp}\left(g\mathrm{U}g^\sharp, g\mathrm{V}g^\sharp\right) = \omega_\mathrm{X}(\mathrm{U},\mathrm{V}) \qquad (\mathrm{X} \in \mathscr{P}_\mathrm{J}, \mathrm{U},\mathrm{V} \in \mathfrak{p}_\mathrm{J}).$$

事实上，

$$\begin{aligned}&\omega_{g\mathrm{X}g^\sharp}\left(g\mathrm{U}g^\sharp, g\mathrm{V}g^\sharp\right) = \widetilde{\mathrm{tr}}\left((g\mathrm{X}g^\sharp)^{-1}g\mathrm{U}g^\sharp(g\mathrm{X}g^\sharp)^{-1}g\mathrm{V}g^\sharp\right)\\ =\ &\widetilde{\mathrm{tr}}((g^\sharp)^{-1}\mathrm{X}^{-1}g^{-1}g\mathrm{U}g^\sharp(g^\sharp)^{-1}\mathrm{X}^{-1}g^{-1}g\mathrm{V}g^\sharp) = \widetilde{\mathrm{tr}}((g^\sharp)^{-1}\mathrm{X}^{-1}\mathrm{U}\mathrm{X}^{-1}\mathrm{V}g^\sharp)\\ =\ &\widetilde{\mathrm{tr}}(\mathrm{X}^{-1}\mathrm{U}\mathrm{X}^{-1}\mathrm{V}) = \omega_\mathrm{X}(\mathrm{U},\mathrm{V}).\end{aligned}$$

此外，按构造，型是正定的。事实上，$\langle\cdot,,\cdot\rangle$ 是正定的，并且对于所有非零 $\mathrm{U} \in \mathfrak{p}_\mathrm{J}$ 和 $\mathrm{P} \in \mathscr{P}_\mathrm{J}$，有

$$\omega_\mathrm{P}(\mathrm{U},\mathrm{U}) = \langle \mathrm{J}\mathrm{U},\mathrm{J}\mathrm{U}\rangle_{\mathrm{J}\mathrm{P}} > 0.$$

现在我们给出锥 $\mathscr{P}_\mathrm{J}$上的方程 (9) 的一个类似结果。

定理 5.2. 设 A、B $\in \mathscr{P}_\mathrm{J}$。映射 γ： $[0,1] \to \mathscr{P}_\mathrm{J}$ 定义为

$$\gamma(t) := \mathrm{A}_\mathrm{J}^{\frac{1}{2}} \bullet \left(\mathrm{A}_\mathrm{J}^{-\frac{1}{2}} \bullet \mathrm{B} \bullet \mathrm{A}_\mathrm{J}^{-\frac{1}{2}}\right)_\mathrm{J}^t \bullet \mathrm{A}_\mathrm{J}^{\frac{1}{2}}, \qquad (t \in [0,1]) \tag{11}$$

它是 $\mathscr{P}_\mathrm{J}$ 中连接A与B 的测地线方程，其中相对于 ω （特别地， $\gamma(0) = \mathrm{A}$ 且 $\gamma(1) = \mathrm{B}$）。

在证明定理 5.2 之前，我们先回顾以下引理。

引理 5.3. 设(M ,g) 和(N ,h) 是两个光滑黎曼流形，并设$\Phi : \mathrm{M} \to \mathrm{N}$ 是一个等距映射（即满足 $g = \Phi^* h$ 的微分同胚）。那么， Φ将M 中的测地线映射到N中的测地线。换言之， γ 是M 中的测地线，当且仅当 $\Phi \circ \gamma$ 是N中的测地线。

证明。 这是黎曼几何中的一个标准结果；参见 [17, 命题 5.4]。

□

定理 *5.2* 的证明。 令 δ 为$\mathscr{P}$中 JA 与 JB 之间的测地线方程。特别地，利用方程 (9)，我们有

$$\delta(t) = (\mathrm{JA})^{\frac{1}{2}} \left((\mathrm{JA})^{-\frac{1}{2}} (\mathrm{JB}) (\mathrm{JA})^{-\frac{1}{2}} \right)^t (\mathrm{JA})^{\frac{1}{2}} .$$

根据引理 5.3，只需证明对所有 $t \in [0,1]$， $\gamma(t) = \mathrm{J}\delta(t)$。由命题 3.3，我们有

$$\mathrm{X}_{\mathrm{J}}^t = \mathrm{J}(\mathrm{JX})^t , \qquad (\mathrm{X} \in \mathscr{P}_{\mathrm{J}} , t \in \mathbb{R}) .$$

因此，$\mathrm{A}^{\frac{1}{2}}\mathrm{J} = \mathrm{J}(\mathrm{JA})^{\frac{1}{2}}$且$\mathrm{A}^{-\frac{1}{2}}\mathrm{J} = \mathrm{J}(\mathrm{JA})^{-\frac{1}{2}}$。此外，

$$\begin{aligned} \mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}} \bullet \mathrm{B} \bullet \mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}} &= \left(\mathrm{J}(\mathrm{JA})^{-\frac{1}{2}} \right) (\mathrm{JBJ}) \left(\mathrm{J}(\mathrm{JA})^{-\frac{1}{2}} \right) = \mathrm{J}(\mathrm{JA})^{-\frac{1}{2}} (\mathrm{JBJJ}) (\mathrm{JA})^{-\frac{1}{2}} \\ &= \mathrm{J}(\mathrm{JA})^{-\frac{1}{2}} (\mathrm{JB}) (\mathrm{JA})^{-\frac{1}{2}} . \end{aligned}$$

Then

$$\begin{aligned} \gamma(t) &= \mathrm{A}_{\mathrm{J}}^{\frac{1}{2}} \bullet \left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}} \bullet \mathrm{B} \bullet \mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}} \right)_{\mathrm{J}}^t \bullet \mathrm{A}_{\mathrm{J}}^{\frac{1}{2}} \\ &= \mathrm{J}(\mathrm{JA})^{\frac{1}{2}} \mathrm{J}^2 \left(\mathrm{J}^2 (\mathrm{JA})^{-\frac{1}{2}} (\mathrm{JB}) (\mathrm{JA})^{-\frac{1}{2}} \right)^t \mathrm{J} \left(\mathrm{J}(\mathrm{JA})^{\frac{1}{2}} \right) \\ &= \mathrm{J}(\mathrm{JA})^{\frac{1}{2}} \left((\mathrm{JA})^{-\frac{1}{2}} (\mathrm{JB}) (\mathrm{JA})^{-\frac{1}{2}} \right)^t (\mathrm{JA})^{\frac{1}{2}} \\ &= \mathrm{J}\delta(t) , \end{aligned}$$

定理得证。

□

注 5.4。 特别地，可以看出，方程（11）中给出的映射 γ 满足方程（10）中给出的常微分方程。事实上，利用 $\gamma(t) = \mathrm{J}\delta(t)$ 以及$\ddot{\delta}(t) = \dot{\delta}(t)\delta(t)^{-1}\dot{\delta}(t)$，可得 $\dot{\gamma}(t) = \mathrm{J}\dot{\delta}(t)$， $\gamma(t)^{-1} = (\mathrm{J}\delta(t))^{-1} = \delta(t)^{-1}\mathrm{J}$。于是

$$\dot{\gamma}(t)\gamma(t)^{-1}\dot{\gamma}(t) = \left(\mathrm{J}\dot{\delta}(t) \right) \left(\delta(t)^{-1}\mathrm{J} \right) \left(\mathrm{J}\dot{\delta}(t) \right) = \mathrm{J}\dot{\delta}(t)\delta(t)^{-1}\dot{\delta}(t) = \mathrm{J}\ddot{\delta}(t) = \ddot{\gamma}(t) ,$$

这恰好是方程（10）中给出的常微分方程。

6. Geometric mean on the cone $\mathscr{P}_{\mathrm{J}}$

本节首先回顾 [15] 关于 $\mathscr{P}$ 上几何平均的主要结果（请注意， [15] 的结果适用于 $\mathbb{D} \in \{\mathbb{R}, \mathbb{C}\}$；四元数情形的类似结果见

附录 A）。对于两个矩阵 A、B $\in \mathscr{P}$ 和 $t \in [0,1]$，我们用 $A\sharp_t B$ 表示 Mat($n \times n$) 中由下式给出的矩阵

$$A\sharp_t B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}},$$

且令 $A\sharp B = A\sharp_{\frac{1}{2}} B$。如方程 (9) 所述，对所有 $t \in [0,1]$，$A\sharp_t B \in \mathscr{P}$。

注记 6.1。以下是 $\sharp_t$ 在 $\mathscr{P}$ 上满足的一些良好性质：

(1) 对所有 A、B $\in \mathscr{P}$ 以及 $t \in (0,1)$，$A\sharp_t A = A$，且 $A\sharp_t B = A$ 当且仅当 $A = B$。(2) 对所有 A、B $\in \mathscr{P}$，$A\sharp B$ 是里卡蒂方程的唯一解

$$XA^{-1}X = B,$$

(3) 对所有 A、B $\in \mathscr{P}$ 以及 C $\in GL(n)$，有

$$(C^*AC)\sharp(C^*BC) = C^*(A\sharp B)C.$$

(4) 对所有满足 $A_1, A_2, B_1, B_2 \in \mathscr{P}$ 的 A 和 B，使得 $A_1 \le A_2$ 且 $B_1 \le B_2$，，可得

$$A_1\sharp B_1 \le A_2\sharp B_2.$$

(5) 对所有 A、B $\in \mathscr{P}$，则 $(A\sharp B)^{-1} = A^{-1}\sharp B^{-1}$。

本节的目标是在我们的锥 $\mathscr{P}_J$ 上给出几何平均的一个类似物。我们从定义开始。

定义 6.2. 设 A, B $\in \mathscr{P}_J$。对所有 $t \in [0,1]$，我们将 $A\sharp_t^J B$ 记为 Mat($n \times n$) 中由下式给出的元素

$$A\sharp_t^J B := A_J^{\frac{1}{2}} \bullet \left(A_J^{-\frac{1}{2}} \bullet B \bullet A_J^{-\frac{1}{2}}\right)_J^t \bullet A_J^{\frac{1}{2}},$$

并令 $A\sharp^J B = A\sharp_{\frac{1}{2}}^J B$。

注 6.3. (1) 如定理 5.2 所述，对所有 A, B $\in \mathscr{P}_J$，对所有 $t \in [0,1]$，我们得到 $A\sharp_t^J B \in \mathscr{P}_J$。

(2) 由定理 5.2 的证明可知，对所有 A, B $\in \mathscr{P}_J$，我们有

$$A\sharp_t^J B = \Phi_J^{-1}\left(\Phi_J(A)\sharp_t\Phi_J(B)\right), \qquad (t \in [0,1]).$$

命题 6.4（黎卡提方程）**.** 设 A, B $\in \mathscr{P}_J$。则 $A\sharp^J B$ 是方程在 $\mathscr{P}_J$ 上的唯一解

$$XA^{-1}X = B. \tag{12}$$

证明。首先，我们证明 $A\sharp^J B$ 确实是 (12) 的一个解。根据命题 3.3，可得 $A_J^1 J A_J^{-1} = J$（即 $A^{-1} = JA_J^{-1}J$）以及 $A_J^{-1} = A^{-\frac{1}{2}}J\,JA^{-\frac{1}{2}}J$。因此

$$XA^{-1}X = XJA_J^{-1}JX = XJA_J^{-\frac{1}{2}}JA_J^{-\frac{1}{2}}JX.$$

Thus

$$\left(A\sharp^J B\right)A^{-1}\left(A\sharp^J B\right) = \left(A\sharp^J B\right)JA_J^{-\frac{1}{2}}JA_J^{-\frac{1}{2}}J\left(A\sharp^J B\right)$$

$$= \quad A_J^{\frac{1}{2}} J\left(A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}}\right)_J^{\frac{1}{2}} J\left(A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}}\right)_J^{\frac{1}{2}} JA_J^{\frac{1}{2}} = A_J^{\frac{1}{2}} J\left(A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}}\right) JA_J^{\frac{1}{2}}$$

$$= \quad \left(A_J^{\frac{1}{2}} JA_J^{-\frac{1}{2}}\right) JBJ\left(A_J^{-\frac{1}{2}} JA_J^{\frac{1}{2}}\right) = J^2BJ^2 = B\,,$$

其中，最后两个等式由命题 3.3 得出。
我们现在证明该解的唯一性。假设 X 满足

$$XJA_J^{-\frac{1}{2}} JA_J^{-\frac{1}{2}} JX = B\,.$$

Then

$$\left(A_J^{-\frac{1}{2}} JXJA_J^{-\frac{1}{2}}\right) J\left(A_J^{-\frac{1}{2}} JXJA_J^{-\frac{1}{2}}\right) = A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}}\,.$$

利用命题 3.3 的第 (4) 部分，我们有

$$\left(A_J^{-\frac{1}{2}} JXJA_J^{-\frac{1}{2}}\right)_J^2 = A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}}\,.$$

现在，我们对两边取 J 平方根，得到

$$A_J^{-\frac{1}{2}} JXJA_J^{-\frac{1}{2}} = (A_J^{-\frac{1}{2}} JBJA_J^{-\frac{1}{2}})_J^{\frac{1}{2}}\,.$$

为了在保持位于 $\mathscr{P}_J$ 内的同时求解 X，我们应用合同映射 $Y \to CYC^{\sharp_1}$，其中 $C = A^2 \sharp J$，该映射保持 $\mathscr{P}_J$。最后

$$J^2XJ^2 = A\sharp^J B\,,$$

即 $X = A\sharp^J B$。 □

我们现在证明 $\sharp^J$ 的一些性质。

引理 6.5. 对于所有的 $A, B \in \mathscr{P}_J$，我们得到

(1) $甲\sharp^J乙 = 乙\sharp^J甲$, *(2)* $\left(甲\sharp^J乙\right)^{-1} = 甲^{-1}\sharp^J乙^{-1}$.

证明. （1）对 (12) 两边取逆，得到

$$X^{-1}AX^{-1} = B^{-1}\,.$$

在等式两边分别左乘和右乘 B，得到

$$\left(BX^{-1}B\right)B^{-1}\left(BX^{-1}B\right) = A\,.$$

令 $Y := BX^{-1}B$。由于 $\mathscr{P}_J$ 在求逆和合同运算下保持不变，我们有 $Y \in \mathscr{P}_J$。因此，Y 是黎卡提方程 $ZB^{-1}Z = A$ 在 $\mathscr{P}_J$ 中的一个解。由 $\mathscr{P}_J$中解的唯一性可知,

$$B\sharp^J A = Y = BX^{-1}B\,.$$

将 A 和 B 互换并应用同样的论证可得 $X = A\left(B\sharp^J A\right)^{-1} A$，因此

$$A\sharp^J B = B\sharp^J A\,.$$

(2) 再次从方程（12）出发，取逆可得

$$X^{-1}AX^{-1} = B^{-1}.$$

等价地，

$$X^{-1}\left(A^{-1}\right)^{-1}X^{-1} = B^{-1}.$$

因此，$X^{-1} \in \mathscr{P}_J$ 是与对 $\left(A^{-1}, B^{-1}\right)$ 相关的黎卡提方程的一个解。

由 $\mathscr{P}_J$ 中解的唯一性可知，

$$A^{-1}\sharp^J B^{-1} = X^{-1} = \left(A\sharp^J B\right)^{-1}.$$

证明至此完成。

□

引理 6.6. 设 $A,B \in \mathscr{P}_J$. 则

$$A\sharp^J B = \max\left\{X \in \mathscr{P}_J, \begin{bmatrix} JA & JX \\ JX & JB \end{bmatrix} \geq 0\right\},$$

其中，最大值是相对于J-*Loewner* 序 $\preceq_J$ 而言的。

证明。 注意到.....的舒尔补

$$\begin{bmatrix} JA & JX \\ JX & JB \end{bmatrix}$$

是 $JB-(JX)(JA)^{-1}(JX) = J\left(B - XA^{-1}X\right)$。当 $X = A\sharp^J B$ 时，舒尔补为零。由于 $JA > 0$，因此有

$$\begin{bmatrix} JA & JX \\ JX & JB \end{bmatrix} \geq 0.$$

另一方面，若 $XA^{-1}X \preceq_J B$，则

$$JB \geq (JX)(JA)^{-1}(JX).$$

这意味着

$$(JA)^{-\frac{1}{2}}(JB)(JA)^{-\frac{1}{2}} \geq \left((JA)^{-\frac{1}{2}}(JX)(JA)^{-\frac{1}{2}}\right)^2.$$

根据平方根的算子单调性，并在两侧乘以 $(JA)^{\frac{1}{2}}$,，可得 $JX \preceq (JA)\sharp(JB)$。因此 $X \preceq_J A\sharp^J B$。

□

命题 6.7. 设 $t \in (0,1)$。

(1) 对所有 $A \in \mathscr{P}_J$，有 $A\sharp_t^J A = A$。此外，$A\sharp_t^J B = A$ 当且仅当 $A = B$。

(2) 对所有 $A, B \in \mathscr{P}_J$ 和 $\alpha, \beta > 0$，有

$$(\alpha A)\sharp_t^J(\beta B) = \alpha^{1-t}\beta^t A\sharp_t^J B.$$

（*3*）对于所有A、$B \in \mathscr{P}_J$，有

$$A\sharp_t^J B = B\sharp_{1-t}^J A.$$

(4) If $\mathrm{A},\mathrm{B},\mathrm{C},\mathrm{D}\in\mathscr{P}_{\mathrm{J}}$ *are such that* $\mathrm{A}\preceq_{\mathrm{J}}\mathrm{C}$ *and* $\mathrm{B}\preceq_{\mathrm{J}}\mathrm{D}$*, then*

$$\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\preceq_{\mathrm{J}}\mathrm{C}\sharp_t^{\mathrm{J}}\mathrm{D}.$$

证明。 回顾一下，加权 J-几何平均定义为

$$\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}:=\mathrm{J}^{-1}\left((\mathrm{JA})\sharp_t(\mathrm{JB})\right),\qquad(\mathrm{A},\mathrm{B}\in\mathscr{P}_{\mathrm{J}}),$$

其中， $\sharp_t$ 表示 $\mathscr{P}$上的通常加权几何平均。由于 $\mathrm{J}^{-1}=\mathrm{J}$，这等价于

$$\mathrm{J}\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right)=(\mathrm{JA})\sharp_t(\mathrm{JB}).\tag{13}$$

我们通过公式 (13) 将相应的标准性质从 $\mathscr{P}$传递到 $\mathscr{P}_{\mathrm{J}}$，从而证明每一项。

(1) 设 $\mathrm{A}\in\mathscr{P}_{\mathrm{J}}$。利用公式 (13) 和备注 6.1，我们得到

$$\mathrm{J}\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{A}\right)=(\mathrm{JA})\sharp_t(\mathrm{JA})=\mathrm{JA},$$

因此 $\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{A}=\mathrm{A}$。

现在假设 $\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}=\mathrm{A}$。两边乘以 J 并利用公式 (13)，得到

$$(\mathrm{JA})\sharp_t(\mathrm{JB})=\mathrm{JA}.$$

同样，由备注 6.1 可知 $\mathrm{JA}=\mathrm{JB}$，即 $\mathrm{A}=\mathrm{B}$。

(2) 设 α、$\beta>0$，且 A、B $\in\mathscr{P}_{\mathrm{J}}$。首先注意到

$$(\alpha\mathrm{A})_{\mathrm{J}}^{\frac{1}{2}}=\alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}},\qquad(\alpha\mathrm{A})_{\mathrm{J}}^{-\frac{1}{2}}=\alpha^{-\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}.$$

我们得到

$$\begin{aligned}(\alpha\mathrm{A})\sharp_t^{\mathrm{J}}(\beta\mathrm{B}) &= (\alpha\mathrm{A})_{\mathrm{J}}^{\frac{1}{2}}\,\mathrm{J}\left((\alpha\mathrm{A})_{\mathrm{J}}^{-\frac{1}{2}}\,\mathrm{J}(\beta\mathrm{B})\,\mathrm{J}(\alpha\mathrm{A})_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^{t}\mathrm{J}(\alpha\mathrm{A})_{\mathrm{J}}^{\frac{1}{2}}\\ &= \alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\mathrm{J}\left(\alpha^{-\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}(\beta\mathrm{JB})\,\alpha^{-\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^{t}\mathrm{J}\alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\\ &= \alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\mathrm{J}\left(\frac{\beta}{\alpha}\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{JBJA}_{\mathrm{J}}^{-\frac{1}{2}}\right)\right)_{\mathrm{J}}^{t}\mathrm{J}\alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}.\end{aligned}$$

利用 $(\lambda\mathrm{A})_{\mathrm{J}}^{t}=\lambda^{t}\mathrm{A}_{\mathrm{J}}^{t}$，我们得到

$$\left(\frac{\beta}{\alpha}\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{JBJA}_{\mathrm{J}}^{-\frac{1}{2}}\right)\right)_{\mathrm{J}}^{t}=\left(\frac{\beta}{\alpha}\right)^{t}\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{JBJA}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^{t}.$$

因此，

$$(\alpha\mathrm{A})\sharp_t^{\mathrm{J}}(\beta\mathrm{B})=\alpha^{\frac{1}{2}}\left(\frac{\beta}{\alpha}\right)^{t}\alpha^{\frac{1}{2}}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\mathrm{J}\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{JBJA}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^{t}\mathrm{JA}_{\mathrm{J}}^{\frac{1}{2}}=\alpha^{1-t}\beta^{t}\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right),$$

这就证明了 (2)。

(3) 考虑曲线

$$\gamma(t):=\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B},\qquad(t\in[0,1]).$$

根据构造， $\gamma(0)=\mathrm{A}$ 且 $\gamma(1)=\mathrm{B}$。令 $\widetilde{\gamma}(t):=\gamma(1-t)$。于是$\widetilde{\gamma}(0)=\mathrm{B}$ 且 $\widetilde{\gamma}(1)=\mathrm{A}$。此外， $\widetilde{\gamma}$ 满足与 γ相同的测地线方程。根据连接给定两个端点的 $\mathscr{P}_{\mathrm{J}}$ 中测地线的唯一性，

可知 $\widetilde{\gamma}$ 与测地线$t \mapsto \mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{A}$ 重合。因此，对每个$t \in [0,1]$,

$$\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{A} = \widetilde{\gamma}(t) = \gamma(1-t) = \mathrm{A}\sharp_{\mathrm{J}}^{1-t}\mathrm{B}\,,$$

这等价于（将 t 替换为 $1-t$ 之后）

$$\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B} = \mathrm{B}\sharp_{\mathrm{J}}^{1-t}\mathrm{A}\,.$$

这就证明了 (3)。

(4) 假设 $\mathrm{A} \preceq_{\mathrm{J}} \mathrm{C}$ 且 $\mathrm{B} \preceq_{\mathrm{J}} \mathrm{D}$。根据 $\preceq_{\mathrm{J}}$ 的定义，这等价于在 $\mathscr{P}$ 上通常的 Loewner 序中 $\mathrm{JA} \preceq \mathrm{JC}$ 且 $\mathrm{JB} \preceq \mathrm{JD}$。利用备注 6.1，可得

$$(\mathrm{JA})\sharp_t(\mathrm{JB}) \preceq (\mathrm{JC})\sharp_t(\mathrm{JD})\,.$$

因此

$$\mathrm{J}\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right) \preceq \mathrm{J}\left(\mathrm{C}\sharp_t^{\mathrm{J}}\mathrm{D}\right),$$

这恰好就是 $\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B} \preceq_{\mathrm{J}} \mathrm{C}\sharp_t^{\mathrm{J}}\mathrm{D}$。

□

命题 6.8. *(1)* 对所有 $g \in \mathrm{K}_{\mathrm{J}}$、$t \in [0,1]$ 以及 $\mathrm{A},\mathrm{B} \in \mathscr{P}_{\mathrm{J}}$，我们有

$$\left(g\mathrm{A}g^{\sharp}\right)\sharp_t^{\mathrm{J}}\left(g\mathrm{B}g^{\sharp}\right) = g\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right)g^{\sharp}\,.$$

(2) 对任意A，B，C，$\mathrm{D} \in \mathscr{P}_{\mathrm{J}}$ 以及 t，$s \in [0,1]$，有

$$(1-s)\,\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{C} + s\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{D} \preceq_{\mathrm{J}} ((1-s)\,\mathrm{A} + s\mathrm{B})\sharp_t^{\mathrm{J}}((1-s)\,\mathrm{C} + s\mathrm{D})\,.$$

(3) 对任意A，$\mathrm{B} \in \mathscr{P}_{\mathrm{J}}$ 以及 t，s，$u \in [0,1]$,

$$\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right)\sharp_u^{\mathrm{J}}\left(\mathrm{A}\sharp_s^{\mathrm{J}}\mathrm{B}\right) = \mathrm{A}\sharp_{(1-u)t+us}^{\mathrm{J}}\mathrm{B}\,.$$

(4) 对任意A，$\mathrm{B} \in \mathscr{P}_{\mathrm{J}}$ 以及 $t \in [0,1]$，得到

$$[(1-t)\mathrm{A}_{\mathrm{J}}^{-1} + t\mathrm{B}_{\mathrm{J}}^{-1}]^{-1} \preceq_{\mathrm{J}} \mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B} \preceq_{\mathrm{J}} (1-t)\mathrm{A} + t\mathrm{B}\,.$$

证明。(1) 设$g \in \mathrm{K}_{\mathrm{J}}$，并设 $\widetilde{\mathrm{A}} = g\mathrm{A}g^{\sharp}$ 以及 $\widetilde{\mathrm{B}} = g\mathrm{B}g^{\sharp}$。利用命题 3.4，对于每个 $\mathrm{X} \in \mathscr{P}_{\mathrm{J}}$ 以及$t \in \mathbb{R}$，有 $(g\mathrm{X}g^{\sharp})_{\mathrm{J}}^t = g\mathrm{X}_{\mathrm{J}}^t g^{\sharp}$。特别地，

$$\widetilde{\mathrm{A}}_{\mathrm{J}}^{\pm\frac{1}{2}} = g\mathrm{A}_{\mathrm{J}}^{\pm\frac{1}{2}}g^{\sharp}\,.$$

根据 $\sharp_t^{\mathrm{J}}$ 的定义（见 6.2），有

$$\widetilde{\mathrm{A}}\sharp_t^{\mathrm{J}}\widetilde{\mathrm{B}} = \widetilde{\mathrm{A}}_{\mathrm{J}}^{\frac{1}{2}}\left(\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}}\widetilde{\mathrm{B}}\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^t\widetilde{\mathrm{A}}_{\mathrm{J}}^{\frac{1}{2}}\,.$$

利用$g^{\sharp}g = \mathrm{Id}_n$，得到

$$\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}}\widetilde{\mathrm{B}}\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}} = \left(g\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}g^{\sharp}\right)\left(g\mathrm{B}g^{\sharp}\right)\left(g\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}g^{\sharp}\right) = g\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{B}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\right)g^{\sharp}\,.$$

最后，利用这一点

$$\left(\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}}\widetilde{\mathrm{B}}\widetilde{\mathrm{A}}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^t = g\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{B}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^t g^{\sharp}\,,$$

由此可得

$$\begin{aligned}\widetilde{\mathrm{A}}\sharp_t^{\mathrm{J}}\widetilde{\mathrm{B}} &= \left(g\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}g^{\sharp}\right)\left(g(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{B}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}})_{\mathrm{J}}^{t}g^{\sharp}\right)\left(g\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}g^{\sharp}\right)\\ &= g\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\left(\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\mathrm{B}\mathrm{A}_{\mathrm{J}}^{-\frac{1}{2}}\right)_{\mathrm{J}}^{t}\mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}g^{\sharp}\\ &= g\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{B}\right)g^{\sharp}.\end{aligned}$$

(2) 利用注记 6.3，可得

$$\mathrm{J}\left(\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{C}\right) = (\mathrm{JA})\,\sharp_t\,(\mathrm{JC})\,, \qquad \mathrm{J}\left(\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{D}\right) = (\mathrm{JB})\,\sharp_t\,(\mathrm{JD})\,.$$

因此，

$$\mathrm{J}\left((1-s)\,\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{C} + s\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{D}\right) = (1-s)\mathrm{JA}\sharp_t\mathrm{JC} + s\mathrm{JB}\sharp_t\mathrm{JD}\,.$$

类似地，可得

$$\mathrm{J}\left(((1-s)\,\mathrm{A}+s\mathrm{B})\,\sharp_t^{\mathrm{J}}\,((1-s)\,\mathrm{C}+s\mathrm{D})\right) = ((1-s)\,\mathrm{JA}+s\mathrm{JB})\,\sharp_t\,((1-s)\,\mathrm{JC}+s\mathrm{JD})\,.$$

利用几何平均在 $\mathscr{P}$ 上的凹性（见 [15]），我们有

$$(1-s)\mathrm{JA}\sharp_t\mathrm{JC} + s\mathrm{JB}\sharp_t\mathrm{JD} \le ((1-s)\,\mathrm{JA}+s\mathrm{JB})\,\sharp_t\,((1-s)\,\mathrm{JC}\;+s\mathrm{JD})\,.$$

结合前面的等式，可得

$$\mathrm{J}\left((1-s)\,\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{C} + s\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{D}\right) \le \mathrm{J}\left(((1-s)\,\mathrm{A}+s\mathrm{B})\,\sharp_t^{\mathrm{J}}\,((1-s)\,\mathrm{C}+s\mathrm{D})\right),$$

i.e.

$$(1-s)\mathrm{A}\sharp_t^{\mathrm{J}}\mathrm{C} + s\mathrm{B}\sharp_t^{\mathrm{J}}\mathrm{D} \le_{\mathrm{J}} ((1-s)\,\mathrm{A}+s\mathrm{B})\,\sharp_t^{\mathrm{J}}\,((1-s)\,\mathrm{C}+s\mathrm{D})\,.$$

(3) 使用与 (2) 中类似的论证，可由 [15] 得出。

(4) 这可由正定矩阵的相应不等式推出。

□

注 6.9. 这里有必要作一条说明。鉴于对于可交换的正定矩阵 X、Y $\in \mathscr{P}$，我们有

$$\mathrm{X}\sharp_t\mathrm{Y} = \mathrm{X}^{1-t}\mathrm{Y}^{t}, \qquad (t\in[0\,,1])\,. \tag{14}$$

自然会问，两个可交换的 J-正定厄米矩阵是否满足这样的等式。也就是说，若 A、B $\in \mathscr{P}_{\mathrm{J}}$ 满足 AB $=$BA，是否有 A$\sharp_t^{\mathrm{J}}$B $=\mathrm{A}_{\mathrm{J}}^{1-t}\mathrm{B}_{\mathrm{J}}^{t}$?

令 J $=\mathrm{diag}(1\,,-1)$（即$p=q=1$）。容易看出，矩阵

$$\mathrm{A} = \begin{bmatrix} 2 & 1 \\ -1 & -2 \end{bmatrix} \qquad \text{and} \qquad \mathrm{B} = \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix}$$

都是 $\mathscr{P}_{\mathrm{J}}$ 中的 J-正矩阵。此外，A 和 B 在 Mat(2×2) 中可交换。然而，我们得到

$$\mathrm{A}\sharp^{\mathrm{J}}\mathrm{B} - \mathrm{A}_{\mathrm{J}}^{\frac{1}{2}}\mathrm{B}_{\mathrm{J}}^{\frac{1}{2}} \approx \begin{bmatrix} 0.263207 & 0.768429 \\ -0.857469 & -2.50336 \end{bmatrix}.$$

这是 J-正定矩阵的几何平均与正定矩阵的几何平均之间的一个关键差异。

公式（14）在 $\mathscr{P}_{\mathrm{J}}$ 上的恰当对应形式由以下引理给出，并且与 $\bullet$ 所提供的代数结构一致。

引理 6.10. 对于所有满足 $A, B \in \mathscr{P}_{\mathrm{J}}$ 且使得 $A \bullet B = B \bullet A$ 的情形，我们有

$$A \sharp_t^{\mathrm{J}} B = A_{\mathrm{J}}^{1-t} \bullet B_{\mathrm{J}}^{t}.$$

证明。设 $A, B \in \mathscr{P}_{\mathrm{J}}$ 满足 $AJB = BJA$。令 $P := JA$，$Q := JB \in \mathscr{P}$。由于 $J^2 = \mathrm{Id}_n$，交换关系推出

$$PQ = (JA)(JB) = J(AJB) = J(BJA) = (JB)(JA) = QP.$$

因此 P 和 Q 交换。由此，根据注 6.9，对所有 $t \in [0,1]$

$$P \sharp_t Q = P^{1-t} Q^{t}.$$

因此

$$A \sharp_t^{\mathrm{J}} B = J^{-1} P^{1-t} Q^{t} = J^{-1} (JA)^{1-t} (JB)^{t}.$$

在两个因子之间插入 $J^2 = \mathrm{Id}_n$，并利用 $X_{\mathrm{J}}^{t} := J(JX)^{t}$，可得

$$J^{-1} (JA)^{1-t} (JB)^{t} = \left(J(JA)^{1-t}\right) J \left(J(JB)^{t}\right) = A_{\mathrm{J}}^{1-t} J B_{\mathrm{J}}^{t},$$

引理得证。

□

加权几何平均还满足另一项重要性质，即著名的 Ando-Hiai 不等式 [1],，该不等式指出，对于 $A,B \in \mathscr{P}$

$$A \sharp_t B \le \mathrm{Id}_n \implies A^r \sharp_t B^r \le \mathrm{Id}_n, \qquad (r \ge 1). \tag{15}$$

命题 6.11（J-安藤–日井不等式）。对于所有 $A, B \in \mathscr{P}_{\mathrm{J}}$ 以及 $r \ge 1$，我们有

$$A \sharp_t^{\mathrm{J}} B \le_{\mathrm{J}} J \implies A_{\mathrm{J}}^{r} \sharp_t^{\mathrm{J}} B_{\mathrm{J}}^{r} \le_{\mathrm{J}} J.$$

证明。注意，$A \sharp_t^{\mathrm{J}} B \le_{\mathrm{J}} J$ 当且仅当 $JA \sharp_t^{\mathrm{J}} B \le \mathrm{Id}_n$。然而，

$$JA \sharp_t^{\mathrm{J}} B = (JA) \sharp_t (JB) \le \mathrm{Id}_n,$$

推出

$$(JA)^r \sharp_t (JB)^r \le \mathrm{Id}_n,$$

由 (15) 得。因此，$\left[J(A)_{\mathrm{J}}^{r}\right] \sharp_t \left[J(B)_{\mathrm{J}}^{r}\right] \le \mathrm{Id}_n$，或者等价地，$J(A_{\mathrm{J}}^{r} \sharp_t B_{\mathrm{J}}^{r}) \le \mathrm{Id}_n$。这便在 J-序下得到了所需结果。

□

本文最后给出 J-安藤–日井不等式的一个推论，即 J-Furuta 不等式。证明是对文献 [10]中某个证明在此情形下的改编。

命题 6.12 (J-Furuta 不等式)。对于所有的 $A, B \in \mathscr{P}_{\mathrm{J}}$、$p \ge 0$ 和 $r \ge 1$，我们有 $0 <_{\mathrm{J}} B \le_{\mathrm{J}} A$ 蕴含

$$\left(A_{\mathrm{J}}^{\frac{r}{2}} \bullet B_{\mathrm{J}}^{p} \bullet A_{\mathrm{J}}^{\frac{r}{2}}\right)_{\mathrm{J}}^{\frac{r}{r+p}} \le_{\mathrm{J}} A_{\mathrm{J}}^{r}.$$

证明。 由引理 4.6 可知，$\mathrm{B} \leq_{\mathrm{J}} \mathrm{A}$ 蕴含 $\mathrm{A}_{\mathrm{J}}^{-1} \leq \mathrm{B}_{\mathrm{J}}^{-1}$。根据命题6.7的第(4)部分，我们有

$$\mathrm{A}_{\mathrm{J}}^{-1} \sharp_t^{\mathrm{J}} \mathrm{B}_{\mathrm{J}}^{q} \leq_{\mathrm{J}} \mathrm{B}_{\mathrm{J}}^{t(1+q)-1},$$

每当 $q \geq 0$ 且 $t \in [0,1]$时。特别地，取 $t = \frac{r}{p+r}$ 且 $q = \frac{p}{r}$。于是

$$\mathrm{A}_{\mathrm{J}}^{-1} \sharp_{\frac{r}{p+r}}^{\mathrm{J}} \mathrm{B}_{\mathrm{J}}^{\frac{p}{r}} \leq_{\mathrm{J}} \mathrm{J}.$$

这里，我们用到了 $\mathrm{B}_{\mathrm{J}}^{0} = \mathrm{J}$ 这一事实。因此，我们可以利用命题 6.11 和命题 3.3 的第 (2) 部分，得到

$$\mathrm{A}_{\mathrm{J}}^{-r} \sharp_{\frac{r}{p+r}}^{\mathrm{J}} \mathrm{B}_{\mathrm{J}}^{p} \leq_{\mathrm{J}} \mathrm{J}.$$

此时，我们回顾一下

$$\mathrm{A}_{\mathrm{J}}^{-r} \sharp_{\frac{r}{p+r}}^{\mathrm{J}} \mathrm{B}_{\mathrm{J}}^{p} = \mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}} \bullet \left(\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{B}_{\mathrm{J}}^{p} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \right)_{\mathrm{J}}^{\frac{r}{p+r}} \bullet \mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}}.$$

因此，

$$\mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}} \bullet \left(\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{B}_{\mathrm{J}}^{p} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \right)_{\mathrm{J}}^{\frac{r}{p+r}} \bullet \mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}} \leq_{\mathrm{J}} J.$$

,由于 $\mathrm{A}^2\mathrm{J}$ 是 J-厄米的，

$$\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}} \bullet \left(\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{B}_{\mathrm{J}}^{p} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \right)_{\mathrm{J}}^{\frac{r}{p+r}} \bullet \mathrm{A}_{\mathrm{J}}^{-\frac{r}{2}} \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \leq_{\mathrm{J}} \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}},$$

这意味着

$$\left(\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{B}_{\mathrm{J}}^{p} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \right)_{\mathrm{J}}^{\frac{r}{p+r}} = \mathrm{J} \bullet \left(\mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \bullet \mathrm{B}_{\mathrm{J}}^{p} \bullet \mathrm{A}_{\mathrm{J}}^{\frac{r}{2}} \right)_{\mathrm{J}}^{\frac{r}{p+r}} \bullet \mathrm{J} \leq_{\mathrm{J}} \mathrm{A}_{\mathrm{J}}^{r},$$

如所要求。

□

Appendix A. The quaternionic case

A.1. **除代数** $\mathbb{H}$。 我们用 $\mathbb{H}$ 表示由以下元素生成的实结合代数

$$\mathbb{H} := \{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}, a, b, c, d \in \mathbb{R}\}$$

1、**i**、**j**、**k**，并满足以下关系

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

代数 $\mathbb{H}$ 不是交换的（例如**ij** = –**ji**），其中心 Z($\mathbb{H}$) 是实数 $\mathbb{R}$ 的集合。

注 A.1。 利用 **k** = **ij**，每个 $q \in \mathbb{H}$都可以唯一地写成

$$q = (a + b\mathbf{i}) + (c + d\mathbf{i})\,\mathbf{j},$$

i.e.

$$q = z_1 + z_2\mathbf{j}, \qquad (z_1, z_2 \in \mathbb{C}),$$

其中 $\mathbb{C} = \mathbb{R} \oplus \mathbb{R}\mathbf{i}$ 被视为由 1 和 **i**生成的 $\mathbb{H}$ 子代数。

记号 A.2. 对于四元数 $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$，分别用 $\mathrm{Re}(q)$ 和 $\mathrm{Im}(q)$ 表示 q 的实部和虚部，其定义如下

$$\mathrm{Re}(q) = a\,, \qquad \mathrm{Im}(q) = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\,.$$

此外，我们用 $\mathrm{Im}(\mathbb{H})$ 表示由下式给出的 $\mathbb{H}$ 子集

$$\mathrm{Im}(\mathbb{H}) := \{b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\,, b\,, c\,, d \in \mathbb{R}\}\,.$$

元素$q \in \mathbb{H}$ 的四元数共轭q 定义为

$$\overline{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}\,,$$

相应的映射为 ι

$$\iota : \mathbb{H} \ni q \to \overline{q} \in \mathbb{H} \tag{16}$$

是 $\mathbb{R}$ 上的线性对合。等价地，若$q = z_1 + z_2\mathbf{j}$，则q 的共轭为

$$\overline{q} = \overline{z_1} - \overline{z_2}\mathbf{j}\,.$$

相应的范数为

$$|q| = \sqrt{q\overline{q}} = \sqrt{a^2 + b^2 + c^2 + d^2}\,,$$

且对所有$p\,,q \in \mathbb{H}$，均满足 $|pq| = |p|\,|q|$。

注记 A.3。四元数还可以用另一种优美而实用的形式表示。对于所有$q \in \mathbb{H}$，存在一个三元组$(a\,,b\,,u) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathrm{Im}(\mathbb{H})$，使得$q = a + bu$，其中$u^2 = -1$；当$b > 0$时，该表示唯一。

事实上，设$q = \mathrm{Re}(q) + \mathrm{Im}(q) \in \mathbb{H}$。若$\mathrm{Im}(q) = 0$，则结论显然成立（$b = 0$）。现在假设$\mathrm{Im}(q) \neq 0$。令$a = \mathrm{Re}(q)$，$b = |\mathrm{Im}(q)|$，并令$u = \mathrm{Im}(q)\,|\mathrm{Im}(q)|$。~~特别~~地，我们有$a \in \mathbb{R}$,$b \geq 0$，并且

$$q = \mathrm{Re}(q) + \mathrm{Im}(q) = \mathrm{Re}(q) + |\mathrm{Im}(q)|\,\frac{\mathrm{Im}(q)}{|\mathrm{Im}(q)|} = a + bu\,.$$

利用 $|u| = 1$ 以及 $\bar{}u = -u$ （因为u 是纯虚的），得到

$$u^2 = -u\overline{u} = -\,|u|^2 = -1\,.$$

最后，通过映射$q = z_1 + z_2\mathbf{j}$ 进行的同一化 $\mathbb{H} \simeq \mathbb{C}^2$ 诱导出一个 $\mathbb{R}$-代数的单射嵌入

$$\Psi : \mathbb{H} \hookrightarrow \mathrm{Mat}(2\,, \mathbb{C})\,, \qquad z_1 + z_2\mathbf{j} \mapsto \begin{bmatrix} z_1 & z_2 \\ -\overline{z}_2 & \overline{z}_1 \end{bmatrix}\,, \tag{17}$$

从而将四元数共轭等同于厄米共轭转置，并将四元数范数等同于行列式。

注记 A.4. 利用公式 (17) 中定义的嵌入 Ψ，我们将 $\mathbb{H}$上的约化迹定义为映射

$$\mathrm{trd} : \mathbb{H} \mapsto \mathbb{R}\,, \qquad \mathrm{trd}(q) := \frac{1}{2}\mathrm{tr}\,(\Psi(q))\,, \tag{18}$$

其中，tr 表示 $\mathrm{Mat}(2,\mathbb{C})$ 中通常的矩阵迹。更确切地说，若$q = z_1 + z_2\mathbf{j}$，其中$z_1, z_2 \in \mathbb{C}$，则

$$\mathrm{trd}(q) = \frac{1}{2}(z_1 + \overline{z_1}) = \mathrm{Re}(z_1) \in \mathbb{R}.$$

特别地，约化迹与四元数的实部相同，即

$$\mathrm{trd}(q) = \frac{1}{2}(q + \overline{q}).$$

引理 A.5. 约化迹 trd 在四元数共轭下保持不变（即 $\mathrm{trd}(q) = \mathrm{trd}(q)$），并且对所有 $p, q \in \mathbb{H}$，有

$$\mathrm{trd}(pq) = \mathrm{trd}(qp).$$

证明. 对于 $q \in \mathbb{H}$，有

$$\mathrm{trd}(\overline{q}) = \frac{1}{2}\left(\overline{q} + \overline{\overline{q}}\right) = \frac{1}{2}(\overline{q} + q) = \mathrm{trd}(q).$$

设 $p, q \in \mathbb{H}$。我们先证明 $\mathrm{Re}(pq) = \mathrm{Re}(qp)$。事实上，

$$\mathrm{Re}(pq) = \frac{1}{2}(pq + \overline{pq}) = \frac{1}{2}(pq + \bar{q}\bar{p}),$$

and

$$\mathrm{Re}(qp) = \frac{1}{2}(qp + \overline{qp}) = \frac{1}{2}(qp + \bar{p}\bar{q}).$$

现在将 $p = a + u$ 以及 $q = b + v$写作，其中 $a, b \in \mathbb{R}$ $u, v \in \mathrm{Im}(\mathbb{H})$。于是

$$pq = ab + av + bu + uv, \qquad \bar{q}\bar{p} = (b - v)(a - u) = ab - au - bv + vu,$$

i.e.

$$pq + \bar{q}\bar{p} = 2ab + (uv + vu).$$

同理可得

$$qp + \bar{p}\bar{q} = 2ab + (uv + vu).$$

因此 $pq + \bar{q}\bar{p} = qp + \bar{p}\bar{q}$，即实部$(pq)$ = 实部(qp)，故约化迹(pq) = 约化迹(qp)。

□

推论 A.6. 对所有 $p, q \in \mathbb{H}$，我们有 $pq - qp \in$ 虚部$(\mathbb{H})$。

A.2. **正四元数厄米矩阵的锥**。 我们沿用第 2 节的记号。设 V 为右-$\mathbb{H}$-模，且 $\dim_{\mathbb{H}}(\mathrm{V}) = n$，并设$\langle\cdot,\cdot\rangle : \mathrm{V} \times \mathrm{V} \to \mathbb{H}$为 V 上的正定 $(\iota, 1)$-埃尔米特形式，且 $\iota(\mathrm{X}) = \mathrm{X}$。如定理 2.3 所述，存在基 $\mathscr{B}$，使得 $\mathrm{Mat}_{\mathscr{B}}(\langle\cdot,\cdot\rangle) = \mathrm{Id}_n$。我们用 Sp(V) 表示 GL(V) 中由如下给出的子群

$$\mathrm{Sp}(\mathrm{V}) := \{\mathrm{T} \in \mathrm{GL}(\mathrm{V}), \langle \mathrm{T}(u), \mathrm{T}(v)\rangle = \langle u, v\rangle, (u, v \in \mathrm{V})\}.$$

类似地，使用基 $\mathscr{B}$并将 T 替换为 $\mathrm{Mat}_{\mathscr{B}}(\mathrm{T})$，我们用 $\mathrm{Sp}(n, \mathbb{H})$ 表示 $\mathrm{GL}(n)$ 的子群，其定义为

$$\mathrm{Sp}(n, \mathbb{H}) = \left\{g \in \mathrm{GL}(n), \iota(g)^t g = \mathrm{Id}_n\right\} = \{g \in \mathrm{GL}(n), g^* g = \mathrm{Id}_n\}.$$

我们用 sp(V) 表示 Sp(V) 的李代数，因此

$$\mathfrak{sp}(\mathrm{V}) := \{\mathrm{X} \in \mathrm{End}(\mathrm{V}), \langle \mathrm{X}(u), v\rangle + \langle u, \mathrm{X}(v)\rangle = 0, (u, v \in \mathrm{V})\},$$

并令 $\mathrm{sp}(n,\mathbb{H}) = \{X \in \mathrm{Mat}(n), X^* + X = 0\}$。
为简化记号，我们将用 $\mathrm{Sp}(n)$ 和 $\mathrm{sp}(n)$ 分别代替 $\mathrm{Sp}(n,\mathbb{H})$ 和 $\mathrm{sp}(n,\mathbb{H})$。

备注A.7. (1) 四元数辛群 $\mathrm{Sp}(n)$ 是连通且紧致的（见 [14]）。(2) 对于两个矩阵 $A、B \in \mathrm{Mat}(n)$，一般而言，我们有 $(AB)^t \neq B^tA^t$。不过，我们有 $(AB)^* = B^*A^*$。此外，若 A、B 可逆，一般而言，我们有$\iota(A)^{-1} \neq \iota(A^{-1})$，但是

$$(A^*)^{-1} = \left(A^{-1}\right)^*, \qquad (AB)^{-1} = B^{-1}A^{-1}.$$

该结果的证明见 [21, 定理 4.1]。

我们用 $\mathfrak{p}$ 表示 $\mathrm{Mat}(n)$ 的如下子集

$$\mathfrak{p} = \{X \in \mathrm{Mat}(n), X = X^*\}.$$

集合 $\mathfrak{p}$ 称为四元数厄米矩阵集合。此外，我们得到

$$[\mathfrak{sp}(n), \mathfrak{sp}(n)] \subseteq \mathfrak{sp}(n), \qquad [\mathfrak{sp}(n), \mathfrak{p}] \subseteq \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{sp}(n),$$

其中 $[\cdot,\cdot]$ 是 $\mathrm{Mat}(n)$ 上的李括号，定义为 $[A、B] = AB - BA$。

现在回顾 [21] 关于四元数矩阵特征值的一些结果。本文只讨论右特征值，即称$\lambda \in \mathbb{H}$ 为 $A \in \mathrm{Mat}(n)$ 的一个特征值，是指存在非零向量 $v \in \mathbb{H}^n$，使得 $Av = v\lambda$。

四元数矩阵与实数/复数矩阵的主要区别在于，一般来说，一个矩阵有无穷多个特征值。事实上，设 $A \in \mathrm{Mat}(n)$，且 $\lambda \in \mathbb{H}$是 A 的一个右特征值。那么对于任意 $q \in \mathbb{H}^*$，四元数 $q^{-1}\lambda q$ 也是 A 的一个右特征值。事实上，设 $v \in \mathbb{H}^n \setminus \{0\}$ 满足 $Av = v\lambda$。对于任意 $q \in \mathbb{H}^*$，定义 $w := vq$。那么$w \neq 0$并且

$$Aw = A(vq) = (Av)q = v\lambda q = (vq)(q^{-1}\lambda q) = w(q^{-1}\lambda q),$$

这表明$q^{-1}\lambda q$ 也是 A 的一个右特征值。

注记 A.8。对于所有 $\lambda \in \mathbb{H}$，我们用 $[\lambda]$ 表示 $\mathbb{H}^n$ 中由下式给出的子集

$$[\lambda] := \left\{q\lambda q^{-1}, q \in \mathbb{H}^*\right\}.$$

如 [21], 中所述，对于所有满足 $\mathrm{Im}(q \in \mathbb{H}) \neq 0$ 的 q，集合

$$[q] \cap \{z \in \mathbb{C}, \mathrm{Im}(z) > 0\}$$

有唯一的元素。此外，每个 $A \in \mathrm{Mat}(n)$ 都恰有 n 个右特征值（相差共轭），而前一个方程为 A 的谱给出了一个规范选择，该谱由虚部为正的复数构成。

然而，对于厄米矩阵，情况稍微简单一些。

引理 A.9. 设$A \in \mathfrak{p} \subseteq \mathrm{Mat}(n)$。那么A 的每个右特征值都是实数。

证明。 我们使用上文定义的 $\mathbb{H}^n$ 上的正定厄米形式

$$\langle x, y\rangle = x^* y .$$

设$v \in \mathbb{H}^n$ 非零且 $\lambda \in \mathbb{H}$ 满足 $\mathrm{A}v = v\lambda$。利用 $\mathrm{A} = \mathrm{A}^*$，得到

$$\langle \mathrm{A}x, y\rangle = \langle x, \mathrm{A}y\rangle, \qquad (x, y \in \mathbb{H}^n) .$$

特别地,

$$\langle \mathrm{A}v, v\rangle = \langle v, \mathrm{A}v\rangle.$$

利用 $\mathrm{A}v = v\lambda$，得到

$$\langle \mathrm{A}v, v\rangle = \langle v\lambda, v\rangle = \overline{\lambda}\langle v, v\rangle ,$$

and

$$\langle v, \mathrm{A}v\rangle = \langle v, v\lambda\rangle = \langle v, v\rangle\lambda .$$

因此

$$\overline{\lambda}\langle v, v\rangle = \langle v, v\rangle\lambda .$$

最后，利用 $\langle v、 v\rangle \in \mathbb{R}^*$，以及 $\mathbb{R}$ 与 $\mathbb{H}$可交换，得到 $\lambda = \lambda$。因此$\lambda \in \mathbb{R}$。

□

定义 A.10. 我们称矩阵 $\mathrm{X} \in \mathrm{Mat}(n)$ 是正定的，如果对所有非零$v \in \mathbb{H}^n$，都有 $\langle \mathrm{X}v, v\rangle > 0$ 。

我们用 $\mathscr{P}$ 表示正定厄米矩阵锥，并令

$$\exp : \mathrm{Mat}(n) \mapsto \mathrm{GL}(n)$$

为指数映射。

引理 A.11. 对于所有 $\mathrm{X} \in \mathfrak{p}$，有 $\exp(\mathrm{X}) \in \mathscr{P}$。

证明。 设 $\mathrm{X} \in \mathfrak{p}$，即 $\mathrm{X} = \mathrm{X}^*$。利用对于所有 $k \geq 0$，都有 $(\mathrm{X}^*)^k = \left(\mathrm{X}^k\right)^*$，可得 $\exp(\mathrm{X})^* = \exp(\mathrm{X}^*) = \exp(\mathrm{X})$，即 $\exp(\mathrm{X}) \in \mathfrak{p}$。此外，由注 A.8 和引理 A.9 可知，$\mathrm{Spec}(\mathrm{X}) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$，其中 $\lambda_i \in \mathbb{R}, 1 \leq i \leq n$。因此，$\mathrm{Spec}(\exp(\mathrm{X})) = \left\{e^{\lambda_1}, \ldots, e^{\lambda_n}\right\} \subseteq \mathbb{R}_{>0}$，即 $\exp(\mathrm{X})$ 是正定的。于是 $\exp(\mathrm{X}) \in \mathscr{P}$。

□

在下述定理中，我们回顾 [21, 推论 6.2] 中的一项重要结果；该结果对于证明 $\exp : \mathfrak{p} \to \mathscr{P}$ 是双射至关重要。

定理 A.12. 设 $\mathrm{X} \in \mathfrak{p}$。那么存在一个矩阵$\mathrm{U} \in \mathrm{Sp}(n)$ 以及实数$\lambda_1, \ldots, \lambda_n \in \mathbb{R}$，使得

$$\mathrm{X} = \mathrm{U}\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\mathrm{U}^* .$$

现在我们可以证明以下结果。

定理 A.13. 矩阵指数

$$\exp : \mathfrak{p} \mapsto \mathscr{P}, \qquad \mathrm{H} \mapsto \exp(\mathrm{H})$$

是定义良好的双射。

证明。 我们首先证明 $\exp : \mathfrak{p} \to \mathscr{P}$ 的满射性。令 $P \in \mathscr{P}$。根据四元数谱定理，存在 $U \in Sp(n)$ 以及实数 $\mu_1, \ldots, \mu_n > 0$，使得

$$P = U\mathrm{diag}(\mu_1, \ldots, \mu_n)U^* .$$

定义

$$H := U\mathrm{diag}(\log(\mu_1), \ldots, \log(\mu_n))U^* .$$

于是 $H^* = H$，即 $H \in \mathfrak{p}$，并且

$$\exp(H) = U\mathrm{diag}(e^{\log(\mu_1)}, \ldots, e^{\log(\mu_n)})U^* = U\mathrm{diag}(\mu_1, \ldots, \mu_n)U^* = P .$$

因此 exp 是满射。

现在我们证明 exp 在 $\mathfrak{p}$ 上的单射性。 假设 $H_1, H_2 \in \mathfrak{p}$ 满足 $\exp(H_1) = \exp(H_2)$。由谱定理，我们取 $U_1, U_2 \in Sp(n, \mathbb{H})$ 以及实数 $\lambda_1, \ldots, \lambda_n, \nu_1, \ldots, \nu_n$，使得

$$H_1 = U_1\mathrm{diag}(\lambda_1, \ldots, \lambda_n)U_1^*, \qquad H_2 = U_2\mathrm{diag}(\nu_1, \ldots, \nu_n)U_2^* .$$

Then

$$\exp(H_1) = U_1\mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})U_1^*, \qquad \exp(H_2) = U_2\mathrm{diag}(e^{\nu_1}, \ldots, e^{\nu_n})U_2^* .$$

由于 $\exp(H_1) = \exp(H_2)$ 是一个正定厄米矩阵，其谱分解除对角元素的排列以及各特征子空间内的酉基变换外是唯一的。特别地，我们有集合等式

$$\left\{e^{\lambda_1}, \ldots, e^{\lambda_n}\right\} = \{e^{\nu_1}, \ldots, e^{\nu_n}\} .$$

由于实指数函数 $t \mapsto e^t$ 在 $\mathbb{R}$ 上是单射，因此

$$\{\lambda_1, \ldots, \lambda_n\} = \{\nu_1, \ldots, \nu_n\} .$$

因此，H_1 和 H_2 可酉对角化，且具有相同的实特征值，所以 $H_1 = H_2$。

□

特别地，由定理 A.13 可知，映射 $\exp : \mathfrak{p} \to \mathscr{P}$ 可逆。令 $\log := \exp^{-1} : \mathscr{P} \to \mathfrak{p}$ 为 exp 的逆。

定义 A.14. 对于所有 $t \in \mathbb{R}$ 和 $X \in \mathscr{P}$，我们将 X^t 表示为 $\mathscr{P}$ 中由下式给出的矩阵

$$X^t := \exp\left(t\log(X)\right) .$$

如果$t = \frac{1}{2}$，则矩阵 $X^{\frac{1}{2}}$是 X 的平方根。

我们以一个引理结束本节。

引理 A.15。 对于所有 $g \in GL(n)$ 以及 $P \in \mathscr{P}$，我们有 $gPg^* \in \mathscr{P}$。此外，$GL(n)$ 在 $\mathscr{P}$ 上的相应作用是传递的。

证明。 利用注 A.7 可知，对所有 $g \in GL(n)$ 以及 $P \in \mathscr{P}$（即 $P = P^*$且 $P > 0$），我们得到

$$(gPg^*)^* = (g^*)^* P^* g^* = gPg^* ,$$

即 $gPg^* \in \mathfrak{p}$。此外，利用 $P > 0$，我们得到对所有非零 $x \in \mathbb{H}^n$

$$x^*(gPg^*)x = (x^*g)P(g^*x) = (g^*x)^* P(g^*x) > 0 ,$$

即 $g\mathrm{P}g^* \in \mathscr{P}$。

要证明相应的作用 $\mathrm{GL}(n) \curvearrowright \mathscr{P}$ 是传递的，只需证明每个 $\mathrm{P} \in \mathscr{P}$ 都属于单位矩阵 $\mathrm{Id}_n \in \mathscr{P}$的轨道 $\mathscr{O}_{\mathrm{Id}_n}$。设 $\mathrm{P} \in \mathscr{P}$。根据定理 A.12，存在矩阵 $\mathrm{U} \in \mathrm{Sp}(n)$ 和对角矩阵 $\mathrm{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathrm{Mat}(n)$，其中 $\lambda_i > 0$，使得 $\mathrm{P} = \mathrm{UDU}^*$。定义 $\mathrm{D}^{\frac{1}{2}} := \mathrm{diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\right)$，并令 $g := \mathrm{UD}^{\frac{1}{2}} \in \mathrm{GL}(n)$。于是

$$gg^* = \mathrm{UD}^{\frac{1}{2}}\left(\mathrm{UD}^{\frac{1}{2}}\right)^* = \mathrm{UD}^{\frac{1}{2}}\left(\mathrm{D}^{\frac{1}{2}}\right)^*\mathrm{U}^* = \mathrm{UD}^{\frac{1}{2}}\mathrm{D}^{\frac{1}{2}}\mathrm{U}^* = \mathrm{UDU}^* = \mathrm{P},$$

即 $\mathrm{P} \in \mathscr{O}_{\mathrm{Id}_n}$，引理得证。

□

注 A.16。 锥 $\mathscr{P}$ 是非紧型黎曼对称空间，即

$$\mathscr{P} \cong \mathrm{GL}(n)/\mathrm{Sp}(n).$$

关于这类空间的详细描述见 [11]。

A.3. 关于 $\mathscr{P}$ 的黎曼形式。 对于矩阵 $\mathrm{A} \in \mathrm{Mat}(n)$，记其标准迹为 $\mathrm{tr}_{\mathbb{H}}$，即

$$\mathrm{tr}_{\mathbb{H}}(\mathrm{A}) = \sum_{i=1}^{n} a_{i,i}.$$

注记 A.17。 由于 $\mathbb{H}$ 不是交换的，一般而言可以推出

$$\mathrm{tr}_{\mathbb{H}}(\mathrm{AB}) \neq \mathrm{tr}_{\mathbb{H}}(\mathrm{BA}),$$

对于 $\mathrm{A}、\mathrm{B} \in \mathrm{Mat}(n)$。

一种解决方法是使用约化迹，而不是标准迹。在方程（18）中，我们定义了 $\mathbb{H}$上的约化迹 trd。约化迹可以扩展到$\mathrm{Mat}(n)$，定义如下：

$$\mathrm{trd}(\mathrm{A}) := \mathrm{trd}(\mathrm{tr}_{\mathbb{H}}(\mathrm{A})) = \sum_{i=1}^{n} \mathrm{trd}(a_{i,i}), \qquad (\mathrm{A} \in \mathrm{Mat}(n)).$$

引理 A.18。 对于所有 $\mathrm{A}、\mathrm{B} \in \mathrm{Mat}(n)$，我们有

$$\mathrm{trd}(\mathrm{AB}) = \mathrm{trd}(\mathrm{BA}).$$

证明。 设 $\mathrm{A} = \left(a_{i,\ j}\right)_{i,\ j}$，$\mathrm{B} = \left(b_{i,\ j}\right)_{i,\ j}$。根据矩阵乘法的定义，AB 的对角元素为

$$(\mathrm{AB})_{i,i} = \sum_{j=1}^{n} a_{i,j}b_{j,i} \qquad (1 \leq i \leq n).$$

利用约化迹 trd 的可加性，我们得到

$$\mathrm{trd}(\mathrm{AB}) \quad = \quad \sum_{i=1}^{n} \mathrm{trd}((\mathrm{AB})_{i,i}) = \sum_{i=1}^{n} \mathrm{trd}\left(\sum_{j=1}^{n} a_{i,j}b_{j,i}\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{trd}(a_{i,j}b_{j,i})\,.$$

现在，利用引理 A.5，我们可以交换 trd 内部的因子：

$$\mathrm{trd}(a_{i,j}b_{j,i}) = \mathrm{trd}(b_{j,i}a_{i,j})\,,$$

so

$$\mathrm{trd}(\mathrm{AB}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{trd}(b_{j,i}a_{i,j})\,.$$

类似地，我们得到

$$(\mathrm{BA})_{j,j} = \sum_{i=1}^{n} b_{j,i}a_{i,j} \qquad (1 \le j \le n)\,,$$

因此

$$\mathrm{trd}(\mathrm{BA}) = \sum_{j=1}^{n} \mathrm{trd}((\mathrm{BA})_{j,j}) = \sum_{j=1}^{n}\sum_{i=1}^{n} \mathrm{trd}(b_{j,i}a_{i,j})\,.$$

最后，$\mathrm{trd}(\mathrm{AB}) = \mathrm{trd}(\mathrm{BA})$。

□

对于每个 $\mathrm{A} \in \mathscr{P}$，我们将切空间识别为

$$\mathrm{T}_{\mathrm{A}}(\mathscr{P}) \cong \mathfrak{p}\,.$$

在单位元 $\mathrm{Id}_n \in \mathscr{P}$ 处，我们在 $\mathrm{T}_{\mathrm{Id}_n}(\mathscr{P}) \cong \mathfrak{p}$ 上定义一个双线性形式 $\omega := \omega_{\mathrm{Id}_n}$，其定义为

$$\omega(\mathrm{X},\mathrm{Y}) := \mathrm{trd}(\mathrm{XY})\,, \qquad (\mathrm{X},\mathrm{Y} \in \mathfrak{p})\,. \tag{19}$$

该型取实值、对称（见引理 A.18），且正定。事实上，对于所有非零 $\mathrm{X} \in \mathfrak{p}$，由定理 A.12 可知

$$\omega(\mathrm{X},\mathrm{X}) = \mathrm{trd}(\mathrm{X}^2) = \sum_{j=1}^{n} \lambda_j^2 > 0\,,$$

其中，$\mathrm{Spec}(\mathrm{X}) = \{\lambda_1, \dots, \lambda_n\} \subseteq \mathbb{R}$ 是 X 的特征值。此外，对于所有$g \in \mathrm{Sp}(n)$以及 $\mathrm{X}, \mathrm{Y} \in \mathfrak{p}$，均有

$$\omega(g\mathrm{X}g^{-1}, g\mathrm{Y}g^{-1}) = \mathrm{trd}\left((g\mathrm{X}g^{-1})(g\mathrm{Y}g^{-1})\right) = \mathrm{trd}(g\mathrm{XY}g^{-1}) = \mathrm{trd}(\mathrm{XY}) = \omega(\mathrm{X},\mathrm{Y})\,,$$

也就是说，型 $\omega := \omega_{\mathrm{Id}_n}$ 是 $\mathrm{Sp}(n)$-不变的。

现在，我们利用 $\mathrm{GL}(n)$ 在 $\mathscr{P}$ 上的传递作用（见引理 A.15），在 $\mathscr{P}$ 上定义黎曼度量。对于所有$g \in \mathrm{GL}(n)$，我们将由 $\Phi_g : \mathscr{P} \to \mathscr{P}$ 给出的映射记为

$$\Phi_g(\mathrm{X}) = g\mathrm{X}g^*\,, \qquad (\mathrm{X} \in \mathscr{P})\,.$$

设 $g \in \mathrm{GL}(n)$ 且 $\mathrm{A} \in \mathscr{P}$。对于 $\mathrm{X} \in \mathrm{T}_{\mathrm{A}}(\mathscr{P}) \cong \mathfrak{p}$，我们考虑曲线

$$\gamma_{\mathrm{A}}^{\mathrm{X}}(t) = \mathrm{A} + t\mathrm{X}\,.$$

Then

$$\left(d\Phi_g\right)_{\mathrm{A}}(\mathrm{X}) = \left.\frac{d}{dt}\right|_{t=0} \Phi_g(\gamma_{\mathrm{A}}^{\mathrm{X}}(t)) = \left.\frac{d}{dt}\right|_{t=0} g\,(\mathrm{A} + t\mathrm{X})\,g^* = g\mathrm{X}g^*\,.$$

Then

$$\left(d\Phi_g\right)_{\mathrm{A}} : \mathrm{T}_{\mathrm{A}}(\mathscr{P}) \mapsto \mathrm{T}_{\Phi_g(\mathrm{A})}(\mathscr{P}), \qquad \mathrm{X} \mapsto g\mathrm{X}g^{*}. \tag{20}$$

由于 $\mathrm{A} = \Phi_{\mathrm{A}^{\frac{1}{2}}}(\mathrm{Id}_n)$，我们用 $\mathrm{A}^{-\frac{1}{2}}$将 A 处的切向量拉回到单位元处，得到

$$\omega(\mathrm{X},\mathrm{Y})_{\mathrm{A}} := \omega\left((d\Phi_{\mathrm{A}^{-\frac{1}{2}}})_{\mathrm{A}}(\mathrm{X}),(d\Phi_{\mathrm{A}^{-\frac{1}{2}}})_{\mathrm{A}}(\mathrm{Y})\right)_{\mathrm{Id}_n}.$$

利用单位元处度量的定义（见公式（19）），可得

$$\omega(\mathrm{X},\mathrm{Y})_{\mathrm{A}} = \mathrm{trd}\left((\mathrm{A}^{-\frac{1}{2}}\mathrm{X}\mathrm{A}^{-\frac{1}{2}})(\mathrm{A}^{-\frac{1}{2}}\mathrm{Y}\mathrm{A}^{-\frac{1}{2}})\right).$$

将迹下的因子重新组合，得到

$$\omega(\mathrm{X},\mathrm{Y})_{\mathrm{A}} = \mathrm{trd}(\mathrm{A}^{-1}\mathrm{X}\mathrm{A}^{-1}\mathrm{Y}). \tag{21}$$

引理 A.19. 正定对称双线性型族 $(\omega_{\mathrm{P}})_{\mathrm{P}\in\mathscr{P}}$ 在 $\mathscr{P}$上定义了一个黎曼度量。

A.4. **上的测地线** $\mathscr{P}$**.** 本节旨在证明以下定理。

定理 A.20. 设 A,B $\in \mathscr{P}$。那么，在 $\mathscr{P}$ 中连接A 和B 的唯一测地线（相对于黎曼度量 ω）由下式给出

$$\gamma(t) = \mathrm{A}^{\frac{1}{2}}\left(\mathrm{A}^{-\frac{1}{2}}\mathrm{B}\mathrm{A}^{-\frac{1}{2}}\right)^{t}\mathrm{A}^{\frac{1}{2}}. \tag{22}$$

在证明定理 A.20 之前，我们需要先证明几个中间结果。令 $\Psi\colon \mathbb{H} \to \mathrm{Mat}(2,\mathbb{C})$ 为公式 (17) 中给出的映射。利用 $\mathbb{H} = \mathbb{C}\oplus\mathbb{C}\mathbf{j}$，可得每个矩阵 $\mathrm{X} \in \mathrm{Mat}(n)$ 都可以写成 $\mathrm{X} = \mathrm{A} + \mathrm{B}\mathbf{j}$，其中 $\mathrm{A},\mathrm{B} \in \mathrm{Mat}(n,\mathbb{C})$。映射 Ψ 可扩展为映射 $\Psi\colon \mathrm{Mat}(n) \mapsto \mathrm{Mat}(2n,\mathbb{C})$，如下所示

$$\Psi(\mathrm{A}+\mathrm{B}\mathbf{j}) = \begin{bmatrix} \mathrm{A} & \mathrm{B} \\ -\overline{\mathrm{B}} & \overline{\mathrm{A}} \end{bmatrix}, \qquad (\mathrm{A},\mathrm{B} \in \mathrm{Mat}(n,\mathbb{C})). \tag{23}$$

对于所有 X、Y $\in \mathrm{Mat}(n)$，都有 $\Psi(\mathrm{XY}) = \Psi(\mathrm{X})\Psi(\mathrm{Y})$。此外，若 X 可逆，则 $\Psi(\mathrm{X})$ 也可逆，并且有 $\Psi(\mathrm{X}^{-1}) = \Psi(\mathrm{X})^{-1}$。
对于矩阵 $\mathrm{X} \in \mathrm{Mat}(2n,\mathbb{C})$，记 X^{*} 为矩阵 $\mathrm{X}^{*} = \mathrm{X}^{t}$。特别地，有 $\overline{\Psi}(\mathrm{X}^{*}) = \Psi(\mathrm{X})^{*}$。

记号 A.21。 我们将 $\mathrm{Mat}(2n,\mathbb{C})$ 中由下式给出的矩阵记为 J_{2n}

$$\mathrm{J}_{2n} := \begin{bmatrix} 0 & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0 \end{bmatrix}.$$

特别地，$\mathrm{J}_{2n}^{2} = -\mathrm{Id}_{2n}$。
引理 A.22。 我们得到 $\Psi(\mathrm{Mat}(n)) = \left\{\mathrm{M} \in \mathrm{Mat}(2n,\mathbb{C}),\ \mathrm{M}\mathrm{J}_{2n} = \mathrm{J}_{2n}\mathrm{M}\right\}$。

在以下引理中，我们将给出 $\Psi(\mathrm{Sp}(n))$和$\Psi(\mathscr{P})$的显式描述。

注记 A.23. 设 $\langle\cdot、\ \cdot\rangle$ 为第A.2节中给出的右$\mathbb{H}$模$\mathrm{V}=\mathbb{H}^n$ 上的埃尔米特形式。我们用$\mathrm{V}_{\mathbb{C}}$ 表示将向量空间V视为复向量空间。特别地，我们有$\dim_{\mathbb{C}}(\mathrm{V}_{\mathbb{C}})=2n$。我们用B表示$\mathrm{V}_{\mathbb{C}}$上由下式给出的型

$$\mathrm{B}(u,v)=\mathrm{pr}_{\mathbb{C}}(\langle u,v\rangle),\qquad (u,v\in\mathrm{V}_{\mathbb{C}}),$$

其中 $\mathrm{pr}_{\mathbb{C}}:\ \mathbb{H}\to\mathbb{C}$ 是由 $\mathrm{pr}_{\mathbb{C}}(a+b\mathbf{j})=a$ 给出的映射。
对所有 u、v、$w\in\mathrm{V}_{\mathbb{C}}$ 以及 $\lambda\in\mathbb{C}$，我们有

$$\begin{aligned}\mathrm{B}(u,v+w\lambda) &= \mathrm{pr}_{\mathbb{C}}(\langle u,v+w\lambda\rangle)=\mathrm{pr}_{\mathbb{C}}(\langle u,v\rangle+\langle u,w\rangle\lambda)\\ &= \mathrm{pr}_{\mathbb{C}}(\langle u,v\rangle)+\mathrm{pr}_{\mathbb{C}}(\langle u,w\rangle)\lambda=\mathrm{B}(u,v)+\mathrm{B}(u,w)\lambda,\end{aligned}$$

and

$$\begin{aligned}\mathrm{B}(u+v\lambda,w) &= \mathrm{pr}_{\mathbb{C}}(\langle u+v\lambda,w\rangle)=\mathrm{pr}_{\mathbb{C}}(\langle u,w\rangle+\overline{\lambda}\langle v,w\rangle)\\ &= \mathrm{pr}_{\mathbb{C}}(\langle u,w\rangle)+\overline{\lambda}\mathrm{pr}_{\mathbb{C}}(\langle v,w\rangle)=\mathrm{B}(u,w)+\overline{\lambda}\mathrm{B}(v,w),\end{aligned}$$

即 B 是 $\mathrm{V}_{\mathbb{C}}$ 上的半双线性形式。此外，利用对所有$q\in\mathbb{H}$ 有 $\mathrm{pr}_{\mathbb{C}}(\iota(q))=\overline{\mathrm{pr}_{\mathbb{C}}(q)}$，我们得到

$$\mathrm{B}(u,v)=\mathrm{pr}_{\mathbb{C}}(\langle u,v\rangle)=\mathrm{pr}_{\mathbb{C}}(\iota(\langle v,u\rangle))=\mathrm{pr}_{\mathbb{C}}\left(\overline{\langle v,u\rangle}\right)=\overline{\mathrm{B}(v,u)},$$

因此 B 是厄米的。最后，利用对所有非零的 u 有 $\langle u,u\rangle>0$，可知型 B 是正定的。我们将 $\mathrm{GL}(\mathrm{V}_{\mathbb{C}})$ 中如下给出的子群记为 $\mathrm{U}(\mathrm{V}_{\mathbb{C}})$

$$\mathrm{U}(\mathrm{V}_{\mathbb{C}}):=\{g\in\mathrm{GL}(\mathrm{V}_{\mathbb{C}}),\mathrm{B}(gu,gv)=\mathrm{B}(u,v),(u,,v\in\mathrm{V}_{\mathbb{C}})\}.$$

利用型 $\langle\cdot,\cdot\rangle$ 是 $\mathrm{Sp}(n)$-不变的，可知 $\Psi(\mathrm{Sp}(n))$ 是 $\mathrm{U}(\mathrm{V}_{\mathbb{C}})$ 的一个子群。

类似地，设 J 为 $\mathrm{V}_{\mathbb{C}}$ 上如下给出的映射

$$\mathrm{J}(v)=v\mathbf{j},\qquad (v\in\mathrm{V}_{\mathbb{C}}),$$

并令 $\mathrm{C}:\mathrm{V}_{\mathbb{C}}\times\mathrm{V}_{\mathbb{C}}\to\mathbb{C}$ 为如下给出的型

$$\mathrm{C}(u,v)=\mathrm{B}(\mathrm{J}(u),v),\qquad (u,v\in\mathrm{V}_{\mathbb{C}}).$$

利用 $z\mathbf{j}=\mathbf{j}z$ 对所有 $z\in\mathbb{C}$ 成立，可得 $\mathrm{J}(uz)=\mathrm{J}(u)z$。因此对所有$u$，$v\in\mathrm{V}_{\mathbb{C}}$ 以及 $z\in\mathbb{C}$，我们有

$$\mathrm{C}(uz,v)=z\mathrm{C}(u,v),\qquad \mathrm{C}(u,vz)=\mathrm{C}(u,v)z,$$

即 C 是双线性的。此外，我们有

$$\begin{aligned}\mathrm{C}(u,v) &= \mathrm{B}(\mathrm{J}(u),v)=\overline{\mathrm{B}(v,\mathrm{J}(u))}=\overline{\mathrm{pr}_{\mathbb{C}}(\langle v,u\mathbf{j}\rangle)}=\overline{\mathrm{pr}_{\mathbb{C}}(\langle v,u\rangle\mathbf{j})}\\ &= \overline{\mathrm{pr}_{\mathbb{C}}(\mathbf{j}\overline{\langle v,u\rangle})}=-\overline{\mathrm{pr}_{\mathbb{C}}(\overline{\langle v\mathbf{j},u\rangle})}=-\overline{\mathrm{pr}_{\mathbb{C}}(\overline{\langle \mathrm{J}(v),u\rangle})}\\ &= -\overline{\overline{\mathrm{C}(v,u)}}=-\mathrm{C}(v,u)\end{aligned}$$

即 C 是斜对称的。此外，利用 B 是非退化的，可得 t
C 是非退化的。我们将 $\mathrm{GL}(\mathrm{V}_{\mathbb{C}})$ 中如下给出的子群记为 $\mathrm{Sp}(\mathrm{V}_{\mathbb{C}})$

$$\mathrm{Sp}(\mathrm{V}_{\mathbb{C}}):=\{g\in\mathrm{GL}(\mathrm{V}_{\mathbb{C}}),\mathrm{C}(g(u),g(v))=\mathrm{C}(u,v),(u,v\in\mathrm{V}_{\mathbb{C}})\},$$

并且 $\Psi(\mathrm{Sp}(n))$ 是 $\mathrm{Sp}(\mathrm{V}_{\mathbb{C}})$ 的子群。

若 $\mathscr{B} = \{v_1, ..., v_n\}$ 是 $V = \mathbb{H}^n$ 的右 $\mathbb{H}$-基，则 $\mathscr{B}_{\mathbb{C}} = \{v_1, ..., v_n, v_1\mathbf{j}, ..., v_n\mathbf{j}\}$ 是 $V_{\mathbb{C}}$ 的复基。此外，可以看出

$$\mathrm{Mat}_{\mathscr{B}_{\mathbb{C}}}(B) = \mathrm{Id}_{2n}, \qquad \mathrm{Mat}_{\mathscr{B}_{\mathbb{C}}}(C) = J_{2n}.$$

因此，

$$U(V_{\mathbb{C}}) \cong U(2n, \mathbb{C}) = \{g \in GL(2n, \mathbb{C}), g^* g = \mathrm{Id}_{2n}\},$$

and

$$Sp(V_{\mathbb{C}}) \cong Sp(2n, \mathbb{C}) = \left\{g \in GL(2n, \mathbb{C}), g^t J_{2n} g = J_{2n}\right\}.$$

引理 A.24. 我们有 $\Psi(\mathrm{Sp}(n)) = U(2n, \mathbb{C}) \cap \mathrm{Sp}(2n, \mathbb{C})$。

证明。 如注 A.23 所述，我们有 $\Psi(\mathrm{Sp}(n)) \subseteq U(2n, \mathbb{C}) \cap \mathrm{Sp}(2n, \mathbb{C})$。为证明另一包含关系，利用三个群 $\mathrm{Sp}(n)$、$U(2n, \mathbb{C})$ 与 $\mathrm{Sp}(2n, \mathbb{C})$ 均为连通的，只需证明

$$\dim_{\mathbb{R}}(\mathrm{Lie}(\mathrm{Sp}(n))) = \dim_{\mathbb{R}}(\mathrm{Lie}\,(U(2n, \mathbb{C}) \cap \mathrm{Sp}(2n, \mathbb{C}))).$$

利用

$$\begin{aligned} \dim_{\mathbb{R}}(\mathrm{Lie}\,(U(2n, \mathbb{C}) \cap \mathrm{Sp}(2n, \mathbb{C}))) &= \dim_{\mathbb{C}}((\mathfrak{u}(2n, \mathbb{C}) \cap \mathfrak{sp}(2n, \mathbb{C})) \otimes_{\mathbb{R}} \mathbb{C}) \\ &= \dim_{\mathbb{C}}(\mathfrak{sp}(2n, \mathbb{C})), \end{aligned}$$

由此可得 $\dim_{\mathbb{C}}(\mathrm{sp}(2n, \mathbb{C})) = \dim_{\mathbb{C}}(S^2(\mathbb{C}^{2n})) = \frac{2n(2n+1)}{2} = n(2n+1)$。类似地，容易看出

$\dim_{\mathbb{R}}(\text{李群}(\mathrm{Sp}(n))) = 3n + \frac{4(n-1)n}{2} = 3n + 2n(n-1) = n(2n-2+3) = n(2n+1)$
,
由此得到该引理。

□

类似地，用 $\mathfrak{s}$ 表示 $\mathrm{Mat}(2n, \mathbb{C})$ 中厄米矩阵的集合，即

$$\mathfrak{s} := \{X \in \mathrm{Mat}(2n, \mathbb{C}), X^* = X\},$$

并用 $\mathscr{S}$ 表示 $\mathfrak{s}$ 中正矩阵的集合。

注 A.25. 复锥 $\mathscr{S}$ 恰好就是我们在第 2 节中定义的那个锥；我们更改了记号，以避免混淆。

引理 A.26. 我们得到

$$\Psi(\mathscr{P}) = \left\{X \in \mathscr{S}, XJ_{2n} = J_{2n}\overline{X}\right\}. \tag{24}$$

特别地， $\Psi(\mathscr{P})$ 是 $\mathscr{S}$ 的一个闭子流形。

证明。 将 $X \in \mathrm{Mat}(2n, \mathbb{C})$ 分块写为 $X = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$。直接计算可得

$$XJ = \begin{bmatrix} -Q & P \\ -S & R \end{bmatrix}, \qquad J\overline{X} = \begin{bmatrix} \overline{R} & \overline{S} \\ -\overline{P} & -\overline{Q} \end{bmatrix}.$$

因此，$XJ = \overline{J}\overline{X}$ 等价于 $R = -Q$ 和 $\overline{S} = P$，即

$$X = \begin{bmatrix} P & Q \\ -\overline{Q} & \overline{P} \end{bmatrix} = \Psi(P + Q\mathbf{j}) \in \Psi(\mathrm{Mat}(n)) .$$

与 $\mathscr{S}$ 相交（即在 X 为正定时施加 $X = X^*$）得到 $X = \Psi(A)$，其中 $A = A^*$ 且 A 正定，因此 $A \in \mathscr{P}$ 且 $X \in \Psi(\mathscr{P})$。
反之，若 $A \in \mathscr{P}$，则由上述分块形式可得 $\Psi(A) \in \mathscr{S}$ 且 $\overline{\Psi(A)}J_{2n} = J_{2n}\Psi(A)$。

□

我们将 τ（参见第5节）所定义在 $\mathscr{S}$ 上的黎曼度量记为

$$\tau_Q(X, Y) = \mathrm{tr}(Q^{-1}XQ^{-1}Y), \qquad (X, Y \in T_A(\mathscr{S}) \cong \mathfrak{s}) . \tag{25}$$

引理 A.27. 子流形 $\Psi(\mathscr{P})$ 在求逆、合同变换和实数幂运算下保持不变：

(1) 若 $X \in \Psi(\mathscr{P})$，则 $X^{-1} \in \Psi(\mathscr{P})$。（2）若 $X \in \Psi(\mathscr{P})$，则 $X^{\frac{1}{2}} \in \Psi(\mathscr{P})$，因此对所有 $t \in \mathbb{R}$，$X^t \in \Psi(\mathscr{P})$。（3）若 A、$B \in \Psi(\mathscr{P})$，则对所有 $t \in \mathbb{R}$，$A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}} \in \Psi(\mathscr{P})$。（4）若 A、$B \in \Psi(\mathscr{P})$，则连接 A 与 B 的测地线在 $\mathscr{S}$ 中（关于 τ）对所有 $t \in [0, 1]$ 均保持在 $\Psi(\mathscr{P})$ 内。

证明。(1) 如果 $XJ = J\overline{X}$，那么在左侧乘以 X^{-1} 得到 $J = X^{-1}JX$；在右侧乘以 $X^{-1} = \overline{X^{-1}}$ 则得到 $X^{-1}J = JX^{-1}$，因此 $X^{-1} \in \overline{\Psi(\mathscr{P})}$。

(2) 令 $X \in \Psi(\mathscr{P})$。则 $X \in \mathscr{S}$ 是正定厄米矩阵，并且 $XJ_{2n} = J_{2n}X$。令 $Y := X^{\frac{1}{2}}$ 为 X 在 $\mathscr{S}$ 中唯一的正定厄米平方根。对 $Y^2 = X$ 取复共轭可得 $\overline{Y}^2 = X$，因此

$$\left(J_{2n}\overline{Y}J_{2n}^{-1}\right)^2 = J_{2n}\overline{Y}^2J_{2n}^{-1} = J_{2n}\overline{X}J_{2n}^{-1} .$$

利用 $XJ_{2n} = J_{2n}X$ 和 $\overline{J_{2n}^{-1}} = -J_{2n}$，可得 $J_{2n}XJ_{2n}^{-1} = \overline{X}$，因此

$$\left(J_{2n}\overline{Y}J_{2n}^{-1}\right)^2 = X .$$

此外，$J_{2n}YJ_{2n}^{-1}$ 仍然是厄米正定矩阵（因为 Y 是）。因此 $J_{2n}YJ_{2n}^{-1}$ 是 X 的一个正定厄米平方根。由正定厄米平方根的唯一性，我们得到

$$J_{2n}\overline{Y}J_{2n}^{-1} = Y ,$$

这等价于 $YJ_{2n} = J_{2n}Y$。因此 $X^{\frac{1}{2}} = \overline{Y} \in \Psi(\mathscr{P})$。最后，对于任意 $t \in \mathbb{R}$，通过 $\mathscr{S}$ 中的函数演算定义 X^t。由于 $MJ_{2n} = J_{2n}M$ 这一关系在 X 的谱上进行全纯（因而连续）函数演算时保持不变，因此可得 $X^tJ_{2n} = J_{2n}X^t$，即对所有 $t \in \mathbb{R}$，$X^t \in \Psi(\mathscr{P})$。

(3) 若 $A, B \in \Psi(\mathscr{P})$，则由 (1) 和 (2) 可得 $A^{\pm\frac{1}{2}} \in \Psi(\mathscr{P})$；并且，满足 $XJ = JX$ 关系的两个元素的乘积仍满足这一关系：事实上，若 $XJ = JX$ 且 $YJ = JY$，则

$$(XY)J = X(YJ) = X(J\overline{Y}) = (XJ)\overline{Y} = J\overline{XY} = J\overline{XY}.$$

因此，$M := A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \in \Psi(\mathscr{P})$，由此根据 (2)，对所有 t，都有 $M^t \in \Psi(\mathscr{P})$。再次在两侧乘以 $A^{\frac{1}{2}}$，可得对所有 $t \in \mathbb{R}$，都有 $A^{\frac{1}{2}}M^tA^{\frac{1}{2}} \in \Psi(\mathscr{P})$。

(4) 根据 [6, 第 6], 章，A 和 B 之间的 $\widetilde{\gamma}$ 测地线 $\mathscr{S}$ 由 $\widetilde{\gamma}(t) = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}}$ 给出，它是相对于 τ 的、在 $\mathscr{S}$ 中连接 A 和 B 的测地线方程。利用 (3)，可得对所有 $t \in [0,1]$，都有 $\gamma(t) \in \Psi(\mathscr{P})$。

□

注A.28。利用方程 (21) 和 (25)，可知 $\mathscr{P}$ 上的度量 ω 是 $\Psi(\mathscr{P}) \subseteq \mathscr{S}$ 上的度量 τ 的拉回。

定理 *A.20* 的证明。 方程（23）中定义的映射 Ψ

$$\Psi : (\mathscr{P}, \omega) \longrightarrow \left(\Psi(\mathscr{P}), \tau_{|\Psi(\mathscr{P})}\right)$$

是一个等距映射。令 $A、B \in \mathscr{P}$，并设 $A = \Psi(A)$、$B = \Psi(B)$ 属于 $\Psi(\mathscr{P})$。如引理 A.27 所述，$\Psi(\mathscr{P})$ 中连接 A 和 B 的测地线方程为

$$\widetilde{\gamma}(t) = \widetilde{A}^{\frac{1}{2}}\left(\widetilde{A}^{-\frac{1}{2}}\widetilde{B}\widetilde{A}^{-\frac{1}{2}}\right)^t \widetilde{A}^{\frac{1}{2}}, \qquad (t \in [0,1]).$$

利用映射 Ψ 是等距映射这一点，由 [17, 命题 5.4] 可知，$\mathscr{P}$ 中连接 A 和 B 的测地线方程为

$$\gamma(t) = \Psi^{-1}(\widetilde{\gamma}(t)), \qquad (t \in [0,1]),$$

并利用对所有 $M \in \mathscr{P}$ 以及 $t \in [0,1]$，都有 $\Psi^{-1}(\Psi(M)^t) = M^t$，可得

$$\gamma(t) = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}}.$$

□

A.5. **上的几何平均**。 对于两个矩阵 $A、B \in \mathscr{P}$ 以及 $t \in [0,1]$，我们用 $A\sharp_t B$ 表示 $\mathrm{Mat}(n)$ 中由下式给出的矩阵

$$A\sharp_t B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}},$$

并且令 $A\sharp B = A\sharp_{\frac{1}{2}} B$。如 A.4 节所述，$A\sharp_t B \in \mathscr{P}$ 对所有 $t \in [0,1]$ 均成立。现在我们汇集几何平均 $\sharp_t$ 在 $\mathscr{P}$ 上的若干基本性质。由于证明与实数和复数情形中的证明相同，故从略。

命题 A.29。 • 对于所有 $A, B \in \mathscr{P}$ 以及 $t \in (0,1)$，$A\sharp_t A = A$ 且 $A\sharp_t B = A$ 当且仅当 $A = B$。

• 对于所有 $A, B \in \mathscr{P}$，$A\sharp B$ 都是里卡蒂方程的唯一解

$$XA^{-1}X = B,$$

- *For all* $\mathrm{A},\mathrm{B}\in\mathscr{P}$ *and* $\mathrm{C}\in\mathrm{GL}(n)$*, we have*
$$(\mathrm{C}^{*}\mathrm{AC})\sharp(\mathrm{C}^{*}\mathrm{BC})=\mathrm{C}^{*}(\mathrm{A}\sharp\mathrm{B})\mathrm{C}.$$
- *For all* $\mathrm{A}_1,\mathrm{A}_2,\mathrm{B}_1,\mathrm{B}_2\in\mathscr{P}$ *such that* $\mathrm{A}_1\preceq\mathrm{A}_2$ *and* $\mathrm{B}_1\preceq\mathrm{B}_2$*, we get*
$$\mathrm{A}_1\sharp\mathrm{B}_1\preceq\mathrm{A}_2\sharp\mathrm{B}_2.$$
- *For all* $\mathrm{A},\mathrm{B}\in\mathscr{P}$*, then* $(\mathrm{A}\sharp\mathrm{B})^{-1}=\mathrm{A}^{-1}\sharp\mathrm{B}^{-1}$.

我们将第 4 节中所述的 $\mathscr{P}$ 上的 Loewner 序记为 $\preceq$。

命题 A.30。对于所有 A、B $\in\mathscr{P}$，我们得到

(1) 甲$\sharp$乙 $=$ 乙$\sharp$甲*, (2)* (甲$\sharp$乙$)^{-1}=$ 甲$^{-1}\sharp$乙$^{-1}$.

此外，令 $t\in(0,1)$。

(1) 对于所有 $\mathrm{A}\in\mathscr{P}$，有 $\mathrm{A}\sharp_t\mathrm{A}=\mathrm{A}$。此外，有$\mathrm{A}\sharp_t\mathrm{B}=\mathrm{A}$当且仅当 $\mathrm{A}=\mathrm{B}$。*(2)* 对于所有 A、B $\in\mathscr{P}$和 α，$\beta>0$，有 *(3)* 对于所有 A、B $\in\mathscr{P}$，有
$$(\alpha\mathrm{A})\sharp_t(\beta\mathrm{B})=\alpha^{1-t}\beta^{t}\mathrm{A}\sharp_t\mathrm{B}.$$

$$\mathrm{A}\sharp_t\mathrm{B}=\mathrm{B}\sharp_{1-t}\mathrm{A}.$$

(4) 若A、B、C、D $\in\mathscr{P}$满足A $\preceq$C 且B $\preceq$D，则
$$\mathrm{A}\sharp_t\mathrm{B}\preceq\mathrm{C}\sharp_t\mathrm{D}.$$

我们以如下命题结束本附录。

命题 A.31。*(1)* 对所有 $g\in\mathrm{Sp}(n)$、$t\in[0,1]$ 以及A,B $\in\mathscr{P}$，有
$$\left(g\mathrm{A}g^{\sharp}\right)\sharp_t\left(g\mathrm{B}g^{\sharp}\right)=g(\mathrm{A}\sharp_t\mathrm{B})g^{\sharp}.$$

(2) 对所有A, B ,C , D $\in\mathscr{P}$ 以及 $t,s\in[0,1]$，有
$$(1-s)\mathrm{A}\sharp_t\mathrm{C}+s\mathrm{B}\sharp_t\mathrm{D}\preceq((1-s)\mathrm{A}+s\mathrm{B})\sharp_t((1-s)\mathrm{C}+s\mathrm{D}).$$

(3) 对所有A, B $\in\mathscr{P}$ 以及 $t,s,u\in[0,1]$,
$$(\mathrm{A}\sharp_t\mathrm{B})\sharp_u(\mathrm{A}\sharp_s\mathrm{B})=\mathrm{A}\sharp_{(1-u)t+us}\mathrm{B}.$$

参考文献

[1] 安藤毅和Fumio Hiai。对数优超与互补Golden-Thompson型不等式。第197/198卷，第113–131页。1994年。国际线性代数学会第二届会议（ILAS）（里斯本，1992年）。[2] N. Bebiano、R. Lemos、J. da Providência和G. Soares。关于*J*-压缩的算子不等式。数学不等式与应用，15(4):883–897，2012年。[3] N. Bebiano、R. Lemos和G. Soares。*J*-自伴矩阵平均及其不定不等式。塞格德数学学报，90(3-4):513–525，2024年。[4] N. Bebiano、R. Lemos和G. Soares。关于Krein空间数值域的双曲性。线性与多重线性代数，72(14):2267–2287，2024年。[5] N. Bebiano、R. Lemos和G. Soares。具有双曲Krein空间数值域的矩阵。算子理论进展，10(1):第13号论文，19页，2025年。

[6] Rajendra Bhatia。正定矩阵。普林斯顿应用数学系列。普林斯顿大学出版社，普林斯顿，新泽西州，2007年。 [2015] 2007年原版的平装版 [MR2284176]。[7] Rajendra Bhatia和John Holbrook。黎曼几何与矩阵几何平均。线性代数及其应用，413(2-3):594–618，2006年。[8] Hayoung Choi、Sejong Kim和Yongdo Lim。幂零矩阵加权几何平均的二项展开公式。线性与多重线性代数，72(4):615–630，2024年。[9] M. Dehghani和S. M. S. Modarres Mosadegh。Krein空间上的算子算术-调和平均不等式。数学拓展杂志，8(1):59–68，2014年。[10] Masatoshi Fujii和Eizaburo Kamei。Ando-Hiai 不等式与 Furuta 不等式。线性代数及其应用，416(2-3):541–545，2006年。[11] Sigurdur Helgason。微分几何、李群与对称空间，数学研究生教程第34卷。美国数学学会，普罗维登斯，罗得岛州，2001年。1978年原版的勘误重印本。[12] Fumio Hiai和Hideki Kosaki。无界算子的联络及相关主题：冯·诺伊曼代数情形。国际数学杂志，32(5):第2150024号论文，88页，2021年。[13] Roger A. Horn和Charles R. Johnson。矩阵分析。剑桥大学出版社，剑桥，第二版，2013年。[14] Anthony W. Knapp。李群导论（进阶篇），数学进展第140卷。Birkhäuser Boston, Inc.，波士顿，马萨诸塞州，第二版，2002年。[15] Fumio Kubo和安藤毅。正线性算子的平均。数学年刊，246(3):205–224，1979/80年。[16] Jimmie Lawson和Yongdo Lim。正算子的加权平均与Karcher方程。美国国家科学院院刊，110(39):15626–15632，2013年。[17] John M. Lee。黎曼流形导论，数学研究生教程第176卷。施普林格，Cham，2018年。第二版，系根据 [MR1468735] 出版。[18] Ming Liao、Xuhua Liu和Tin-Yau Tam。非紧型对称空间上的几何平均。李理论杂志，24(3):725–736，2014年。[19] W. Pusz和S. L. Woronowicz。半双线性形式的函数演算与纯化映射。数学物理报告，8(2):159–170，1975年。[20] Winfried Scharlau。二次型与埃尔米特形式，数学基础科学丛书第270卷 *[*数学科学基本原理 *]*。施普林格出版社，柏林，1985年。[21] Fuzhen Zhang。四元数与四元数矩阵。线性代数及其应用，251 57，1997 21：–。

数学与统计系 , 北佛罗里达大学 , UNF大道1号, 杰克-逊维尔, 佛罗里达州 32224, USA
电子邮箱：jose.franco@unf.edu

数学与统计系 , 北佛罗里达大学 , UNF大道1号, 杰克-逊维尔, 佛罗里达州 32224, USA
电子邮件地址: allan.merino@unf.edu