

OPERATOR INEQUALITIES AND GYROLINES OF THE WEIGHTED GEOMETRIC MEANS

SEJONG KIM

(Communicated by J.-C. Bourin)

Abstract. We consider in this paper two different types of the weighted geometric means of positive definite operators. We show the component-wise bijection of these geometric means and give a geometric property of the spectral geometric mean as a metric midpoint. Moreover, several interesting inequalities related with the geometric means of positive definite operators will be shown. We also see the meaning of weighted geometric means in the gyrogroup structure with finite dimension and find the formulas of weighted geometric means of 2-by-2 positive definite matrices and density matrices.

1. Introduction

A geometric mean of two positive real numbers is the length of the side of the square with the same area of the rectangle with sides of given positive real numbers. It was introduced first in Euclid's *Elements* (Book II, Proposition 14), and since then many characterizations have been studied (see Section 1 in [18]). During several decades, a variety of construction schemes of the geometric means for positive operators and matrices together with their properties and applications have been widely developed.

Since Pusz and Woronowicz [26] has defined a geometric mean $A\#B$ of positive definite matrices A and B ,

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$$

many properties on the finite- and infinite-dimensional settings have been found. The geometric mean of positive definite matrices A and B is a unique midpoint of the Riemannian geodesic, called the weighted geometric mean of A and B :

$$\gamma(t) = A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2} =: A\#_t B, \quad t \in [0, 1]$$

connecting two points A and B with respect to the Riemannian trace metric $\delta(A, B) = \|\log A^{-1/2}BA^{-1/2}\|_2$. Since Kubo and Ando [17] have established the geometric mean of positive definite operators, many theoretical and computational research topics including the operator inequality and the extension theory to multi-variable geometric means have been widely studied.

Mathematics subject classification (2010): 47B65, 15B48.

Keywords and phrases: Positive definite operator, Loewner order, geometric mean, spectral geometric mean, gyrogroup, gyroline.

On the other hand, Fiedler and Pták [8] have introduced and studied a new geometric mean $A\sharp B$ of two positive definite matrices A and B :

$$A\sharp B = (A^{-1}\#B)^{1/2}A(A^{-1}\#B)^{1/2}$$

having analogous properties of the geometric mean. Since the eigenvalues of $A\sharp B$ are the same as the positive square roots of the eigenvalues of AB , we call it the spectral geometric mean. It can be naturally generalized to the weighted spectral geometric mean as a differentiable curve

$$\beta(t) = (A^{-1}\#B)^t A (A^{-1}\#B)^t =: A\sharp_t B, \quad t \in [0, 1].$$

Note that $A\sharp B = A\sharp_{1/2}B$ and $A\sharp_t B = A^{1-t}B^t$ for any commuting A and B . Several interesting properties of the weighted spectral geometric mean have been found [1, 16, 21], but not as much as the geometric mean.

Those weighted geometric means can be extended to all real numbers $t \in \mathbb{R}$. The main goal of this paper is to investigate those weighted geometric means of positive definite operators with operator inequality and applications to the finite-dimensional gyrogroup structure, which provides an algebraic tool for the hyperbolic geometry and the theory of special relativity. In more details, we give a geometric property of the spectral geometric mean for positive definite operators as a metric midpoint, see the meaning of two weighted geometric means on the gyrogroup structure of positive definite Hermitian matrices and density matrices, respectively, and find the explicit formulas of two weighted geometric means for 2×2 positive definite matrices and density matrices.

The structure of this article is organized as follows. We show in Section 2 the component-wise bijection of these weighted geometric means for positive definite operators and give a geometric property of the spectral geometric mean as a midpoint with respect to the new semi-metric $d(A, B) = 2\|\log(A^{-1}\#B)\|$. In Section 3 we prove several operator inequalities of two weighted geometric means including the Ando-Hiai inequality with sufficient condition of the spectral geometric mean. In Section 4 we see a connection between the weighted (spectral) geometric means and the gyroline (cogyroline, respectively) on the open convex cone of positive definite matrices and on the gyrovector space of invertible density matrices. We also provide in Section 5 the explicit formulas of the weighted geometric means of 2×2 positive definite matrices and invertible density matrices as a linear combination. Finally, we close with some remarks about the semi-metric and an inequality relation with the Riemannian trace metric.

2. Two-variable geometric means of positive definite operators

Let $B(\mathcal{H})$ be the Banach space of all bounded linear operators on a Hilbert space \mathcal{H} equipped with the inner product $\langle \cdot, \cdot \rangle$ and the operator norm $\|\cdot\|$. Let $S(\mathcal{H}) \subset B(\mathcal{H})$ be the closed subspace of all self-adjoint operators, and let $\mathbb{P} \subseteq S(\mathcal{H})$ be the open convex cone of all positive definite operators. Note that $A \in \mathbb{P}$ means that $\langle x, Ax \rangle > 0$ for all nonzero $x \in \mathcal{H}$. For $X, Y \in S(\mathcal{H})$ we denote as $X \leq Y$ if $Y - X$ is

positive semi-definite, and as $X < Y$ if $Y - X$ is positive definite: this is known as the Loewner order on $S(\mathcal{H})$. For the case that the dimension of the Hilbert space \mathcal{H} is finite, say $\dim(\mathcal{H}) = n < \infty$, we denote as \mathbb{P}_n instead of \mathbb{P} .

2.1. Geometric means

Pusz and Woronowicz [26] have first introduced a geometric mean $A\#B$ of positive definite Hermitian matrices A and B :

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.$$

Since then Kubo and Ando [17] have established the geometric mean of positive definite operators in \mathbb{P} with several remarkable properties. It coincides with the unique positive definite solution $X \in \mathbb{P}$ of the Riccati equation $XA^{-1}X = B$, and moreover, from [17]

$$A\#B = \max \left\{ X \in S(\mathcal{H}) : \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geq 0 \right\}. \tag{1}$$

Note that (1) is proved by the Schur complement, the Loewner-Heinz inequality in Lemma 3 and the order preserving of congruence transformation.

The Thompson metric on \mathbb{P} is defined by $d_T(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|$, where $\|X\|$ denotes the operator norm of X . It is known that d_T is a complete metric on \mathbb{P} and that

$$d_T(A, B) = \max\{\log M(B/A), \log M(A/B)\},$$

where $M(B/A) = \inf\{\alpha > 0 : B \leq \alpha A\} = \lambda_{\max}(A^{-1/2}BA^{-1/2})$, the largest eigenvalue of $A^{-1/2}BA^{-1/2}$: see [27]. The geometric mean curve

$$[0, 1] \ni t \mapsto A\#_t B := A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}. \tag{2}$$

is a minimal geodesic from A to B for the Thompson metric. It can be extended to the curve $t \mapsto A\#_t B$ for all real numbers t . We call $A\#_t B$ for $t \in [0, 1]$ the *weighted geometric mean* of A and B on \mathbb{P} , and especially, its metric midpoint $A\#B := A\#_{1/2} B$.

THEOREM 1. *Let $A, B \in \mathbb{P}$.*

(1) *The map $X \mapsto A\#_t X$ is a bijection on \mathbb{P} for any $t \in (0, 1]$.*

(2) *The map $X \mapsto X\#_t B$ is a bijection on \mathbb{P} for any $t \in [0, 1)$.*

Proof. We first show (1). Then (2) holds easily from (1), since $X\#_t B = B\#_{1-t} X$. Let $t \in (0, 1]$.

(i) For $X, Y \in \mathbb{P}$, assume that $A\#_t X = A\#_t Y$. Then $(A^{-1/2}XA^{-1/2})^t = (A^{-1/2}YA^{-1/2})^t$ by congruence transformation. So $A^{-1/2}XA^{-1/2} = A^{-1/2}YA^{-1/2}$ by taking the $(1/t)$ -power map. Thus, $X = Y$, that is, the map $X \mapsto A\#_t X$ is injective.

(ii) For any $C \in \mathbb{P}$, set $X = A\#_{1/t} C$. Then $X \in \mathbb{P}$ and $A\#_t X = A\#_t(A\#_{1/t} C) = C$. So the map $X \mapsto A\#_t X$ is surjective. \square

REMARK 1. On the open convex cone \mathbb{P}_n of positive definite Hermitian matrices, first Moakher [24] and then Bhatia and Holbrook [6] suggested the multi-variable geometric mean by taking the mean to be the unique minimizer of the weighted sum of squares of Riemannian trace distances to each variable A_1, \dots, A_m :

$$\Lambda(\omega; A_1, \dots, A_m) = \operatorname{arg\,min}_{X \in \mathbb{P}_n} \sum_{i=1}^m w_i \delta^2(X, A_i)$$

where $\delta(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_2$ denotes the Riemannian trace distance and $\omega = (w_1, \dots, w_m)$ is the positive probability vector in \mathbb{R}^m . This idea in a Riemannian manifold has been anticipated by Élie Cartan, and later by Karcher [11] $\Lambda(\omega; A_1, \dots, A_m)$ coincides with a unique solution $X \in \mathbb{P}_n$ of the non-linear equation, named the Karcher equation,

$$\sum_{i=1}^m w_i \log(X^{1/2} A_i^{-1} X^{1/2}) = 0. \tag{3}$$

We call $\Lambda(\omega; A_1, \dots, A_m)$ the Cartan mean, Riemannian mean or Karcher mean.

This theory does not readily carry over to the setting \mathbb{P} of positive definite operators on a Hilbert space, because there is neither such Riemannian structure nor non-positive curvature metric. Nevertheless, it has been shown in [20] that the Karcher equation (3) has a unique solution in the infinite-dimensional setting \mathbb{P} , so the unique solution can be defined as the Karcher mean $\Lambda(\omega; A_1, \dots, A_m)$. Since $\Lambda(1-t, t; A, B) = A \#_t B$ for any $t \in [0, 1]$ we have that $A \#_t B$ is the unique solution $X \in \mathbb{P}$ of the equation

$$(1-t) \log(X^{1/2} A^{-1} X^{1/2}) + t \log(X^{1/2} B^{-1} X^{1/2}) = 0.$$

2.2. Spectral geometric means

Fiedler and Pták have introduced and studied in [8] a new geometric mean of two positive definite matrices, called the *spectral geometric mean*, which possesses analogous properties of the geometric mean. The spectral geometric mean of A and B , denoted by $A \natural B$, is defined by

$$A \natural B := (A^{-1} \# B)^{1/2} A (A^{-1} \# B)^{1/2}.$$

One of the most important properties is that $(A \natural B)^2$ is positively similar to AB , and hence, the eigenvalues of $A \natural B$ are the same as the positive square roots of the eigenvalues of AB .

H. Lee and Y. Lim [21] have extended the theory of spectral geometric means of positive definite matrices to symmetric cones, and also provided a weighted version of spectral geometric mean with interesting properties. We in this section consider the spectral geometric mean of positive definite operators with its properties. For any $t \in [0, 1]$ the *weighted spectral geometric mean* of A and B in \mathbb{P} is defined as a differentiable curve

$$t \in [0, 1] \mapsto A \natural_t B := (A^{-1} \# B)^t A (A^{-1} \# B)^t. \tag{4}$$

We can easily see that $A \natural_0 B = A$, $A \natural_1 B = B$, and $A \natural_t B = A \natural_{1/2} B$. Analogous to the weighted geometric mean, the weighted spectral geometric mean can be also extended to the curve $t \mapsto A \natural_t B$ for all real numbers $t \in \mathbb{R}$. From the Riccati equation of the geometric mean, we obtain

PROPOSITION 1. *Let $A, B \in \mathbb{P}$ and $t \in \mathbb{R}$. Then $A \natural_t B$ is a unique positive definite solution $X \in \mathbb{P}$ of the equation*

$$(A^{-1} \# B)^t = A^{-1} \# X.$$

The following properties of the weighted spectral geometric mean in [21], which have been proved on $[0, 1]$, can hold for all real numbers by definition.

LEMMA 1. *Let $A, B \in \mathbb{P}$ and $s, t, u \in \mathbb{R}$.*

- (1) $A \natural_t B = A^{1-t} B^t$ if A and B commute.
- (2) $(aA) \natural_t (bB) = a^{1-t} b^t (A \natural_t B)$ for any $a, b > 0$.
- (3) $U^*(A \natural_t B)U = (U^*AU) \natural_t (U^*BU)$ for any unitary operator U .
- (4) $A \natural_t B = B \natural_{1-t} A$.
- (5) $(A \natural_s B) \natural_t (A \natural_u B) = A \natural_{(1-t)s+tu} B$.
- (6) $(A \natural_t B)^{-1} = A^{-1} \natural_t B^{-1}$.

THEOREM 2. *Let $A, B \in \mathbb{P}$.*

- (1) *The map $X \mapsto A \natural_t X$ is a bijection on \mathbb{P} for any $t \in (0, 1]$.*
- (2) *The map $X \mapsto X \natural_t B$ is a bijection on \mathbb{P} for any $t \in [0, 1)$.*

Proof. We first show (1). Then (2) holds easily from (1), since $X \natural_t B = B \natural_{1-t} X$ by Lemma 1 (4). Let $t \in (0, 1]$.

- (i) For $X, Y \in \mathbb{P}$, assume that $A \natural_t X = A \natural_t Y$. Taking congruence transformation by $A^{1/2}$ yields

$$[A^{1/2}(A^{-1} \# X)^t A^{1/2}]^2 = A^{1/2}(A \natural_t X)A^{1/2} = A^{1/2}(A \natural_t Y)A^{1/2} = [A^{1/2}(A^{-1} \# Y)^t A^{1/2}]^2.$$

Then $A^{1/2}(A^{-1} \# X)^t A^{1/2} = A^{1/2}(A^{-1} \# Y)^t A^{1/2}$, so $A^{-1} \# X = A^{-1} \# Y$. By Theorem 1 (1) $X = Y$, that is, the map $X \mapsto A \natural_t X$ is injective.

- (ii) For any $C \in \mathbb{P}$, set $X = A \natural_{1/t} C$. Then $X \in \mathbb{P}$ and $A \natural_t X = A \natural_t (A \natural_{1/t} C) = C$ by Lemma 1 (5). So the map $X \mapsto A \natural_t X$ is surjective. \square

Unfortunately, the spectral geometric mean does not satisfy the monotonicity and arithmetic-geometric-harmonic mean inequalities with respect to the Loewner order on $S(\mathcal{H})$. In [16, Proposition 2.3], on the other hand, an upper bound and a lower bound of the weighted spectral geometric mean are provided.

PROPOSITION 2. Let $A, B \in \mathbb{P}$ and $t \in [0, 1]$. Then

$$2^{1+t}(A + B^{-1})^{-t} - A^{-1} \leq A \#_t B \leq [2^{1+t}(A^{-1} + B)^{-t} - A]^{-1}.$$

No metric property for the spectral geometric mean of positive definite operators and matrices is known. We show that the spectral geometric mean is a midpoint with respect to the following semi-metric on \mathbb{P} .

LEMMA 2. The map $d : \mathbb{P} \times \mathbb{P} \rightarrow [0, \infty)$ given by

$$d(A, B) = 2 \|\log(A^{-1} \# B)\|$$

is a semi-metric, where $\|\cdot\|$ denotes the operator norm.

Proof. We need to show that the map $d : \mathbb{P} \times \mathbb{P} \rightarrow [0, \infty)$ satisfies all axioms of metric except the triangle inequality.

- (1) Obviously $d(A, B) \geq 0$ for all $A, B \in \mathbb{P}$.
- (2) If $A = B$ then $A^{-1} \# B = I$, and $\log(A^{-1} \# B) = O$. So $d(A, B) = 0$. Conversely, $d(A, B) = 0$ implies $\log(A^{-1} \# B) = O$, that is, $A^{-1} \# B = I$. Then $(A^{1/2} B A^{1/2})^{1/2} = A$, and $A^{1/2} B A^{1/2} = A^2$. Thus, $B = A$.
- (3) By the invariance under inversion and symmetry of the geometric mean

$$\begin{aligned} d(A, B) &= 2 \|\log(A^{-1} \# B)\| = 2 \|-\log(A^{-1} \# B)\| = 2 \|\log(A^{-1} \# B)^{-1}\| \\ &= 2 \|\log(A \# B^{-1})\| = 2 \|\log(B^{-1} \# A)\| = d(B, A). \quad \square \end{aligned}$$

We see several interesting properties of the semi-metric d .

PROPOSITION 3. Let $A, B \in \mathbb{P}$.

- (1) $d(\alpha A, \alpha B) = d(A, B)$ for any $\alpha > 0$.
- (2) $d(A^{-1}, B^{-1}) = d(A, B)$.
- (3) $d(UAU^*, UBU^*) = d(A, B)$ for any unitary operator U .
- (4) $d(A^t, B^t) = |t|d(A, B)$ for any commuting operators A, B and $t \in \mathbb{R}$.

Proof. The items (1) and (2) follow from the homogeneity and invariance under inversion of the geometric mean $\#$.

- (3) For any unitary operator U , $(UAU^*)^{-1} \# UBU^* = UA^{-1}U^* \# UBU^* = U(A^{-1} \# B)U^*$, so

$$\log[(UAU^*)^{-1} \# UBU^*] = U \log(A^{-1} \# B) U^*.$$

Since the operator norm is unitarily invariant, $d(UAU^*, UBU^*) = 2 \|\log(A^{-1} \# B)\| = d(A, B)$.

- (4) Since $A^{-1}\#B = A^{-1/2}B^{1/2}$ for any commuting A, B , we have $d(A, B) = \|\log A - \log B\|$. So

$$d(A^t, B^t) = \|\log A^t - \log B^t\| = |t| \|\log A - \log B\| = |t|d(A, B). \quad \square$$

THEOREM 3. *The spectral geometric mean $A\sharp_t B$ is a midpoint of A and B with respect to the semi-metric d on \mathbb{P} .*

Proof. By Proposition 1 for $t = 1/2$ we have $(A^{-1}\#B)^{1/2} = A^{-1}\#(A\sharp_{1/2}B)$ for positive definite operators A and B . Then

$$d(A, A\sharp_{1/2}B) = 2\|\log(A^{-1}\#(A\sharp_{1/2}B))\| = 2\|\log(A^{-1}\#B)^{1/2}\| = \|\log(A^{-1}\#B)\| = \frac{1}{2}d(A, B).$$

Similarly, by Lemma 1 (4), the above result and the symmetry of semi-metric d

$$d(B, A\sharp_{1/2}B) = d(B, B\sharp_{1/2}A) = \frac{1}{2}d(B, A) = \frac{1}{2}d(A, B). \quad \square$$

REMARK 2. We give some remarks and an open question for the semi-metric.

- (1) Note that the semi-metric d can be considered as a symmetric divergence. Unfortunately, the semi-metric d does not hold the triangle inequality. For instance, let

$$A = \begin{pmatrix} 5 & 0 \\ 0 & 1/5 \end{pmatrix}, B = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}.$$

Then $A, B, C \in \mathbb{P}_2$ with determinant 1, and by using MATLAB

$$d(A, B) = 1.117270\dots, d(B, C) = 0.173732\dots, d(A, C) = 1.305274\dots.$$

So $d(A, C) > d(A, B) + d(B, C)$.

- (2) Note that for any $t \in \mathbb{R}$

$$d(A, A\sharp_t B) = |t|d(A, B) \quad \text{and} \quad d(B, A\sharp_t B) = |1 - t|d(A, B).$$

Indeed, by Proposition 1

$$\begin{aligned} d(A, A\sharp_t B) &= 2\|\log(A^{-1}\#(A\sharp_t B))\| = 2\|\log(A^{-1}\#B)^t\| \\ &= 2|t| \cdot \|\log(A^{-1}\#B)\| = |t|d(A, B). \end{aligned}$$

Similarly, $d(B, A\sharp_t B) = d(B, B\sharp_{1-t}A) = |1 - t|d(B, A) = |1 - t|d(A, B)$ by Lemma 1 (4). It gives an affirmative answer that $A\sharp_t B$ is a geodesic for the semi-metric d , but it is a challengeable problem. That is, for any $s, t \in \mathbb{R}$

$$d(A\sharp_s B, A\sharp_t B) = |s - t|d(A, B).$$

3. Operator inequalities

In this section we see several operator inequalities of weighted geometric means. The following are crucial to prove the main results.

LEMMA 3. (The Loewner-Heinz inequality) *Let $C \in S(\mathcal{H})$ and let $A, B \geq 0$. If $C^2 \leq A \leq B$, then $C \leq A^{1/2} \leq B^{1/2}$.*

Lemma 3 yields that for positive semi-definite operators A and B ,

$$A \leq B \text{ implies } A^t \leq B^t \text{ for any } t \in [0, 1].$$

This is also known as the Loewner-Heinz inequality, which is the most used in the fields of operator inequalities and operator means.

The following are well-known results of the Loewner partial order applied to positive operators.

LEMMA 4. *Let $X, Y \in S(\mathcal{H})$ and $S \in B(\mathcal{H})$.*

$$(1) \ 0 \leq X \leq Y \text{ implies } 0 \leq SXS^* \leq SYS^*.$$

$$(2) \ 0 < X \leq Y \text{ if and only if } 0 < Y^{-1} \leq X^{-1}.$$

The following has been proved for positive definite Hermitian matrices in [5, Corollary 4.4.5], but it is also valid for positive definite operators.

LEMMA 5. *Let $0 \leq B \leq A$. Then for any $p > 0$*

$$A^p \# B^{-p} \geq I.$$

THEOREM 4. *Let $A, B \in \mathbb{P}$ with $A \# B \leq I$. Then*

$$A^{p+1} \# (A \#_{\frac{p}{2}} B) \leq A$$

for any $p > 0$. Especially, $A^3 \# B \leq A$ for $p = 2$.

Proof. Assume that $A \# B \leq I$ for $A, B \in \mathbb{P}$. Taking the congruence transformation by $A^{-1/2}$ on both sides implies $(A^{-1/2}BA^{-1/2})^{1/2} \leq A^{-1}$ by Lemma 4 (1). By applying Lemma 5 we have

$$A^{-p} \# (A^{-1/2}BA^{-1/2})^{-p/2} \geq I$$

for any $p > 0$. Taking inverse on both sides and congruence transformation by $A^{1/2}$ yield

$$A^{p+1} \# (A \#_{\frac{p}{2}} B) = A^{p+1} \# A^{1/2} (A^{-1/2}BA^{-1/2})^{p/2} A^{1/2} \leq A$$

from Lemma 4. \square

THEOREM 5. *Let $A, B \in \mathbb{P}$. Then the following are equivalent.*

- (1) $A^{-1}\sharp B \leq I$;
- (2) $A\sharp B^{-1} \geq I$;
- (3) $A\#B \leq A$;
- (4) $A\#B \geq B$;
- (5) $B \leq A$.

Proof. By Lemma 1 (6), items (1) and (2) are equivalent. Moreover, items (3) and (4) are equivalent from [8, Theorem 4.2]

We now show (1) \implies (3) \implies (5) \implies (1).

- (1) \implies (3) Assume that $A^{-1}\sharp B \leq I$, that is,

$$(A\#B)^{1/2}A^{-1}(A\#B)^{1/2} \leq I.$$

By Lemma 4 (1) we have $A^{-1} \leq (A\#B)^{-1}$, equivalently $A\#B \leq A$ by Lemma 4 (2).

- (3) \implies (5) Assume that $A\#B \leq A$. Using Lemma 4 (1) again, we have

$$(A^{-1/2}BA^{-1/2})^{1/2} \leq I.$$

This implies $A^{-1/2}BA^{-1/2} \leq I$, and so $B \leq A$.

- (5) \implies (1) This follows from the opposite steps shown as (1) \implies (3) \implies (5). \square

LEMMA 6. Let $S \in S(\mathcal{H})$ and let $X \in \mathbb{P}$. Then $SXS \leq X$ implies $S \leq I$.

Proof. Assume that $SXS \leq X$. Then by Lemma 4 (1)

$$(X^{1/2}SX^{1/2})^2 = X^{1/2}(SXS)X^{1/2} \leq X^2.$$

By Lemma 3, $X^{1/2}SX^{1/2} \leq X$. Using Lemma 4 (1) again, we obtain $S \leq I$. \square

REMARK 3. In general, the converse of Lemma 6 does not hold. For instance, let

$$S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, X = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathbb{P}_2,$$

where $a, c > 0$, $a \neq c$ and $ac - b^2 > 0$. Note that S is a permutation matrix, so $S^2 = I$. That is, S is a contraction. However,

$$X - SXS = \begin{pmatrix} a & b \\ b & c \end{pmatrix} - \begin{pmatrix} c & b \\ b & a \end{pmatrix} = \begin{pmatrix} a - c & 0 \\ 0 & c - a \end{pmatrix},$$

which is not positive semi-definite, in general.

Ando and Hiai [3] have established a characterization of two-variable geometric means such as

$$A\#B \leq I \text{ implies } A^p\#B^p \leq I$$

for all $p \geq 1$ and positive definite matrices A and B . We call it the *Ando-Hiai inequality*. Uchiyama [28] has extended the the Ando-Hiai inequality to the weighted geometric mean, and Yamazaki [30] generalized it to the Riemannian mean (or Cartan mean, Karcher mean) of positive definite matrices, which is the multi-variable geometric mean defined as the unique minimizer of the weighted sum of squares of the Riemannian trace distances to each variable. Lawson and Lim [20] naturally defined the Karcher mean of positive definite operators as the unique positive definite solution of the Karcher equation, and extended Yamazaki's result to positive definite operators.

We give a different sufficient condition for the Ando-Hiai inequality in terms of the spectral geometric mean.

THEOREM 6. *Let $A, B \in \mathbb{P}$ and $t \in (0, 1]$. Then $A^{-1}\natural_t B \leq A^{-1}$ implies $A^p\#B^p \leq I$ for any $p \geq 1$.*

Proof. Assume that $A^{-1}\natural_t B = (A\#B)^t A^{-1} (A\#B)^t \leq A^{-1}$. By Lemma 6 we have $(A\#B)^t \leq I$, equivalently $A\#B \leq I$. By the Ando-Hiai inequality we obtain $A^p\#B^p \leq I$ for any $p \geq 1$. \square

REMARK 4. The following are sufficient conditions of the inequality $A\#B \leq I$ for $A, B \in \mathbb{P}$:

- (1) $A \leq I$ and $B \leq I$;
- (2) $\log A + \log B \leq 0$;
- (3) $A^{-1}\natural_t B \leq A^{-1}$ for any $t \in (0, 1]$.

Obviously, (1) implies (2) since the logarithmic map is operator monotone, but the converse is not true in general. On the other hand, the relations between (1) and (3), or between (2) and (3) do not know yet. So (3) is a new sufficient condition of the inequality $A\#B \leq I$, and also of the Ando-Hiai inequality: $A^p\#B^p \leq I$ for any $p \geq 1$.

4. Gyrogroups and gyrolines

We review in this section the algebraic structure of a gyrogroup as a natural extension of a group into the regime of the nonassociative algebra. We then introduce a gyrovector space providing the setting for hyperbolic geometry just as a vector space provides the setting for Euclidean geometry. A. A. Ungar has introduced and intensely studied the theory of gyrogroups and gyrovector spaces in a series of papers and books; see [29] and its bibliography. We also see what geometric means of positive definite Hermitian matrices are in the sense of gyrovector spaces.

4.1. Gyrogroups and gyrovector spaces

The binary operation in a gyrogroup is not associative, in general. The breakdown of associativity for gyrogroup operations is salvaged in a modified form, called gyroassociativity. The axioms for a (gyrocommutative) gyrogroup G is reminiscent of those for a (commutative) group.

DEFINITION 1. A binary system (G, \oplus) is a *gyrogroup* if it satisfies the following axioms for all $a, b, c \in G$:

(G1) $e \oplus a = a \oplus e = a$ (existence of identity);

(G2) $a \oplus (\ominus a) = (\ominus a) \oplus a = e$ (existence of inverses);

(G3) There is an automorphism $\text{gyr}[a, b] : G \rightarrow G$ for each $a, b \in G$ such that

$$a \oplus (b \oplus c) = (a \oplus b) \oplus \text{gyr}[a, b]c \quad (\text{gyroassociativity});$$

(G4) $\text{gyr}[a \oplus b, b] = \text{gyr}[a, b]$ (loop property).

A gyrogroup (G, \oplus) is *gyrocommutative* if it satisfies

$$a \oplus b = \text{gyr}[a, b](b \oplus a) \quad (\text{gyrocommutativity}).$$

A gyrogroup is *uniquely 2-divisible* if for every $b \in G$, there exists a unique element $a \in G$ such that $a \oplus a = b$.

In (G3) the automorphism $\text{gyr}[a, b]$ for each $a, b \in G$ is called the *Thomas gyration* or the *gyroautomorphism*, simply *gyration*. From (G2) and (G3) we have

$$\text{gyr}[a, b]c = \ominus(a \oplus b) \oplus [a \oplus (b \oplus c)] \tag{5}$$

for all $a, b, c \in G$. In Euclidean space it plays a role of rotation in the plane spanned by $\{a, b\}$ leaving the orthogonal complement fixed.

DEFINITION 2. Let (G, \oplus) be a gyrogroup. The *gyrogroup cooperation* is a binary operation in G given by

$$a \boxplus b = a \oplus \text{gyr}[a, \ominus b]b$$

for all $a, b \in G$. The groupoid (G, \boxplus) is said to be the *cogyrogroup* associated with the gyrogroup (G, \oplus) .

The gyrogroup cooperation \boxplus gives a useful criterion for the gyrogroup to be gyrocommutative. The following has been shown in Theorem 3.2 and Theorem 3.4 [29] and Lemma 2.1 [13].

THEOREM 7. *Let (G, \oplus) be a gyrogroup. For any $a, b \in G$ the following are equivalent.*

- (1) G is gyrocommutative.
- (2) (Automorphic inverse property) $\ominus(a \oplus b) = \ominus a \oplus b$.
- (3) (Bruck identity) $(a \oplus b) \oplus (a \oplus b) = a \oplus (b \oplus (b \oplus a))$.
- (4) (Commutativity of coaddition) $a \boxplus b = b \boxplus a$.

In the same way that vector spaces are commutative groups of vectors admitted scalar multiplication, gyrovector spaces are gyrocommutative gyrogroups of gyrovectors admitted properly scalar multiplication. We give a definition of gyrovector spaces slightly different from Definition 6.2 in [29].

DEFINITION 3. A gyrocommutative gyrogroup (G, \oplus) equipped with a scalar multiplication

$$(t, x) \mapsto t \otimes x : \mathbb{R} \times G \rightarrow G$$

is called a *gyrovector space* if it satisfies the following for $s, t \in \mathbb{R}$ and $a, b, c \in G$.

- (V1) $1 \otimes a = a$, $0 \otimes a = t \otimes e = e$, and $(-1) \otimes a = \ominus a$.
- (V2) $(s + t) \otimes a = s \otimes a \oplus t \otimes a$.
- (V3) $(st) \otimes a = s \otimes (t \otimes a)$.
- (V4) $\text{gyr}[a, b](t \otimes c) = t \otimes \text{gyr}[a, b]c$.

We now see some typical examples of the gyrovector space in the Euclidean space \mathbb{R}^n . We consider elements in \mathbb{R}^n naturally as column vectors, so that $\mathbf{u}^T \mathbf{v}$ is the usual inner product of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ written in matrix form. A. Ungar has introduced in [29] two distinctive examples of gyrovector spaces in the open unit ball \mathbf{B} of the n -dimensional vector space \mathbb{R}^n , also corresponding to two models of hyperbolic geometry.

EXAMPLE 1. We define the binary operations \oplus_E and \oplus_M in \mathbf{B} by

$$\mathbf{u} \oplus_E \mathbf{v} = \frac{1}{1 + \mathbf{u}^T \mathbf{v}} \left\{ \mathbf{u} + \frac{1}{\gamma_{\mathbf{u}}} \mathbf{v} + \frac{\gamma_{\mathbf{u}}}{1 + \gamma_{\mathbf{u}}} (\mathbf{u}^T \mathbf{v}) \mathbf{u} \right\}, \tag{6}$$

$$\mathbf{u} \oplus_M \mathbf{v} = \frac{1}{1 + 2\mathbf{u}^T \mathbf{v} + \|\mathbf{u}\|^2 \|\mathbf{v}\|^2} \left\{ (1 + 2\mathbf{u}^T \mathbf{v} + \|\mathbf{v}\|^2) \mathbf{u} + (1 - \|\mathbf{u}\|^2) \mathbf{v} \right\} \tag{7}$$

for any $\mathbf{u}, \mathbf{v} \in \mathbf{B}$, where $\gamma_{\mathbf{v}} = \frac{1}{\sqrt{1 - \|\mathbf{v}\|^2}}$ is the well-known Lorentz (gamma) factor. The binary systems (\mathbf{B}, \oplus_E) and (\mathbf{B}, \oplus_M) form (uniquely 2-divisible) gyrocommutative gyrogroups called the *Einstein gyrogroup* and *Möbius gyrogroup*, respectively.

We define a map $\otimes : \mathbb{R} \times \mathbf{B} \rightarrow \mathbf{B}$ by

$$\begin{aligned} t \otimes \mathbf{v} &= \frac{(1 + \|\mathbf{v}\|)^t - (1 - \|\mathbf{v}\|)^t}{(1 + \|\mathbf{v}\|)^t + (1 - \|\mathbf{v}\|)^t} \frac{\mathbf{v}}{\|\mathbf{v}\|} \\ &= \tanh(t \tanh^{-1} \|\mathbf{v}\|) \frac{\mathbf{v}}{\|\mathbf{v}\|}, \end{aligned} \tag{8}$$

for $t \in \mathbb{R}$ and $\mathbf{v}(\neq \mathbf{0}) \in \mathbf{B}$, and define $t \otimes \mathbf{0} := \mathbf{0}$. We call $(\mathbf{B}, \oplus_E, \otimes)$ and $(\mathbf{B}, \oplus_M, \otimes)$ the *Einstein gyrovector space* and the *Möbius gyrovector space*, respectively.

The Beltrami-Klein ball model of hyperbolic geometry is algebraically regulated by Einstein gyrovector spaces. The geodesics of this model, called gyrolines, are Euclidean straight lines in the open unit ball. On the other hand, the Poincaré ball model of hyperbolic geometry is algebraically regulated by Möbius gyrovector spaces. The geodesics of this model are Euclidean circular arcs in the open unit ball that intersect the boundary of the ball orthogonally.

EXAMPLE 2. [13, Example 2.2, Example 3.2] We define the binary operation \oplus and a scalar multiplication \circ on the open convex cone \mathbb{P}_n of positive definite matrices by

$$\begin{aligned} \oplus : \mathbb{P}_n \times \mathbb{P}_n &\rightarrow \mathbb{P}_n, A \oplus B = A^{1/2}BA^{1/2}, \\ \circ : \mathbb{R} \times \mathbb{P}_n &\rightarrow \mathbb{P}_n, t \circ A = A^t \end{aligned}$$

for any $A, B \in \mathbb{P}_n$ and $t \in \mathbb{R}$. Then the system $(\mathbb{P}_n, \oplus, \circ)$ forms a gyrovector space, and the gyroautomorphism generated by A and B is given by

$$\text{gyr}[A, B]X = U(A, B)XU(A, B)^{-1}, \tag{9}$$

where $U(A, B) = (A^{1/2}BA^{1/2})^{-1/2}A^{1/2}B^{1/2}$ is a unitary part of the polar decomposition for $A^{1/2}B^{1/2}$ such that

$$A^{1/2}B^{1/2} = (A \oplus B)^{1/2}U(A, B).$$

REMARK 5. The inner product on $M_n(\mathbb{C})$, the vector space of all $n \times n$ matrices with complex entries, is naturally defined as $\langle A, B \rangle = \text{tr}(AB^*)$. The gyroautomorphism on \mathbb{P}_n preserves the inner product, and so the norm induced by inner product. Indeed, for any $A, B, X, Y \in \mathbb{P}_n$

$$\begin{aligned} \langle \text{gyr}[A, B]X, \text{gyr}[A, B]Y \rangle &= \text{tr}[U(A, B)XU(A, B)^{-1}(U(A, B)YU(A, B)^{-1})^*] \\ &= \text{tr}[U(A, B)XY^*U(A, B)^*] \\ &= \text{tr}[XY^*] = \langle X, Y \rangle. \end{aligned}$$

While A. Ungar has explained a gyrogroup structure for 2×2 density matrices in Chapter 9, [29], we now see an example of gyrovector space for arbitrary dimensional density matrices.

EXAMPLE 3. [12] Let \mathbb{D}_n be a set of all $n \times n$ invertible density matrices, which are positive definite Hermitian matrices of trace 1. We define a binary operation \odot and a scalar multiplication \star given by

$$\begin{aligned} \odot : \mathbb{D}_n \times \mathbb{D}_n &\rightarrow \mathbb{D}_n, \rho \odot \sigma = \frac{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}{\text{tr}(\rho\sigma)} = \frac{\rho \oplus \sigma}{\text{tr}(\rho \oplus \sigma)} \\ \star : \mathbb{R} \times \mathbb{D}_n &\rightarrow \mathbb{D}_n, t \star \rho = \frac{\rho^t}{\text{tr}(\rho^t)} = \frac{t \circ \rho}{\text{tr}(t \circ \rho)} \end{aligned}$$

for any $\rho, \sigma \in \mathbb{D}_n$ and $t \in \mathbb{R}$. Then $(\mathbb{D}_n, \odot, \star)$ is a gyrovector space. Note that the identity element in $(\mathbb{D}_n, \odot, \star)$ is $\frac{1}{n}I_n$ and the inverse of ρ is $(-1) \star \rho = \frac{1}{\text{tr}(\rho^{-1})}\rho^{-1}$, where I_n denotes the $n \times n$ identity matrix.

4.2. Gyrolines and cogyrolines

The *gyroline* passing through the points a and b in the gyrovector space (G, \oplus, \otimes) is defined by

$$L : \mathbb{R} \times G \times G \rightarrow G, L(t; a, b) = a \oplus t \otimes (\ominus a \oplus b). \tag{10}$$

The gyroline is uniquely determined by given points, and a left gyrotranslation of a gyroline is again a gyroline by Theorem 6.21 in [29]. In other words,

$$x \oplus L(t; a, b) = L(t; x \oplus a, x \oplus b)$$

for any $x \in G$. The point $L(1/2; a, b) = a \oplus \frac{1}{2} \otimes (\ominus a \oplus b)$ is called the *gyromidpoint* of given two points a and b in the gyrovector space (G, \oplus, \otimes) . By using the gyrogroup cooperation, we can write alternatively

$$L\left(\frac{1}{2}; a, b\right) = \frac{1}{2} \otimes (a \boxplus b). \tag{11}$$

Indeed,

$$\begin{aligned} 2 \otimes L\left(\frac{1}{2}; a, b\right) &= a \oplus [(\ominus a) \oplus b] \oplus a \\ &= a \oplus [\ominus(a \oplus b) \oplus \{a \oplus (\ominus b \oplus b)\}] \\ &= a \oplus \text{gyr}[a, \ominus b]b \\ &= a \boxplus b. \end{aligned}$$

The first equality follows from the Bruck identity in Theorem 7 (3), the second from the automorphic inverse property in Theorem 7 (2) together with (G1) and (G2), the third from (5), and the last from the definition of coaddition in Definition 2. Thus, (11) holds by (V3) of Definition 3.

On the other hand, the *cogyroline* passing through the points a and b in the gyrovector space (G, \oplus, \otimes) is defined by

$$L^c : \mathbb{R} \times G \times G \rightarrow G, L^c(t; a, b) = t \otimes (\ominus a \boxplus b) \oplus a. \tag{12}$$

The gyroline is uniquely determined by given points, and the point $L^c(1/2; a, b) = \frac{1}{2} \otimes (\ominus a \boxplus b) \oplus a$ is called the *cogyromidpoint* of given two points a and b in the gyrovector space (G, \oplus, \otimes) .

Gyrolines and cogyrolines play an important role in the hyperbolic analytic geometry regulated by the gyrovector space. See Chapter 6 in [29] for more information. We here see the connection of the weighted (spectral) geometric mean and the gyroline (cogyroline, respectively) on the open convex cone \mathbb{P} of positive definite matrices and on the gyrovector space \mathbb{D}_n of invertible density matrices.

THEOREM 8. For any $A, B \in (\mathbb{P}_n, \oplus, \circ)$ and $t \in \mathbb{R}$

$$L(t; A, B) = A \#_t B, \quad L^c(t; A, B) = A \natural_t B.$$

Proof. From Example 2 we can easily obtain the gyroline on $(\mathbb{P}_n, \oplus, \circ)$ passing through A and B such as

$$\begin{aligned} L(t; A, B) &= A \oplus t \circ (((-1) \circ A) \oplus B) \\ &= A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2} = A \#_t B \end{aligned}$$

for any $t \in \mathbb{R}$. Moreover, since

$$\begin{aligned} ((-1) \circ A) \boxplus B &= A^{-1} \oplus \text{gyr}[A^{-1}, B^{-1}] B \\ &= A^{-1/2} (A^{1/2} B A^{1/2})^{1/2} A^{-1} (A^{1/2} B A^{1/2})^{1/2} A^{-1/2} \\ &= (A^{-1} \# B)^2, \end{aligned}$$

the cgyroline passing through A and B is

$$L^c(t; A, B) = t \circ (((-1) \circ A) \boxplus B) \oplus A = (A^{-1} \# B)^t A (A^{-1} \# B)^t = A \natural_t B. \quad \square$$

THEOREM 9. For any $\rho, \sigma \in (\mathbb{D}_n, \odot, \star)$ and $t \in \mathbb{R}$

$$\begin{aligned} L(t; \rho, \sigma) &= \frac{1}{\text{tr}(\rho \#_t \sigma)} \rho \#_t \sigma, \\ L^c(t; \rho, \sigma) &= \frac{1}{\text{tr}(\rho \natural_t \sigma)} \rho \natural_t \sigma. \end{aligned}$$

Proof. Let $\rho, \sigma \in (\mathbb{D}_n, \odot, \star)$ and $t \in \mathbb{R}$. The formula of $L(t; \rho, \sigma)$ has been shown in [14, Theorem 4.2]. Note that

$$\begin{aligned} ((-1) \star \rho) \boxtimes \sigma &= ((-1) \star \rho) \odot \text{gyr}[(-1) \star \rho, (-1) \star \sigma] \sigma \\ &= ((-1) \star \rho) \odot \text{gyr}[\rho, \sigma] \sigma \\ &= \frac{((-1) \circ \rho) \boxplus \sigma}{\text{tr}(((-1) \circ \rho) \boxplus \sigma)}. \end{aligned}$$

The first equality follows from Definition 2 and the second follows from Theorem 2.34 in [29]. Thus, by the formula of $L^c(t; A, B)$ in Theorem 8 we obtain

$$\begin{aligned} L^c(t; \rho, \sigma) &= t \star [((-1) \star \rho) \boxtimes \sigma] \odot \rho \\ &= \frac{t \circ [((-1) \circ \rho) \boxplus \sigma] \oplus \rho}{\text{tr}(t \circ [((-1) \circ \rho) \boxplus \sigma] \oplus \rho)} \\ &= \frac{\rho \natural_t \sigma}{\text{tr}(\rho \natural_t \sigma)}. \quad \square \end{aligned}$$

5. Special cases: 2×2 positive definite matrices

We compute in this section the weighted geometric means of 2×2 positive definite matrices as the special cases, but important examples in the theoretical and applied areas.

5.1. 2×2 positive definite matrices with determinant one

For any $A, B \in \mathbb{P}_2$ with determinant 1, we can write the geometric mean of A and B to linear combination of A and B . For this goal we introduce a map $L_f : (0, \infty) \rightarrow \mathbb{R}$ constructed by a differentiable function $f : (0, \infty) \rightarrow \mathbb{R}$:

$$L_f(x) = \begin{cases} \frac{f(x) - f(x^{-1})}{x - x^{-1}}, & x \neq 1; \\ f'(1), & x = 1. \end{cases}$$

For $f(x) = x^t$ we simply write $L_f(x)$ as $L_t(x)$.

LEMMA 7. For any $A, B \in \mathbb{P}_2$ with determinant 1,

$$A \#_t B = L_{1-t}(\lambda)A + L_t(\lambda)B, \quad t \in [0, 1]$$

where λ is an eigenvalue of AB^{-1} .

Proof. Assume that $A \neq B$ for $A, B \in \mathbb{P}_2$ with determinant 1. It has been shown from [22] that there exists a unique minimal geodesic connecting A and B for the Thompson metric d_T on \mathbb{P}_2 . Note from [7, 19] that $t \mapsto A \#_t B$ is a minimal geodesic for the Thompson metric, and from [25] that $t \mapsto L_{1-t}(\lambda)A + L_t(\lambda)B$ is a minimal geodesic for the Thompson metric. By uniqueness, we obtain the desired identity. \square

REMARK 6. It is known from [5, Proposition 4.1.12] that

$$A \# B = \frac{A + B}{\sqrt{\det(A + B)}}$$

for any $A, B \in \mathbb{P}_2$ with determinant 1. So by Lemma 7 we have

$$L_{\frac{1}{2}}(\lambda) = \frac{1}{\sqrt{\det(A + B)}},$$

where λ is an eigenvalue of AB^{-1} .

LEMMA 8. [2] For any $A, B, X \in B(\mathcal{H})$,

$$\begin{pmatrix} A & X \\ X^* & B \end{pmatrix} \geq 0,$$

implies $\|X\| \leq \sqrt{\|A\| \cdot \|B\|}$ for operator norm $\|\cdot\|$.

Applying Lemma 8 to \mathbb{P}_2 we obtain some interesting inequalities.

COROLLARY 1. *Let $A, B \in \mathbb{P}_2$ with determinants 1. Then*

$$\|A + B\| \leq \sqrt{\det(A + B)\|A\| \cdot \|B\|}.$$

Proof. Since

$$\begin{pmatrix} A & A\#B \\ A\#B & B \end{pmatrix} \geq 0,$$

Lemma 8 together with Proposition 4.1.12 in [5] yields the desired inequality. \square

REMARK 7. Let $A, B \in \mathbb{P}_2$ with determinants $\det A = \alpha^2$ and $\det B = \beta^2$, respectively. Then replacing A and B by $\alpha^{-1}A$ and $\beta^{-1}B$ in Corollary 1 implies

$$\sqrt{\alpha\beta}\|\alpha^{-1}A + \beta^{-1}B\| \leq [\det(\alpha^{-1}A + \beta^{-1}B)\|A\| \cdot \|B\|]^{1/2}.$$

We give a simple formula for the spectral geometric mean of 2×2 positive definite matrices. The following can be derived from the Cayley-Hamilton theorem for any 2×2 matrix.

LEMMA 9. *For any 2×2 matrix X and any $c \in \mathbb{R}$,*

$$\det(cI + X) = c^2 + \text{ctr}(X) + \det(X).$$

PROPOSITION 4. *Let $A, B \in \mathbb{P}_2$ with determinants 1. Then for any $t \in \mathbb{R}$*

$$A\sharp_t B = \frac{(A^{-1} + B)^t A (A^{-1} + B)^t}{[2 + \text{tr}(AB)]^t}.$$

Proof. It has been proved in [5, Proposition 4.1.12] that

$$A\#B = \frac{A + B}{\sqrt{\det(A + B)}}.$$

Since $\det(A^{-1}) = 1$, we obtain from the equation (4)

$$A\sharp_t B = \left[\frac{A^{-1} + B}{\sqrt{\det(A^{-1} + B)}} \right]^t A \left[\frac{A^{-1} + B}{\sqrt{\det(A^{-1} + B)}} \right]^t = \frac{(A^{-1} + B)^t A (A^{-1} + B)^t}{\det(A^{-1} + B)^t}.$$

Since $\det(A^{-1} + B) = \det(A^{-1}(I + AB)) = \det(I + AB) = 2 + \text{tr}(AB)$ by Lemma 9, the proof is complete. \square

COROLLARY 2. *Let $A, B \in \mathbb{P}_2$ such that $\det(A) = \alpha^2$ and $\det(B) = \beta^2$ for some $\alpha, \beta > 0$. Then for any $t \in \mathbb{R}$*

$$A\sharp_t B = \left[\frac{1}{2\alpha\beta + \text{tr}(AB)} \right]^t (\alpha\beta A^{-1} + B)^t A (\alpha\beta A^{-1} + B)^t.$$

Proof. By Proposition 4 and Lemma 9

$$\begin{aligned} \left(\frac{A}{\alpha}\right) \natural_t \left(\frac{B}{\beta}\right) &= \frac{(\alpha A^{-1} + \beta^{-1} B)^t \alpha^{-1} A (\alpha A^{-1} + \beta^{-1} B)^t}{\det(I + (\alpha\beta)^{-1} AB)^t} \\ &= \frac{\alpha^{2t-1} (\alpha\beta A^{-1} + B)^t A (\alpha\beta A^{-1} + B)^t}{\det(\alpha\beta I + AB)^t} \\ &= \frac{\alpha^{2t-1} (\alpha\beta A^{-1} + B)^t A (\alpha\beta A^{-1} + B)^t}{[2(\alpha\beta)^2 + \alpha\beta \operatorname{tr}(AB)]^t} \end{aligned}$$

Since $\left(\frac{A}{\alpha}\right) \natural_t \left(\frac{B}{\beta}\right) = \alpha^{t-1} \beta^{-t} A \natural_t B$ by Lemma 1 (2), it leads our result by a simple calculation. \square

5.2. 2×2 invertible density matrices

In quantum information theory, a physical state can be described as a density matrix ρ on a complex Hilbert space, which is a positive semi-definite Hermitian matrix with trace 1. In particular, a qubit state is the 2×2 density matrix whose form is given by

$$\rho_{\mathbf{v}} = \frac{1}{2} \begin{pmatrix} 1 + v_3 & v_1 - iv_2 \\ v_1 + iv_2 & 1 - v_3 \end{pmatrix},$$

where $\mathbf{v} = (v_1 \ v_2 \ v_3)^T \in \mathbb{R}^3$. So the qubit state $\rho_{\mathbf{v}}$ is parameterized by some $\mathbf{v} \in \mathbb{R}^3$ such that $\|\mathbf{v}\| \leq 1$. In this case the vector \mathbf{v} is known as the Bloch vector or Bloch vector representation of the state $\rho_{\mathbf{v}}$.

REMARK 8. Via a characteristic equation of the qubit state $\rho_{\mathbf{v}}$, we obtain that its eigenvalues are

$$\frac{1 + \|\mathbf{v}\|}{2}, \frac{1 - \|\mathbf{v}\|}{2},$$

and its determinant is

$$\det \rho_{\mathbf{v}} = \frac{1 - \|\mathbf{v}\|^2}{4} = \frac{1}{4\gamma_{\mathbf{v}}^2}.$$

So the mixed states are parameterized by the open unit ball \mathbf{B} in \mathbb{R}^3 , meanwhile the pure states are parameterized by the boundary of \mathbf{B} , the unit sphere (Bloch sphere). See [12, 29] for more information.

It has been shown in [12, 14] that the map

$$\rho : (\mathbf{B}, \oplus_E, \otimes) \rightarrow (\mathbb{D}_2, \odot, \star), \mathbf{v} \mapsto \rho_{\mathbf{v}}$$

is an isomorphism between two gyrovector spaces. On the gyrovector space \mathbb{D}_2 , moreover, we have that the identity is $\frac{1}{2}I_2$ and the inverse for $\rho_{\mathbf{u}}$ is

$$\rho_{-\mathbf{u}} = \frac{1}{4\gamma_{\mathbf{u}}^2} \rho_{\mathbf{u}}^{-1}. \tag{13}$$

PROPOSITION 5. For any $\mathbf{u}, \mathbf{v} \in \mathbf{B}$ and $t \in [0, 1]$

$$\rho_{\mathbf{u}\#\mathbf{t}\rho_{\mathbf{v}}} = L_{1-t}(\mu) \left(\frac{\gamma_{\mathbf{u}}}{\gamma_{\mathbf{v}}}\right)^t \rho_{\mathbf{u}} + L_t(\mu) \left(\frac{\gamma_{\mathbf{v}}}{\gamma_{\mathbf{u}}}\right)^{1-t} \rho_{\mathbf{v}},$$

where $\mu = \gamma_{\mathbf{u}}\gamma_{\mathbf{v}}(1 - \mathbf{u}^T\mathbf{v})(1 \pm \|\mathbf{u} \oplus_E (-\mathbf{v})\|)$.

Proof. Let $A = 2\gamma_{\mathbf{u}}\rho_{\mathbf{u}}$ and $B = 2\gamma_{\mathbf{v}}\rho_{\mathbf{v}}$. By Remark 8 one can see that A, B are 2×2 positive definite matrices with determinant 1. By the joint homogeneity of two-variable geometric mean and Lemma 7

$$\begin{aligned} \rho_{\mathbf{u}\#\mathbf{t}\rho_{\mathbf{v}}} &= \left(\frac{1}{2\gamma_{\mathbf{u}}}\right) \#_t \left(\frac{1}{2\gamma_{\mathbf{v}}}\right) = \frac{1}{2\gamma_{\mathbf{u}}^{1-t}\gamma_{\mathbf{v}}^t} A\#_t B \\ &= \frac{1}{2\gamma_{\mathbf{u}}^{1-t}\gamma_{\mathbf{v}}^t} [L_{1-t}(\mu)A + L_t(\mu)B] = L_{1-t}(\mu) \left(\frac{\gamma_{\mathbf{u}}}{\gamma_{\mathbf{v}}}\right)^t \rho_{\mathbf{u}} + L_t(\mu) \left(\frac{\gamma_{\mathbf{v}}}{\gamma_{\mathbf{u}}}\right)^{1-t} \rho_{\mathbf{v}}. \end{aligned}$$

Here, μ is an eigenvalue of AB^{-1} . Note that $AB^{-1} = \frac{\gamma_{\mathbf{u}}}{\gamma_{\mathbf{v}}}\rho_{\mathbf{u}}\rho_{\mathbf{v}}^{-1} = 4\gamma_{\mathbf{u}}\gamma_{\mathbf{v}}\rho_{\mathbf{u}}\rho_{-\mathbf{v}}$ by (13), and $\rho_{\mathbf{u}}\rho_{-\mathbf{v}}$ is similar to

$$\rho_{\mathbf{u}}^{\frac{1}{2}}\rho_{-\mathbf{v}}\rho_{\mathbf{u}}^{\frac{1}{2}} = \text{tr}(\rho_{\mathbf{u}}\rho_{-\mathbf{v}})\rho_{\mathbf{u} \oplus_E (-\mathbf{v})} = \frac{1 - \mathbf{u}^T\mathbf{v}}{2}\rho_{\mathbf{u} \oplus_E (-\mathbf{v})}.$$

The first equality follows from the gyrogroup isomorphism in [12, Theorem 3.4]. Thus, eigenvalues of AB^{-1} are $\gamma_{\mathbf{u}}\gamma_{\mathbf{v}}(1 - \mathbf{u}^T\mathbf{v})(1 \pm \|\mathbf{u} \oplus_E (-\mathbf{v})\|)$. \square

REMARK 9. There is an important and useful connection between the Lorentz factor of the Einstein vector sum and the Lorentz factors of the summands:

$$\gamma_{\mathbf{u} \oplus_E \mathbf{v}} = \gamma_{\mathbf{u}}\gamma_{\mathbf{v}}(1 + \mathbf{u}^T\mathbf{v}). \tag{14}$$

Since $\gamma_{\mathbf{v}} = \gamma_{-\mathbf{v}}$, the value μ in Proposition 5 can be rewritten as

$$\mu = \gamma_{\mathbf{u} \oplus_E (-\mathbf{v})}(1 \pm \|\mathbf{u} \oplus_E (-\mathbf{v})\|) = \left[\frac{1 \pm \|\mathbf{u} \oplus_E (-\mathbf{v})\|}{1 \mp \|\mathbf{u} \oplus_E (-\mathbf{v})\|} \right]^{1/2}.$$

For the Einstein gyrogroup (\mathbf{B}, \oplus_E) , A. Ungar considers what we call the Ungar gyro-metric defined by

$$\rho(\mathbf{u}, \mathbf{v}) = \|(-\mathbf{u}) \oplus_E \mathbf{v}\| = \|\mathbf{u} \oplus_E (-\mathbf{v})\|.$$

He also defines what we call the rapidity metric by $d_E(\mathbf{u}, \mathbf{v}) = \tanh^{-1}\rho(\mathbf{u}, \mathbf{v})$. It is known as the Cayley-Klein metric on the Beltrami-Klein model of hyperbolic geometry (see [9]), or the Bergman metric on the symmetric structure \mathbf{B} with symmetries $S_{\mathbf{w}}(\mathbf{v}) = \mathbf{u} \oplus (-\mathbf{v})$ for some $\mathbf{u} = 2 \otimes \mathbf{w}$ (see [32]). Since $\tanh^{-1}t = \frac{1}{2} \ln \frac{1+t}{1-t}$ for $|t| < 1$, the value μ in Proposition 5 can be rewritten as

$$\mu = e^{\pm d_E(\mathbf{u}, \mathbf{v})}.$$

Applying Proposition 4, (13) and (14) to $A = 2\gamma_u\rho_u$ and $B = 2\gamma_v\rho_v$, we obtain the explicit formula of the weighted spectral geometric mean for 2×2 invertible density matrices.

PROPOSITION 6. For any $\mathbf{u}, \mathbf{v} \in \mathbf{B}$ and $t \in \mathbb{R}$

$$\rho_u \#_t \rho_v = \left[\frac{2\gamma_u}{\gamma_v(1 + \gamma_{\mathbf{u} \oplus_E \mathbf{v}})} \right]^t (\gamma_u \rho_{-\mathbf{u}} + \gamma_v \rho_v)^t \rho_u (\gamma_u \rho_{-\mathbf{u}} + \gamma_v \rho_v)^t.$$

PROPOSITION 7. For any $\mathbf{u}, \mathbf{v} \in \mathbf{B}$,

$$\frac{1 + \|\mathbf{m}\|}{1 - \|\mathbf{m}\|} \leq \left[\frac{(1 + \|\mathbf{u}\|)(1 + \|\mathbf{v}\|)}{(1 - \|\mathbf{u}\|)(1 - \|\mathbf{v}\|)} \right]^{1/2},$$

where $\mathbf{m} = \frac{\gamma_u \mathbf{u} + \gamma_v \mathbf{v}}{\gamma_u + \gamma_v}$.

Proof. Let $A = 2\gamma_u\rho_u$ and $B = 2\gamma_v\rho_v$. Then $A, B \in \mathbb{P}_2$ and their determinants are 1 by Remark 8. Moreover,

$$\|A\| = 2\gamma_u \cdot \frac{1 + \|\mathbf{u}\|}{2} = \gamma_u(1 + \|\mathbf{u}\|) = \sqrt{\frac{1 + \|\mathbf{u}\|}{1 - \|\mathbf{u}\|}},$$

where $\|\cdot\|$ is the operator norm. Similarly $\|B\| = \gamma_v(1 + \|\mathbf{v}\|)$. Applying Corollary 1 to A and B yields

$$\|\gamma_u \rho_u + \gamma_v \rho_v\|^2 \leq \det(\gamma_u \rho_u + \gamma_v \rho_v) \gamma_u(1 + \|\mathbf{u}\|) \gamma_v(1 + \|\mathbf{v}\|).$$

Since $\gamma_u \rho_u + \gamma_v \rho_v = (\gamma_u + \gamma_v)\rho_m$, it reduces to

$$\|\rho_m\|^2 \leq \det(\rho_m) \gamma_u(1 + \|\mathbf{u}\|) \gamma_v(1 + \|\mathbf{v}\|).$$

Since $\|\rho_m\| = \frac{1 + \|\mathbf{m}\|}{2}$ and $\det \rho_m = \frac{1 - \|\mathbf{m}\|^2}{4}$, we obtain the desired inequality from a simplification of the above inequality. \square

REMARK 10. In Theorem 4.2, [15], it has been shown that

$$\frac{\rho_u \# \rho_v}{\text{tr}(\rho_u \# \rho_v)} = \frac{\gamma_u \rho_u + \gamma_v \rho_v}{\gamma_u + \gamma_v} = \rho_m$$

for any $\mathbf{u}, \mathbf{v} \in \mathbf{B}$, where

$$\mathbf{m} = \frac{\gamma_u \mathbf{u} + \gamma_v \mathbf{v}}{\gamma_u + \gamma_v} = \frac{1}{2} \otimes (\mathbf{u} \boxplus_E \mathbf{v}) = L\left(\frac{1}{2}; \mathbf{u}, \mathbf{v}\right)$$

is a gyromidpoint on the Einstein gyrovector space $(\mathbf{B}, \oplus_E, \otimes)$. See Section 6.22 in [29] for more information about the gyromidpoint and gyrocentroid. So Theorem 9 and

Proposition 5 give us generalizations of the result in Theorem 4.2, [15]. Moreover, the inequality in Proposition 7 can be rewritten as

$$2d_E(\mathbf{0}, \mathbf{m}) \leq d_E(\mathbf{0}, \mathbf{u}) + d_E(\mathbf{0}, \mathbf{v}),$$

where $d_E(\mathbf{u}, \mathbf{v}) = \tanh^{-1} \|(-\mathbf{u}) \oplus_E \mathbf{v}\|$ is the rapidity metric on the Einstein gyrovector space $(\mathbf{B}, \oplus_E, \otimes)$: see Remark 9.

6. Final remarks

In this section we investigate the semi-metric d given by

$$d(A, B) = 2 \|\log(A^{-1} \# B)\|_2 \tag{15}$$

on the open convex cone \mathbb{P}_n of positive definite matrices. Even though the semi-metric d in (15) is defined by the Frobenius norm unlike Lemma 2 with the operator norm, all properties in Proposition 3 and Theorem 3 are satisfied for d . The interesting consequence in the geometric point of view is that the geometric mean $A \# B$ of $A, B \in \mathbb{P}_n$ is the unique midpoint with respect to the Riemannian trace metric $\delta(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_2$, meanwhile the spectral geometric mean $A \sharp B$ is the midpoint with respect to the semi-metric d .

REMARK 11. The Riemannian trace metric δ on the cone \mathbb{P} can be rewritten as

$$\delta(A, B) = \|\log(\ominus A \oplus B)\|_2.$$

Meanwhile, the semimetric in Lemma 2 can be rewritten as

$$d(A, B) = \|\log(\boxplus A \boxminus B)\|_2,$$

where $\boxplus A = \ominus A = (-1) \circ A$. Note that the Riemannian metric δ and the semimetric d are related with the gyrogroup operation and cooperation on \mathbb{P} , respectively.

We also compare with the semi-metric d and the Riemannian trace metric δ on the open convex cone \mathbb{P}_n . First, we review the concepts of majorization. For Hermitian matrices H , let $\lambda(H) = (\lambda_1(H), \dots, \lambda_n(H))$ be the vector of all eigenvalues of H , and let $|\lambda(H)|$ be the vector consisting of absolute values of eigenvalues of H . We can rearrange the entries $\lambda_1(H), \dots, \lambda_n(H)$ of the vector $\lambda(H)$ in decreasing order:

$$\lambda_1^\downarrow(H) \geq \lambda_2^\downarrow(H) \geq \dots \geq \lambda_n^\downarrow(H).$$

For Hermitian matrices H and K , we say that $\lambda(K)$ weakly majorizes $\lambda(H)$, written as $\lambda(H) \prec_w \lambda(K)$, if and only if

$$\sum_{i=1}^k \lambda_i^\downarrow(H) \leq \sum_{i=1}^k \lambda_i^\downarrow(K), \quad k = 1, \dots, n.$$

If the equality holds for $k = n$, in addition, then we say that $\lambda(K)$ majorizes $\lambda(H)$, written as $\lambda(H) \prec \lambda(K)$. For positive definite matrices A and B , we say that $\lambda(B)$ weakly log-majorizes $\lambda(A)$, written as $\lambda(A) \prec_{w \log} \lambda(B)$, if and only if

$$\prod_{i=1}^k \lambda_i^\downarrow(A) \leq \prod_{i=1}^k \lambda_i^\downarrow(B), \quad k = 1, \dots, n.$$

If the equality holds for $k = n$, in addition, then we say that $\lambda(B)$ log-majorizes $\lambda(A)$, written as $\lambda(A) \prec_{\log} \lambda(B)$. So $\lambda(A) \prec_{w \log} \lambda(B)$ (or $\lambda(A) \prec_{\log} \lambda(B)$) for $A, B \in \mathbb{P}_n$ implies that $\lambda(\log A) \prec_w \lambda(\log B)$ (respectively, or $\lambda(\log A) \prec \lambda(\log B)$).

THEOREM 10. For any $A, B \in \mathbb{P}_n$,

$$d(A, B) \leq \delta(A, B).$$

The equality holds for commuting matrices $A, B \in \mathbb{P}_n$.

Proof. It has been shown in [23] that

$$\lambda(A \#_t B) \prec_{\log} \lambda(A^{1-t} B^t)$$

for any $A, B \in \mathbb{P}_n$ and $t \in [0, 1]$. So

$$\lambda(\log(A^{-1} \# B)) \prec \lambda(\log(A^{-1/2} B^{1/2})) = \lambda(\log(A^{-1/4} B^{1/2} A^{-1/4})),$$

since $A^{-1/2} B^{1/2}$ is similar to $A^{-1/4} B^{1/2} A^{-1/4}$. By the fact that the singular values of Hermitian matrices are the absolute values of their eigenvalues and by Corollary 10.1 in [31],

$$\begin{aligned} s(\log(A^{-1} \# B)) &= |\lambda(\log(A^{-1} \# B))| \\ &\prec_w |\lambda(\log(A^{-1/4} B^{1/2} A^{-1/4}))| = s(\log(A^{-1/4} B^{1/2} A^{-1/4})), \end{aligned}$$

where $s(H)$ denotes the vector of singular values of a Hermitian matrix H in decreasing order. By the Fan Dominance Theorem in [4, Theorem IV.2.2],

$$||| \log(A^{-1} \# B) ||| \leq ||| \log(A^{-1/4} B^{1/2} A^{-1/4}) |||$$

for any unitarily invariant norm $||| \cdot |||$. Since the Frobenius norm $\| \cdot \|_2$ is unitarily invariant and the Riemannian trace metric is convex by [5, Proposition 6.1.10], we obtain

$$d(A, B) = 2 \| \log(A^{-1} \# B) \|_2 \leq 2 \| \log(A^{-1/4} B^{1/2} A^{-1/4}) \|_2 = 2 \delta(A^{1/2}, B^{1/2}) \leq \delta(A, B).$$

Moreover, any commuting matrices $A, B \in \mathbb{P}_n$ are simultaneously diagonalizable by Theorem 1.3.12 in [10]. So

$$d(A, B) = \| \log A - \log B \|_2 = \delta(A, B). \quad \square$$

According to above remarks and properties of the semi-metric, it is an interesting symmetric divergence. So it would be a good project to find other geometric properties of the semi-metric with the spectral geometric mean such as the metric convexity and the extension to multi-variable spectral geometric means.

Acknowledgement. I deeply thank to anonymous reviewer to improve this paper. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. NRF-2018R1C1B6001394).

REFERENCES

- [1] E. AHN, S. KIM, AND Y. LIM, *An extended Lie-Trotter formula and its applications*, Linear Algebra Appl. **427** (2007), 190–196.
- [2] T. ANDO, *Geometric mean and norm Schwarz inequality*, Ann. Func. Anal. **7** (2016), no. 1, 1–8.
- [3] T. ANDO AND F. HIAI, *Log majorization and complementary Golden-Thompson type inequalities*, Linear Algebra Appl. **197** (1994), 198: 113–131.
- [4] R. BHATIA, *Matrix Analysis*, Springer, 1997.
- [5] R. BHATIA, *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, 2007.
- [6] R. BHATIA AND J. HOLBROOK, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl. **413** (2006), 594–618.
- [7] G. CORACH, H. PORTA, AND L. RECHT, *Convexity of the geodesic distance on spaces of positive operators*, Illinois J. Math. **38** (1994), 87–94.
- [8] M. FIEDLER AND V. PTÁK, *A new positive definite geometric mean of two positive definite matrices*, Linear Algebra Appl. **251** (1997), 1–20.
- [9] Y. FRIEDMAN AND T. SCARR, *Physical Applications of Homogeneous Balls*, Birkhäuser, Basel, 2005.
- [10] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, 2nd edition, Cambridge University Press, 2013.
- [11] H. KARCHER, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math. **30** (1977), no. 5, 509–541.
- [12] S. KIM, *Distances of qubit density matrices on Bloch sphere*, J. Math. Phys. **52** (2011), 102303.
- [13] S. KIM, *Distributivity on the gyrovector spaces*, Kyungpook Math. J. **55** (2015), 13–20.
- [14] S. KIM, *Gyrovector spaces on the open convex cone of positive definite matrices*, Mathematics Interdisciplinary Research **1** (2016), 173–185.
- [15] S. KIM, S. KIM, AND H. LEE, *Factorizations of invertible density matrices*, Linear Algebra Appl. **463** (2014), 190–204.
- [16] S. KIM AND H. LEE, *Relative operator entropy related with the spectral geometric mean*, Anal. Math. Phys. **5**, Vol. 3 (2015), 233–240.
- [17] F. KUBO AND T. ANDO, *Means of positive linear operators*, Math. Ann. **246** (1979/80), no. 3, 205–224.
- [18] J. LAWSON AND Y. LIM, *The geometric mean, matrices, metrics, and more*, Amer. Math. Monthly **108** (2001), 797–812.
- [19] J. LAWSON AND Y. LIM, *Metric convexity of symmetric cones*, Osaka J. Math. **44** (2007), 795–816.
- [20] J. LAWSON AND Y. LIM, *Karcher means and Karcher equations of positive definite operators*, Trans. Amer. Math. Soc. Series B, Vol. 1 (2014), 1–22.
- [21] H. LEE AND Y. LIM, *Metric and spectral geometric means on symmetric cones*, Kyungpook Math. J. **47** (2007), 133–150.
- [22] Y. LIM, *Geometry of midpoint sets for Thompson’s metric*, Linear Algebra Appl. **439** (2013), 211–227.
- [23] J. S. MATHARU AND J. S. AUJLA, *Some inequalities for unitarily invariant norms*, Linear Algebra Appl. **436** (2012), 1623–1631.
- [24] M. MOAKHER, *A differential geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl. **26** (2005), no. 3, 735–747.
- [25] R. D. NUSSBAUM, *Finsler structures for the part metric and Hilbert’s projective metric and applications to ordinary differential equations*, Differential and Integral Equations **7** (1994), 1649–1707.

- [26] W. PUSZ, AND S. L. WORONOWICZ, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys. **8** (1975), 159–170.
- [27] A. C. THOMPSON, *On certain contraction mappings in a partially ordered vector space*, Proc. Amer. Math. Soc. **14** (1963), 438–443.
- [28] M. UCHIYAMA, *Criteria for monotonicity of operator means*, J. Math. Soc. Japan, **53** (2003), 197–207.
- [29] A. A. UNGAR, *Analytic hyperbolic geometry and Albert Einstein's special theory of relativity*, World Scientific Press, 2008.
- [30] T. YAMAZAKI, *The Riemannian mean and matrix inequalities related to the Ando-Hiai inequality and chaotic order*, Oper. Matrices **6** (2012), 577–588.
- [31] F. ZHANG, *Matrix Theory: Basic Results and Techniques*, 2nd edition, Springer, 2011.
- [32] K. ZHU, *Spaces of Holomorphic Functions in the Unit Ball*, Graduate Texts in Mathematics, Springer, New York, 2005.

(Received October 12, 2020)

Sejong Kim
Department of Mathematics
Chungbuk National University
Cheongju 28644, Korea
e-mail: skim@chungbuk.ac.kr