



ON A GEOMETRIC PROPERTY OF POSITIVE DEFINITE MATRICES CONE

MASATOSHI ITO¹, YUKI SEO², TAKEAKI YAMAZAKI^{3*} AND
MASAHIRO YANAGIDA⁴

Communicated by M. S. Moslehian

ABSTRACT. We shall discuss the matrix geometric mean for the positive definite matrices. The set of all $n \times n$ matrices with a suitable inner product will be a Hilbert space, and the matrix geometric mean can be considered as a path between two positive matrices. In this paper, we shall obtain a matrix geometric mean inequality, and as an application of it, a property of Riemannian metric space is given. We also obtain some examples related to our result.

1. INTRODUCTION

Let \mathbb{M}_n be the set of all $n \times n$ matrices on \mathbb{C} . The set \mathbb{M}_n will be a Hilbert space with the inner product $\langle A, B \rangle = \text{tr} B^* A$ for $A, B \in \mathbb{M}_n$, and the associated norm $\|A\|_2 = (\text{tr} A^* A)^{\frac{1}{2}}$. The set of all Hermitian matrices \mathbb{H}_n constitutes a real vector space on \mathbb{M}_n . The subset \mathbb{P}_n of \mathbb{M}_n consisting of positive definite matrices is an open subset in \mathbb{H}_n . Hence it is a differentiable manifold. The tangent space to \mathbb{P}_n at any of its points A is the space $T_A \mathbb{P}_n = \{A\} \times \mathbb{H}_n$, identified for simplicity, with \mathbb{H}_n . The inner product on \mathbb{H}_n leads to a Riemannian metric on the manifold \mathbb{P}_n . At the point A this metric is given by the differential

$$ds = \|A^{-\frac{1}{2}} dA A^{-\frac{1}{2}}\|_2 = [\text{tr}(A^{-1} dA)^2]^{\frac{1}{2}}.$$

Date: Received: 2 April 2009; Accepted: 18 June 2009.

* Corresponding author.

2000 Mathematics Subject Classification. Primary 47A64. Secondary 47A63, 47L25.

Key words and phrases. Positive matrix, Riemannian metric, geometric mean.

This is a mnemonic for computing the length of a (piecewise) differentiable path in \mathbb{P}_n . If $\gamma : [a, b] \rightarrow \mathbb{P}_n$ is such a path, we define its length as

$$L(\gamma) = \int_a^b \|\gamma^{\frac{-1}{2}}(t)\gamma'(t)\gamma^{\frac{-1}{2}}(t)\|_2 dt.$$

For any two points A and B in \mathbb{P}_n , let

$$\delta_2(A, B) = \inf\{L(\gamma) : \gamma \text{ is a path from } A \text{ to } B\}.$$

Bhatia and Holbrook [3] show that the infimum is attained at a unique path joining A and B . This path is called geodesic from A to B , and it is denoted by $[A, B]$. This gives a metric on \mathbb{P}_n , and concrete form of $\delta_2(A, B)$ is given as follows:

Theorem 1.1 ([2]). *Let A and B be any two elements of \mathbb{P}_n . Then there exists a unique geodesic $[A, B]$ joining A and B . This geodesic has a parametrization*

$$\gamma(t) = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^t A^{\frac{1}{2}}, \quad 0 \leq t \leq 1,$$

which is natural in the sense that

$$\delta_2(A, \gamma(t)) = t\delta_2(A, B)$$

for each $t \in [0, 1]$. Furthermore, the metric is precisely estimated by

$$\delta_2(A, B) = \|\log A^{\frac{-1}{2}}BA^{\frac{-1}{2}}\|_2.$$

We remark that the geodesic from A to B is known as the generalized geometric mean $A\sharp_t B$ of A and B , that is,

$$\gamma(t) = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^t A^{\frac{1}{2}} = A\sharp_t B \quad \text{for } t \in [0, 1].$$

In particular, $A\sharp_{\frac{1}{2}}B$ (denoted by $A\sharp B$, simply) is called the matrix geometric mean, simply, and it has many good properties. We shall introduce some properties of the generalized geometric mean in the next section. The metric δ_2 has an important property which is called exponential metric increasing property as follows:

Theorem 1.2 ([2, 3]). *For each pair of points A, B in \mathbb{P}_n , we have*

$$\delta_2(A, B) \geq \|\log A - \log B\|_2.$$

In other words, for any two matrices H and K in \mathbb{H}_n ,

$$\delta_2(e^H, e^K) \geq \|H - K\|_2.$$

So the map

$$\exp : (\mathbb{H}_n, \|\cdot\|_2) \longrightarrow (\mathbb{P}_n, \delta_2) \tag{1.1}$$

increases distances, or is metric increasing.

It is known that \mathbb{P}_n is a Riemannian manifold of nonpositive curvature [4]. Another essential feature of this geometry is the semiparallelogram law for the metric δ_2 . To understand this, recall the parallelogram law in a Hilbert space \mathcal{H} . Let a and b be any two points in \mathcal{H} and let $m = \frac{a+b}{2}$ be their midpoint. Given any other points c and d consider the parallelogram, one of whose diagonals is

$[a, b]$ and the other $[c, d]$. The two diagonals intersect at m and the parallelogram law is the equality

$$\|a - b\|^2 + \|c - d\|^2 = 2(\|a - c\|^2 + \|b - c\|^2).$$

Upon rearrangement this can be written as

$$\|c - m\|^2 = \frac{\|a - c\|^2 + \|b - c\|^2}{2} - \frac{\|a - b\|^2}{4}.$$

In the semiparallelogram law this last equality is replaced by an inequality.

Theorem 1.3 (The Semiparallelogram Law [2, 3]). *Let A and B be any two points of \mathbb{P}_n and let $M = A\sharp B$ be the midpoint of the geodesic $[A, B]$. Then for any C in \mathbb{P}_n we have*

$$\delta_2^2(C, M) \leq \frac{\delta_2^2(A, C) + \delta_2^2(B, C)}{2} - \frac{\delta_2^2(A, B)}{4}.$$

In the Euclidean space, the distance between the midpoints of two sides of a triangle is equal to half the length of the third side. In a space whose metric satisfies the semiparallelogram law this is replaced by an inequality as follows:

Theorem 1.4 ([2]). *Let A, B and C be any three points in \mathbb{P}_n . Then*

$$\delta_2(A\sharp B, A\sharp C) \leq \frac{\delta_2(B, C)}{2}.$$

By Theorem 1.1, the matrix geometric mean is closely related to Riemannian metric on the manifold \mathbb{P}_n . So we can expect that many results of the matrix geometric mean can be applied for the study of Riemannian metric in the manifold \mathbb{P}_n . As a property of the matrix geometric mean, the following result is obtained by Ando–Li–Mathias in [1]:

Theorem 1.5 ([1]). *For positive definite matrices A, B, C, D , $A\sharp B = C\sharp D$ implies*

$$(A\sharp C)\sharp(B\sharp D) = A\sharp B.$$

Roughly speaking, Theorem 1.5 says that for four points $A, B, C, D \in \mathbb{P}_n$, if the diagonals $[A, B]$ and $[C, D]$ have the same midpoint $G = A\sharp B = C\sharp D$ in a quadrilateral $ADBC$, then the line joining their midpoints in opposite sides $[A, C]$ and $[D, B]$ pass through the point G .

In this paper, we shall discuss an extension of Theorem 1.5. In section 2, we shall introduce some properties of matrix means. In section 3, we will give an extension of Theorem 1.5. In section 4, we give examples related to the result in section 3.

2. MATRIX MEAN

In this section, we shall introduce some properties of matrix means. Let M be a binary operation $M : \mathbb{P}_n \times \mathbb{P}_n \longrightarrow \mathbb{P}_n$. Then M is called a matrix mean if and only if the following conditions hold:

- (1) $M(A, B) \leq M(C, D)$ if $0 \leq A \leq C$ and $0 \leq B \leq D$ (monotonicity),

- (2) $XM(A, B)X \leq M(XAX, XBX)$ holds for $X \geq 0$ (transformer inequality),
- (3) M is upper semicontinuous, and
- (4) $M(I, I) = I$.

We remark that if X is invertible, then the transformer inequality (2) can be replaced into the transformer equality. It is well known that the matrix arithmetic and geometric means are typical examples of matrix mean. In this paper, we write the matrix arithmetic and the geometric means by

$$A\nabla B = \frac{A+B}{2} \quad \text{and} \quad A\sharp B = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}},$$

respectively. More generally, the generalized matrix arithmetic mean $A\nabla_t B$ and the geometric mean $A\sharp_t B$ are defined as follows: For $t \in [0, 1]$,

$$A\nabla_t B = (1-t)A + tB \quad \text{and} \quad A\sharp_t B = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t A^{\frac{1}{2}}.$$

We remark that if $t \notin [0, 1]$, then $A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t A^{\frac{1}{2}}$ is not a matrix mean, but we write it $A\sharp_t B$ for the sake of convenience.

Here we introduce some properties of the generalized matrix arithmetic and geometric means. Firstly, the famous arithmetic–geometric means inequality holds, i.e.,

$$A\sharp_t B \leq A\nabla_t B \quad \text{for } t \in [0, 1]. \quad (2.1)$$

Next property is shown by operator concavity of x^r for $r \in [0, 1]$.

$$A^r \nabla_t B^r \leq (A\nabla_t B)^r \quad \text{for } r \in [0, 1] \text{ and } t \in [0, 1]. \quad (2.2)$$

We remark that in the case $r = -1$, the sign of inequality in (2.2) is reversed, i.e.,

$$A^{-1} \nabla_t B^{-1} \geq (A\nabla_t B)^{-1} \quad \text{for } t \in [0, 1]. \quad (2.3)$$

In general, a matrix mean does not satisfy the commutative law, but the matrix arithmetic and the geometric means satisfy

$$A\nabla_t B = B\nabla_{1-t} A \quad \text{and} \quad A\sharp_t B = B\sharp_{1-t} A \quad (2.4)$$

for $t \in [0, 1]$, ($A\sharp_t B = B\sharp_{1-t} A$ for $t \notin [0, 1]$ also holds.) The former equality is obvious, and the latter one is given by

$$(AB^2A)^\alpha = AB(BA^2B)^{\alpha-1}BA \quad \text{for } A, B \geq 0 \text{ and real number } \alpha.$$

in [7, Lemma 1]. We remark that an argument over general matrix means is in [5, 8].

3. RESULTS

In this section, we shall give an extension of Theorem 1.5 as follows:

Theorem 3.1. *Let A, B, C, D be positive invertible matrices with $0 < mI \leq A, B, C, D \leq MI$ for some positive numbers m and M . If $A\sharp_\alpha B = C\sharp_\alpha D = G$ holds for a positive number $\alpha \in (0, 1)$, then for each $\beta \in [0, 1]$,*

$$\left\{ \frac{(h^{2-\frac{1}{\alpha_0}} + 1)^2}{4h^{2-\frac{1}{\alpha_0}}} \right\}^{-\alpha_0} G \leq (A\sharp_\beta C)\sharp_\alpha (B\sharp_\beta D) \leq \left\{ \frac{(h^{2-\frac{1}{\alpha_0}} + 1)^2}{4h^{2-\frac{1}{\alpha_0}}} \right\}^{\alpha_0} G \quad (3.1)$$

holds, where $h = \frac{M}{m}$ and $\alpha_0 = \min\{\alpha, 1 - \alpha\}$.

By putting $\alpha = \frac{1}{2}$ in Theorem 3.1, we obtain a slight extension of Theorem 1.5 as follows:

Corollary 3.2. *Let A, B, C, D be positive invertible matrices. If $A\sharp B = C\sharp D = G$ holds, then for each $\beta \in [0, 1]$,*

$$(A\sharp_{\beta}C)\sharp(B\sharp_{\beta}D) = G \quad (3.2)$$

holds.

Roughly speaking, Corollary 3.2 says that when we consider a quadrilateral $ADBC$ with the same midpoints G in each diagonals, for each internal ratio of $1 - \beta$ to β for $\beta \in [0, 1]$, the line joining their internally dividing point in opposite sides $[A, C]$ and $[D, B]$ pass through the midpoint G . We would like to remark that for any $\alpha \in (0, 1) - \{\frac{1}{2}\}$, “ \sharp ” in the center of the left hand side in (3.2) can not be replaced into “ \sharp_{α} ”. Concrete example and discussion will appear in the next section. From the viewpoint of Riemannian metric, Theorem 3.1 can be written by the following form:

Theorem 3.1’. *Let A, B, C, D be positive invertible matrices with $0 < mI \leq A, B, C, D \leq MI$ for some positive numbers m and M . If $A\sharp_{\alpha}B = C\sharp_{\alpha}D = G$ holds for a positive number $\alpha \in (0, 1)$, then for each $\beta \in [0, 1]$,*

$$\delta_2((A\sharp_{\beta}C)\sharp_{\alpha}(B\sharp_{\beta}D), G) \leq \sqrt{n} \log \left\{ \frac{(h^{2-\frac{1}{\alpha_0}} + 1)^2}{4h^{2-\frac{1}{\alpha_0}}} \right\}^{\alpha_0}$$

holds, where $h = \frac{M}{m}$ and $\alpha_0 = \min\{\alpha, 1 - \alpha\}$.

To prove Theorem 3.1, we need the following lemma which is a kind of reverse inequality of (2.3).

Lemma 3.3 ([6, 9]). *Let A and B be positive invertible matrices satisfying $0 < m_1I \leq A \leq M_1I$ and $0 < m_2I \leq B \leq M_2I$ for some positive numbers $0 < m_1 < M_1$ and $0 < m_2 < M_2$. Then*

$$A\nabla_t B \leq \frac{(h+1)^2}{4h} (A^{-1}\nabla_t B^{-1})^{-1} \quad (3.3)$$

holds for $t \in [0, 1]$, where $h = \max\{\frac{M_2}{m_1}, \frac{M_1}{m_2}\}$.

It is a known result, but for the reader’s convenient, we give a proof.

Proof. Let $X = A^{\frac{-1}{2}}BA^{\frac{-1}{2}}$, and

$$X = \int \lambda dE_{\lambda}$$

be the spectral decomposition of X . Since $h = \max\{\frac{M_2}{m_1}, \frac{M_1}{m_2}\}$, then we have $\frac{1}{h} \leq \lambda \leq h$. We remark that for $t \in (0, 1)$ and $h > 0$, since $h + \frac{1}{h} \geq 2$,

$$(1-t)^2 + t(1-t)(h + \frac{1}{h}) + t^2 = (2-h-\frac{1}{h}) \left(t - \frac{1}{2}\right)^2 + \frac{(h+1)^2}{4h} \leq \frac{(h+1)^2}{4h}. \quad (3.4)$$

Moreover $\lambda + \frac{1}{\lambda} \leq h + \frac{1}{h}$ for $\frac{1}{h} \leq \lambda \leq h$. Then for $t \in [0, 1]$,

$$\begin{aligned}
(1-t) + tX &= \int \{(1-t) + t\lambda\} dE_\lambda \\
&= \int \frac{\{(1-t) + t\lambda\}}{\{(1-t) + t\lambda^{-1}\}^{-1}} \cdot \{(1-t) + t\lambda^{-1}\}^{-1} dE_\lambda \\
&= \int \{(1-t)^2 + t(1-t)(\lambda + \frac{1}{\lambda}) + t^2\} \cdot \{(1-t) + t\lambda^{-1}\}^{-1} dE_\lambda \\
&\leq \int \{(1-t)^2 + t(1-t)(h + \frac{1}{h}) + t^2\} \cdot \{(1-t) + t\lambda^{-1}\}^{-1} dE_\lambda \\
&\leq \int \frac{(h+1)^2}{4h} \cdot \{(1-t) + t\lambda^{-1}\}^{-1} dE_\lambda \quad \text{by (3.4)} \\
&= \frac{(1+h)^2}{4h} \cdot \{(1-t) + tX^{-1}\}^{-1}.
\end{aligned}$$

Hence we have

$$(1-t) + tX \leq \frac{(1+h)^2}{4h} \cdot \{(1-t) + tX^{-1}\}^{-1}.$$

Multiplying both sides of this inequality by $A^{\frac{1}{2}}$, we have

$$(1-t)A + tB \leq \frac{(1+h)^2}{4h} \{(1-t)A^{-1} + tB^{-1}\}^{-1},$$

that is, (3.3). □

Proof of Theorem 3.1. Firstly, we prove the case $\alpha \in (0, \frac{1}{2}]$ (i.e., $\alpha_0 = \alpha$).

$$\begin{aligned}
A \sharp_\alpha B = G &\iff A^{\frac{1}{2}} (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^\alpha A^{\frac{1}{2}} = G \\
&\iff (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^\alpha = A^{-\frac{1}{2}} G A^{-\frac{1}{2}}.
\end{aligned}$$

Hence $B = A^{\frac{1}{2}} (A^{-\frac{1}{2}} G A^{-\frac{1}{2}})^{\frac{1}{\alpha}} A^{\frac{1}{2}} = A \natural_{\frac{1}{\alpha}} G$. Similarly, by (2.4), we have

$$C \sharp_\alpha D = G \iff D \sharp_{1-\alpha} C = G \iff C = D \natural_{\frac{1}{1-\alpha}} G.$$

Then

$$\begin{aligned}
& G^{\frac{-1}{2}} (A\sharp_{\beta}C)\sharp_{\alpha}(B\sharp_{\beta}D)G^{\frac{-1}{2}} \\
&= G^{\frac{-1}{2}} \left\{ A\sharp_{\beta}(D\sharp_{\frac{1}{1-\alpha}}G) \right\} \sharp_{\alpha} \left\{ (A\sharp_{\frac{1}{\alpha}}G)\sharp_{\beta}D \right\} G^{\frac{-1}{2}} \\
&= \left\{ (G^{\frac{-1}{2}}AG^{\frac{-1}{2}})\sharp_{\beta}(G^{\frac{-1}{2}}DG^{\frac{-1}{2}}\sharp_{\frac{1}{1-\alpha}}I) \right\} \sharp_{\alpha} \left\{ (G^{\frac{-1}{2}}AG^{\frac{-1}{2}}\sharp_{\frac{1}{\alpha}}I)\sharp_{\beta}G^{\frac{-1}{2}}DG^{\frac{-1}{2}} \right\} \\
&= \left\{ (G^{\frac{-1}{2}}AG^{\frac{-1}{2}})\sharp_{\beta}(G^{\frac{-1}{2}}DG^{\frac{-1}{2}})^{1-\frac{1}{1-\alpha}} \right\} \sharp_{\alpha} \left\{ (G^{\frac{-1}{2}}AG^{\frac{-1}{2}})^{1-\frac{1}{\alpha}}\sharp_{\beta}G^{\frac{-1}{2}}DG^{\frac{-1}{2}} \right\} \\
&= (X\sharp_{\beta}Y^{\frac{-\alpha}{1-\alpha}})\sharp_{\alpha}(X^{\frac{\alpha-1}{\alpha}}\sharp_{\beta}Y) \quad \text{by putting } X = G^{\frac{-1}{2}}AG^{\frac{-1}{2}}, Y = G^{\frac{-1}{2}}DG^{\frac{-1}{2}} \\
&\leq (X\nabla_{\beta}Y^{\frac{-\alpha}{1-\alpha}})\sharp_{\alpha}(X^{\frac{\alpha-1}{\alpha}}\nabla_{\beta}Y) \quad \text{by (2.1)} \\
&\leq (X^{\frac{1-\alpha}{\alpha}}\nabla_{\beta}Y^{-1})^{\frac{\alpha}{1-\alpha}}\sharp_{\alpha}(X^{\frac{\alpha-1}{\alpha}}\nabla_{\beta}Y) \quad \text{by (2.2) and } \frac{\alpha}{1-\alpha} \in [0, 1] \\
&\leq (X^{\frac{1-\alpha}{\alpha}}\nabla_{\beta}Y^{-1})^{\frac{\alpha}{1-\alpha}}\sharp_{\alpha} \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}}(X^{\frac{1-\alpha}{\alpha}}\nabla_{\beta}Y^{-1})^{-1} \right\} \\
&= \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}} \right\}^{\alpha} (X^{\frac{\alpha-1}{\alpha}}\nabla_{\beta}Y)^{\frac{\alpha}{1-\alpha}(1-\alpha)+(-\alpha)} \\
&= \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}} \right\}^{\alpha} I,
\end{aligned}$$

where the last inequality holds since $\frac{1}{h}I \leq X, Y \leq hI$ ensures

$$X^{\frac{\alpha-1}{\alpha}}\nabla_{\beta}Y \leq \frac{(h_0+1)^2}{4h_0}(X^{\frac{1-\alpha}{\alpha}}\nabla_{\beta}Y^{-1})^{-1},$$

where $h_0 = \max\left\{\frac{h}{\frac{\alpha-1}{\alpha}}, \frac{h^{\frac{\alpha-1}{\alpha}}}{h}\right\} = h^{2-\frac{1}{\alpha}}$ by Lemma 3.3. Hence we have

$$(A\sharp_{\beta}C)\sharp_{\alpha}(B\sharp_{\beta}D) \leq \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}} \right\}^{\alpha} G. \quad (3.5)$$

Now, we shall prove the lower bound of (3.1). By $0 < mI \leq A, B, C, D \leq MI$, we have $0 < M^{-1}I \leq A^{-1}, B^{-1}, C^{-1}, D^{-1} \leq m^{-1}I$, $\frac{m^{-1}}{M^{-1}} = \frac{M}{m} = h$. Hence by (3.5), we obtain

$$(A^{-1}\sharp_{\beta}C^{-1})\sharp_{\alpha}(B^{-1}\sharp_{\beta}D^{-1}) \leq \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}} \right\}^{\alpha} G^{-1}.$$

Taking the inverse of both sides,

$$(A\sharp_{\beta}C)\sharp_{\alpha}(B\sharp_{\beta}D) \geq \left\{ \frac{(h^{2-\frac{1}{\alpha}}+1)^2}{4h^{2-\frac{1}{\alpha}}} \right\}^{-\alpha} G,$$

that is, we get the desired inequality.

Next, we prove the case $\alpha \in [\frac{1}{2}, 1)$ (i.e., $\alpha_0 = 1 - \alpha$). By $A\sharp_\alpha B = C\sharp_\alpha D = G$, we have $B\sharp_{1-\alpha} A = D\sharp_{1-\alpha} C = G$. Since $1 - \alpha \in (0, \frac{1}{2}]$,

$$\left\{ \frac{(h^{2-\frac{1}{1-\alpha}} + 1)^2}{4h^{2-\frac{1}{1-\alpha}}} \right\}^{-1+\alpha} G \leq (B\sharp_\beta D)\sharp_{1-\alpha}(A\sharp_\beta C) \leq \left\{ \frac{(h^{2-\frac{1}{1-\alpha}} + 1)^2}{4h^{2-\frac{1}{1-\alpha}}} \right\}^{1-\alpha} G.$$

So that by (2.4),

$$\left\{ \frac{(h^{2-\frac{1}{\alpha_0}} + 1)^2}{4h^{2-\frac{1}{\alpha_0}}} \right\}^{-\alpha_0} G \leq (A\sharp_\beta C)\sharp_\alpha(B\sharp_\beta D) \leq \left\{ \frac{(h^{2-\frac{1}{\alpha_0}} + 1)^2}{4h^{2-\frac{1}{\alpha_0}}} \right\}^{\alpha_0} G.$$

□

4. COUNTEREXAMPLES

In the previous section, we prove Corollary 3.2 which is an extension of a result by Ando–Li–Mathias. But we do not show that whether “ \sharp ” in the center of the left hand side in (3.2) can be replaced into “ \sharp_α ” for $\alpha \neq \frac{1}{2}$ or not. One might think that Corollary 3.2 can be extended to more general form. The aim of this section is to give a counterexample of the problem as follows:

Proposition 4.1. *There exist positive invertible matrices A, B, C and D satisfying*

$$A\sharp_\alpha B = C\sharp_\alpha D \quad \text{and} \quad (A\sharp C)\sharp_\alpha(B\sharp D) \neq A\sharp_\alpha B$$

for any $\alpha \in (0, 1) - \{\frac{1}{2}\}$.

To prove the above proposition is a little bit complicated, so we shall give a concrete counterexample in the case $\alpha = \frac{1}{3}$, firstly.

Example 4.2. There exist positive invertible matrices A, B, C and D satisfying

$$A\sharp_{\frac{1}{3}} B = C\sharp_{\frac{1}{3}} D \quad \text{and} \quad (A\sharp C)\sharp_{\frac{1}{3}}(B\sharp D) \neq A\sharp_{\frac{1}{3}} B.$$

Proof. Let $A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 160 & 128 \\ 128 & 104 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ and $D = \begin{pmatrix} 580 & 308 \\ 308 & 164 \end{pmatrix}$.

Then

$$A\sharp_{\frac{1}{3}} B = C\sharp_{\frac{1}{3}} D = \begin{pmatrix} 8 & 4 \\ 4 & 4 \end{pmatrix} = G.$$

But

$$(A\sharp C)\sharp_{\frac{1}{3}}(B\sharp D) = \begin{pmatrix} 7.04103 \cdots & 3.12431 \cdots \\ 3.12431 \cdots & 3.65874 \cdots \end{pmatrix} \neq G.$$

□

The above example is not a counterexample for the general case $\alpha \in (0, 1) - \{\frac{1}{2}\}$. Next, we shall prove Proposition 4.1, i.e., general case $\alpha \in (0, 1) - \{\frac{1}{2}\}$. To give a proof of it, we prepare the following discussions:

Proposition 4.3. *For $\alpha \in (0, 1)$, assume that positive invertible matrices A, B, C, D with $A\sharp_\alpha B = C\sharp_\alpha D = G$ satisfy $(A\sharp C)\sharp_\alpha(B\sharp D) = G$. Then there exist positive invertible matrices X and Y such that*

$$X^{\frac{1-\alpha}{\alpha}} \sharp Y^{\frac{1-\alpha}{\alpha}} = (X\sharp Y)^{\frac{1-\alpha}{\alpha}}.$$

Proof. Let $A\#_{\alpha}B = C\#_{\alpha}D = G$. Then $B = A\#_{\frac{1}{\alpha}}G$ and $C = D\#_{\frac{1}{1-\alpha}}G$. Then we have

$$\begin{aligned} & G^{\frac{-1}{2}} \{(A\#C)\#_{\alpha}(B\#D)\} G^{\frac{-1}{2}} \\ &= G^{\frac{-1}{2}} \left[\left\{ A\# \left(D\#_{\frac{1}{1-\alpha}} G \right) \right\} \#_{\alpha} \left\{ \left(A\#_{\frac{1}{\alpha}} G \right) \# D \right\} \right] G^{\frac{-1}{2}} \\ &= \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right) \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \#_{\frac{1}{1-\alpha}} I \right) \right\} \#_{\alpha} \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \#_{\frac{1}{\alpha}} I \right) \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right) \right\} \\ &= \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right) \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right)^{1-\frac{1}{1-\alpha}} \right\} \#_{\alpha} \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right)^{1-\frac{1}{\alpha}} \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right) \right\} \\ &= \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right) \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right)^{\frac{-\alpha}{1-\alpha}} \right\} \#_{\alpha} \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right)^{\frac{\alpha-1}{\alpha}} \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right) \right\}, \end{aligned}$$

so that $(A\#C)\#_{\alpha}(B\#D) = G$ is equivalent to

$$\left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right) \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right)^{\frac{-\alpha}{1-\alpha}} \right\} \#_{\alpha} \left\{ \left(G^{\frac{-1}{2}} A G^{\frac{-1}{2}} \right)^{\frac{\alpha-1}{\alpha}} \# \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right) \right\} = I.$$

Let $X = G^{\frac{1}{2}} A^{-1} G^{\frac{1}{2}}$ and $Y = \left(G^{\frac{-1}{2}} D G^{\frac{-1}{2}} \right)^{\frac{-\alpha}{1-\alpha}}$. Then it is equivalent to

$$(X^{-1}\#Y^{-1})\#_{\alpha}(X^{\frac{1-\alpha}{\alpha}}\#Y^{\frac{1-\alpha}{\alpha}}) = I. \quad (4.1)$$

By the way, for positive invertible matrices S and T ,

$$S\#_{\alpha}T = I \iff T = S^{\frac{\alpha-1}{\alpha}}.$$

Hence (4.1) is equivalent to

$$X^{\frac{1-\alpha}{\alpha}}\#Y^{\frac{1-\alpha}{\alpha}} = (X^{-1}\#Y^{-1})^{\frac{\alpha-1}{\alpha}} = (X\#Y)^{\frac{1-\alpha}{\alpha}}.$$

It completes the proof. \square

Hence by Proposition 4.3, we have only to consider a counterexample of

$$X^r\#Y^r = (X\#Y)^r \quad (4.2)$$

for a positive number r with $r \neq 1$.

Proposition 4.4. *For positive 2×2 matrices X and Y with $\det(X) = \det(Y) = 1$. Then*

$$X\#Y = \frac{X+Y}{\sqrt{\det(X+Y)}}.$$

It is a well-known formula (see [1]), but we shall give a proof for the reader's convenient.

Proof. Let $D = X^{\frac{-1}{2}} Y X^{\frac{-1}{2}}$. Then by the Cayley–Hamilton theorem, we have

$$D^2 - \text{trace}(D)D + I = 0.$$

Then by

$$D = \frac{(D+I)^2}{\text{trace}(D)+2} \geq 0,$$

we obtain

$$D^{\frac{1}{2}} = \frac{D+I}{\sqrt{\text{trace}(D)+2}}.$$

Hence

$$X\sharp Y = X^{\frac{1}{2}}D^{\frac{1}{2}}X^{\frac{1}{2}} = \frac{X+Y}{\sqrt{\text{trace}(D)+2}}.$$

Here by $\det(X\sharp Y) = 1$, we can get

$$X\sharp Y = \frac{X+Y}{\sqrt{\det(X+Y)}}.$$

□

Then we shall give a counterexample to prove Proposition 4.1 based on the above discussions.

Example 4.5. Let

$$X = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } Y = \frac{1}{2\sqrt{2}} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

Then (4.2) does not hold for all positive number $r \neq 1$.

Proof. It is easy to see $\det(X) = \det(Y) = 1$. Then there exist real numbers α and β_r such that

$$\begin{aligned} X\sharp Y &= \alpha(X+Y), \text{ hence } (X\sharp Y)^r = \alpha^r(X+Y)^r, \\ X^r\sharp Y^r &= \beta_r(X^r+Y^r) \end{aligned}$$

by Proposition 4.4. Put

$$\begin{aligned} (X\sharp Y)^r &= \alpha^r(X+Y)^r = \begin{pmatrix} X_1(r) & X_2(r) \\ X_2(r) & X_3(r) \end{pmatrix}, \\ X^r\sharp Y^r &= \beta_r(X^r+Y^r) = \begin{pmatrix} Y_1(r) & Y_2(r) \\ Y_2(r) & Y_3(r) \end{pmatrix}. \end{aligned}$$

If (4.2) holds for some r , then we have $X_1(r) = Y_1(r)$ and $X_2(r) = Y_2(r)$, so that,

$$\frac{X_1(r)}{X_2(r)} = \frac{Y_1(r)}{Y_2(r)}. \quad (4.3)$$

Hence we have only to check that (4.3) does not hold for all positive number r with $r \neq 1$.

Put a unitary matrix U as $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Then we have

$$Y = U \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot U.$$

Hence we obtain

$$X^r + Y^r = \frac{1}{2(\sqrt{2})^r} \begin{pmatrix} 3 \cdot 2^r + 1 & 2^r - 1 \\ 2^r - 1 & 2^r + 3 \end{pmatrix}. \quad (4.4)$$

By putting $r = 1$ in (4.4), we get

$$X + Y = \frac{1}{2\sqrt{2}} \begin{pmatrix} 7 & 1 \\ 1 & 5 \end{pmatrix}.$$

Let V be a unitary matrix as follows:

$$V = \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} 1 & \sqrt{2}-1 \\ \sqrt{2}-1 & -1 \end{pmatrix}.$$

Then we have

$$V(X+Y)V = \frac{1}{2\sqrt{2}} \begin{pmatrix} 6+\sqrt{2} & 0 \\ 0 & 6-\sqrt{2} \end{pmatrix},$$

and then

$$(X+Y)^r \tag{4.5}$$

$$= k(r) \begin{pmatrix} (6+\sqrt{2})^r + (3-2\sqrt{2})(6-\sqrt{2})^r & (\sqrt{2}-1)\{(6+\sqrt{2})^r - (6-\sqrt{2})^r\} \\ (\sqrt{2}-1)\{(6+\sqrt{2})^r - (6-\sqrt{2})^r\} & (3-2\sqrt{2})(6+\sqrt{2})^r + (6-\sqrt{2})^r \end{pmatrix},$$

where $k(r) = \frac{1}{(2\sqrt{2})^r(4-2\sqrt{2})}$.

If $X^r \sharp Y^r = (X \sharp Y)^r$ for some $r > 0$ with $r \neq 1$, then by (4.3), (4.4) and (4.5)

$$\frac{3 \cdot 2^r + 1}{2^r - 1} = \frac{(6+\sqrt{2})^r + (3-2\sqrt{2})(6-\sqrt{2})^r}{(\sqrt{2}-1)\{(6+\sqrt{2})^r - (6-\sqrt{2})^r\}}, \tag{4.6}$$

if and only if

$$(3 \cdot 2^r + 1)(\sqrt{2}-1)(\lambda^r - 1) - (2^r - 1)\{\lambda^r + (3-2\sqrt{2})\} = 0, \text{ where } \lambda = \frac{6+\sqrt{2}}{6-\sqrt{2}},$$

if and only if

$$(\sqrt{2}-1)^2(2\lambda)^r - 2^r + \lambda^r - (\sqrt{2}-1)^2 = 0.$$

Let $f(r) = (\sqrt{2}-1)^2(2\lambda)^r - 2^r + \lambda^r - (\sqrt{2}-1)^2$. Then we have only to show $f(r) \neq 0$ for any $r > 0$ such that $r \neq 1$ by the above argument. By calculation,

$$\begin{aligned} f'(r) &= (\sqrt{2}-1)^2(2\lambda)^r \log(2\lambda) - 2^r \log 2 + \lambda^r \log \lambda \\ &= (2\lambda)^r \{(\sqrt{2}-1)^2 \log(2\lambda) - \lambda^{-r} \log 2 + 2^{-r} \log \lambda\}. \end{aligned}$$

Put $g(r) = (\sqrt{2}-1)^2 \log(2\lambda) - \lambda^{-r} \log 2 + 2^{-r} \log \lambda$. Then

$$g'(r) = -\lambda^{-r} \log \lambda^{-1} \cdot \log 2 + 2^{-r} \log 2^{-1} \cdot \log \lambda = \log 2 \cdot \log \lambda \cdot (\lambda^{-r} - 2^{-r}) > 0$$

for $r > 0$ since $1 < \lambda < 2$. Therefore $f'(r) = (2\lambda)^r g(r)$ is strictly increasing for $r \geq 0$, so that $f(r)$ is a convex function for $r \geq 0$. Hence $f(r) \neq 0$ for any $r > 0$ such that $r \neq 1$ since $f(0) = f(1) = 0$. \square

Remark 4.6. The above calculation is a bit complicated, but only to prove the existence of a counterexample, there is an easier calculation as follows:

For some $r > 0$, we assume that

$$X^r \sharp Y^r = (X \sharp Y)^r \tag{4.2}$$

holds for any positive invertible matrices X and Y . Then by putting $X_1 = X^r$ and $Y_1 = Y^r$, we have

$$(X_1 \sharp Y_1)^{\frac{1}{r}} = X_1^{\frac{1}{r}} \sharp Y_1^{\frac{1}{r}}.$$

Hence (4.2) also holds for $\frac{1}{r}$. So we have only to consider the case $r > 1$. If (4.2) holds for all $X, Y > 0$ and some $r > 1$, then we obtain

$$X^{r^2} \# Y^{r^2} = (X^r \# Y^r)^r = (X \# Y)^{r^2},$$

that is, (4.2) holds for r^2 . By the same way, (4.2) holds for r^n . Hence we have to check it for all sufficiently large r . Here, from the left and right hand sides of (4.6), we have

$$\lim_{r \rightarrow \infty} \frac{3 \cdot 2^r + 1}{2^r - 1} = \lim_{r \rightarrow \infty} \left(\frac{4}{2^r - 1} + 3 \right) = 3,$$

$$\lim_{r \rightarrow \infty} \frac{(6 + \sqrt{2})^r + (3 - 2\sqrt{2})(6 - \sqrt{2})^r}{(\sqrt{2} - 1)\{(6 + \sqrt{2})^r - (6 - \sqrt{2})^r\}} = \lim_{r \rightarrow \infty} \left\{ \frac{2\sqrt{2}}{\left(\frac{6+\sqrt{2}}{6-\sqrt{2}}\right)^r - 1} + \sqrt{2} + 1 \right\} = \sqrt{2} + 1.$$

Hence we obtain that (4.6) does not hold for all sufficiently large r since the above limit points are different. It completes the proof.

Consequently we can prove Proposition 4.1 as follows:

Proof of Proposition 4.1. By Example 4.5, there exist positive invertible matrices X and Y satisfying

$$X^{\frac{1-\alpha}{\alpha}} \# Y^{\frac{1-\alpha}{\alpha}} \neq (X \# Y)^{\frac{1-\alpha}{\alpha}}$$

for all $\alpha \in (0, 1) - \{\frac{1}{2}\}$. Therefore, by scrutinizing the proof of Proposition 4.3, we can get desired matrices A, B, C and D for a given positive invertible matrix G as follows:

$$A = G^{\frac{1}{2}} X^{-1} G^{\frac{1}{2}}, \quad D = G^{\frac{1}{2}} Y^{\frac{1-\alpha}{\alpha}} G^{\frac{1}{2}}, \quad B = A \natural_{\frac{1}{\alpha}} G \quad \text{and} \quad C = D \natural_{\frac{1}{1-\alpha}} G.$$

□

REFERENCES

1. T. Ando, C.K. Li and R. Mathias, *Geometric means*, Linear Algebra Appl., **385** (2004), 305–334.
2. R. Bhatia, *Positive definite matrices*, Princeton University Press, New Jersey, (2007).
3. R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl., **413** (2006), 594–618.
4. G. Corach, H. Porta and L. Recht, *Geodesics and operator means in the space of positive operators*, Int. J. Math., **4** (1993), 193–202.
5. J.I. Fujii and E. Kamei, *Uhlmann’s interpolational method for operator means*, Math. Japon., **34** (1989), 541–547.
6. J.I. Fujii, M. Nakamura, J. Pečarić and Y. Seo, *Bounds for the ratio and difference between parallel sum and series via Mond-Pečarić method*, Math. Inequal. Appl., **9** (2006), 749–759.
7. T. Furuta, *Extension of the Furuta inequality and Ando–Hiai log-majorization*, Linear Algebra Appl., **219** (1995), 139–155.
8. F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann., **246** (1980), 205–224.
9. J. Pečarić and J. Mičić, *Chaotic order among means of positive operators*, Sci. Math. Jpn., **7** (2002), 97–106.

¹ DEPARTMENT OF INTEGRATED DESIGN ENGINEERING, MAEBASHI INSTITUTE OF TECHNOLOGY, MAEBASHI 371-0816, JAPAN

E-mail address: m-ito@maebashi-it.ac.jp

² FACULTY OF ENGINEERING, SHIBAURA INSTITUTE OF TECHNOLOGY, SAITAMA 337-8570, JAPAN

E-mail address: yukis@sic.shibaura-it.ac.jp

³ DEPARTMENT OF MATHEMATICS, KANAGAWA UNIVERSITY, YOKOHAMA 221-8686, JAPAN

E-mail address: yamazt26@kanagawa-u.ac.jp

⁴ DEPARTMENT OF MATHEMATICAL INFORMATION SCIENCE, FACULTY OF SCIENCE, TOKYO UNIVERSITY OF SCIENCE, TOKYO 162-8601, JAPAN

E-mail address: yanagida@rs.kagu.tus.ac.jp