GEOMETRIC MEANS OF POSITIVE OPERATORS II

SAICHI IZUMINO * AND NOBORU NAKAMURA **

Received July 7, 2008; revised August 2, 2008

Dedicated to the memory of Emeritus Professor Masahiro Nakamura

ABSTRACT. Borrowing a technique due to Ando-Li-Mathias, we define a geometric mean of (k + 1) (positive invertible) operators from that of k (or a k-tuple of) operators with a parameter $\lambda \in (0, 1]$. If $\lambda = 1$, then the corresponding geometric mean $G_{\lambda}(=G_1)$ of (k + 1) operators is one defined by Ando-Li-Mathias, and if $\lambda = 2/3$, then G_{λ} is one given by one of the authors in the preceding paper. We also show that a formula due to Yamazaki of the geometric mean for a 3-tuple of 2×2 matrices satisfying a trace condition does not depend on any choice of a parameter in construction.

1. INTRODUCTION

Borrowing a technique presented in Ando-Li-Mathias [2], based on the Riccati equation, one of the authors defined new geometric means of more than two (positive invertible) operators on a Hilbert space [9]. In succession of the paper we shall define such geometric means with parameters.

Let A and B be operators on a Hilbert space H. Then the weighted geometric (or α -power) mean $A \sharp_{\alpha} B$ for $\alpha \in (0, 1]$ is defined [6] by

$$A\sharp_{\alpha}B = A^{\frac{1}{2}} (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^{\alpha} A^{\frac{1}{2}}.$$

The usual geometric mean $A \sharp B$ is given as the case $\alpha = 1/2$. For three operators A, B, Cand for $\lambda \in (0, 1]$, let us define three sequences $\{A_{\lambda,n}\}, \{B_{\lambda,n}\}$ and $\{C_{\lambda,n}\}$, by $A_{\lambda,1} = A$, $B_{\lambda,1} = B$, $C_{\lambda,1} = C$,

(1.1)
$$\begin{aligned} A_{\lambda,n+1} &= A_{\lambda,n} \sharp_{\lambda} (B_{\lambda,n} \sharp C_{\lambda,n}), \\ B_{\lambda,n+1} &= B_{\lambda,n} \sharp_{\lambda} (C_{\lambda,n} \sharp A_{\lambda,n}) \quad \text{and} \\ C_{\lambda,n+1} &= C_{\lambda,n} \sharp_{\lambda} (A_{\lambda,n} \sharp B_{\lambda,n}) \quad \text{for } n \ge 1. \end{aligned}$$

Then we shall obtain a common limit of them, which we define a geometric mean $G_{\lambda} = G_{\lambda}(A, B, C)$ with a parameter λ . If $\lambda = 1$, then the geometric mean $G_{\lambda}(=G_1)$ is one defined in [2], and if $\lambda = 2/3$, then G_{λ} is one defined in [9].

The above technical device using sequential limits due to Ando-Li-Mathias causes the geometric mean G_{λ} to satisfy a property, *permutation invariance* (P3, below) for three operators. In [2], Ando-Li-Mathias stated the following ten postulates for a geometric mean $G(A_1, ..., A_k)$ of k (or a k-tuple of) operators $A_1, ..., A_k$ to be a reasonable one, (the usual geometric mean $G(A_1, A_2) = A_1 \sharp A_2$ is reasonable):

P1 Consistency with scalars. If A_1, A_2, \ldots, A_k commute then

 $G(A_1, A_2, \ldots, A_k) = (A_1 A_2 \ldots A_k)^{\frac{1}{k}}.$

²⁰⁰⁰ Mathematics Subject Classification. 47A63, 47A64.

Key words and phrases. geometric mean of n-operators, arithmetic-geometric mean inequality.

P1' This implies $G(A, \ldots, A) = A$.

P2 Joint homogeneity. $G(a_1A_1, a_2A_2, ..., a_kA_k) = (a_1a_2 \cdots a_k)^{\frac{1}{k}} G(A_1, A_2, ..., A_k)$ for $a_i \ge 0$ with i = 1, ..., k.

P2' This implies $G(aA_1, aA_2, \dots, aA_k) = aG(A_1, A_2, \dots, A_k)$ $(a \ge 0)$.

P3 Permutation invariance. For any permutation $\pi(A_1, A_2, \ldots, A_k)$ of (A_1, A_2, \ldots, A_k) , $G(A_1, A_2, \ldots, A_k) = G(\pi(A_1, A_2, \ldots, A_k))$.

P4 Monotonicity. The map $(A_1, A_2, \ldots, A_n) \mapsto G(A_1, A_2, \ldots, A_n)$ is monotone, i.e., if $A_i \geq B_i$ for $i = 1, \ldots, k$, then $G(A_1, A_2, \ldots, A_k) \geq G(B_1, B_2, \ldots, B_k)$.

P5 Continuity from above. If $\{A_1^{(n)}\}, \{A_2^{(n)}\}, \ldots, \{A_k^{(n)}\}$ are monotonic decreasing sequences converging to A_1, A_2, \ldots, A_k , respectively, then $\{G(A_1^{(n)}, A_2^{(n)}, \ldots, A_k^{(n)})\}$ converges to $G(A_1, A_2, \ldots, A_k)$.

P6 Congruence invariance. For any invertible S,

$$G(S^*A_1S, S^*A_2S, \dots, S^*A_kS) = S^*G(A_1, A_2, \dots, A_k)S$$

P7 Joint concavity. The map $(A_1, A_2, \ldots, A_k) \mapsto G(A_1, A_2, \ldots, A_k)$ is jointly concave:

$$G(\lambda A_1 + (1-\lambda)A'_1, \lambda A_2 + (1-\lambda)A'_2, \dots, \lambda A_k + (1-\lambda)A'_k)$$

$$\geq \lambda G(A_1, A_2, \dots, A_k) + (1 - \lambda) G(A'_1, A'_2, \dots, A'_k) \quad (0 < \lambda < 1).$$

P8 Self-duality. $G(A_1, A_2, \ldots, A_k)^* = G(A_1, A_2, \ldots, A_k)$. The dual $G(A_1, A_2, \ldots, A_k)^*$ is defined by

$$G(A_1, A_2, \dots, A_k)^* = G(A_1^{-1}, A_2^{-1}, \dots, A_k^{-1})^{-1}$$

P9 (In case A_1, A_2, \ldots, A_k are matrices.) Determinant identity.

$$\det G(A_1, A_2, \dots, A_k) = (\det A_1 \cdot \det A_2 \cdot \dots \cdot \det A_k)^{\frac{1}{k}}$$

P10 The arithmetic-geometric-harmonic mean inequality.

$$\frac{A_1 + A_2 + \dots + A_k}{k} \ge G(A_1, A_2, \dots, A_k) \ge \left(\frac{A_1^{-1} + A_2^{-1} + \dots + A_k^{-1}}{k}\right)^{-1}.$$

In this note, we define a geometric mean of (k+1) operators with a parameter λ which still satisfies the above properties P1-P10 from a given geometric mean of k operators satisfying all properties. Based on a method in [2], Yamazaki [11] obtained a formula of the geometric mean of 2×2 matrices under a trace condition. We shall show that the formula does not depend on any special choice of a parameter in the process of construction.

Without occurrence of ambiguity, we shall often abbreviate the letter λ . All operators (or matrices) are assumed to be *positive invertible* (or *positive definite*) if stated otherwise.

2. Definition of geometric means of more than two operators

Let Ω be the set of all (positive invertible) operators on H. Then the Thompson metric on Ω is defined ([10], [3], [4]) by

 $d(A,B) = \max\{\log M(A/B), \log M(B/A)\} \text{ for } A, B \in \Omega,$

where

$$M(A/B) = \inf\{\mu > 0 : A \le \mu B\} \ (= \| B^{-1/2} A B^{-1/2} \|).$$

We remark that Ω is complete with respect to the Thompson metric topology. As a basic inequality with respect to the metric, the following inequality for a weighted geometric mean of two operators holds [3], [4]:

(2.1)
$$\begin{aligned} d(A_1 \sharp_{\alpha} A_2, B_1 \sharp_{\alpha} B_2) &\leq (1 - \alpha) d(A_1, B_1) + \alpha d(A_2, B_2) \\ \text{for } A_1, A_2, B_1, B_2 \in \Omega \text{ and } \alpha \in (0, 1). \end{aligned}$$

Now in order to define our geometric mean $G_{\lambda}(A_1, ..., A_{k+1})$ of (k+1) operators from a given one of $k (\geq 2)$ operators, we want to assume a useful inequality:

(2.2)
$$d(G(A_1, ..., A_k), G(B_1, ..., B_k)) \le \frac{1}{k} \sum_{i=1}^k d(A_i, B_i)$$

for another k-tuple of operators $B_1, ..., B_k$.

Theorem 2.1. The geometric mean $G_{\lambda}(A_1, ..., A_{k+1})$ is always defined as the common limit of the following (k+1) sequences $\{A_1^{(r)}\}, ..., \{A_{k+1}^{(r)}\}$ of (k+1) operators $A_1, ..., A_{k+1}$:

$$(2.3) \qquad \begin{array}{l} A_i^{(1)} = A_i \ for \ i = 1, ..., k+1, \quad and \\ A_i^{(r+1)} = A_i^{(r)} \sharp_{\lambda} G((A_j^{(r)})_{j \neq i}) (= A_i^{(r)} \sharp_{\lambda} G(A_1^{(r)}, ..., A_{i-1}^{(r)}, A_{i+1}^{(r)}, ..., A_{k+1}^{(r)})) \\ for \ r \ge 1, \ i = 1, ..., k+1. \end{array}$$

where $\lambda \in (0,1]$ and $G(A_1, ..., A_k)$ is a geometric mean of k operators satisfying P1-P10 and the inequality (2.2). The geometric mean $G_{\lambda}(A_1, ..., A_{k+1})$ satisfies P1-P10, and furthermore, the following inequality holds:

(2.4)
$$d(G_{\lambda}(A_1, ..., A_{k+1}), G_{\lambda}(B_1, ..., B_{k+1})) \le \frac{1}{k+1} \sum_{i=1}^{k+1} d(A_i, B_i)$$

corresponding to (2.2) for another (k + 1)-tuple $B_1, ..., B_{k+1}$ of operators.

Proof. To see that all sequences $\{A_i^{(r)}\}\$ are convergent with a common limit we first show that for $i, j = 1, ..., k + 1, i \neq j$

(2.5)
$$d(A_i^{(r+1)}, A_j^{(r+1)}) \le \left(1 - \frac{k-1}{k}\lambda\right)^r d(A_i, A_j)$$

By the definition (2.3) of $A_i^{(r)}$ and the inequalities (2.1) and (2.4), we have

$$d(A_i^{(r+1)}, A_j^{(r+1)}) = d(A_i^{(r)} \sharp_{\lambda} G((A_{\ell}^{(r)})_{\ell \neq i}), A_j^{(r)} \sharp_{\lambda} G((A_{\ell}^{(r)})_{\ell \neq j}))$$

$$\leq (1 - \lambda) d(A_i^{(r)}, A_j^{(r)}) + \lambda d(G((A_{\ell}^{(r)})_{\ell \neq i}), G((A_{\ell}^{(r)})_{\ell \neq j}))$$

$$\leq (1 - \lambda) d(A_i^{(r)}, A_j^{(r)}) + \lambda \cdot \frac{1}{k} d(A_i^{(r)}, A_j^{(r)})$$

$$= \left(1 - \frac{k - 1}{k}\lambda\right) d(A_i^{(r)}, A_j^{(r)}).$$

Hence by iteration with respect to r we can obtain the desired inequality. Next we show

(2.6)
$$d(A_i^{(r+1)}, A_i^{(r)}) \le \frac{\lambda}{k} \left(1 - \frac{k-1}{k}\lambda\right)^{r-1} K_i,$$

where
$$K_i = \sum_{\ell=1, \ell \neq i}^{k+1} d(A_i, A_\ell)$$
. Note that

$$A_{i}^{(r)} = A_{i}^{(r)} \sharp_{\lambda} G(A_{i}^{(r)}, ..., A_{i}^{(r)}).$$

Using (2.2), we have

$$d(A_i^{(r+1)}, A_i^{(r)}) \leq \lambda d(G((A_\ell^{(r)})_{\ell \neq i}), G(A_i^{(r)}, \dots, A_i^{(r)})) \leq \lambda \cdot \frac{1}{k} \sum_{\ell=1, \ell \neq i}^{k+1} d(A_i^{(r)}, A_\ell^{(r)}).$$

Hence from (2.5)

$$d(A_i^{(r+1)}, A_i^{(r)}) \le \frac{\lambda}{k} \cdot \sum_{\ell=1, \ell \neq i}^{k+1} \left(1 - \frac{k-1}{k}\lambda\right)^{r-1} d(A_\ell, A_i) = \frac{\lambda}{k} \left(1 - \frac{k-1}{k}\lambda\right)^{r-1} K_i,$$

which is the desired inequality. Now we see that for any i, the sequence $\{A_i^{(r)}\}$ is convergent, or a Cauchy sequence. In fact, for $r \leq s$

$$d(A_{i}^{(r+1)}, A_{i}^{(s+1)}) \leq \sum_{\ell=r+1}^{s} d(A_{i}^{(\ell)}, A_{i}^{(\ell+1)}) \leq \frac{\lambda}{k} K_{i} \sum_{\ell=r+1}^{s} \left(1 - \frac{k-1}{k}\lambda\right)^{\ell-1} \\ \leq \frac{\lambda}{k} K_{i} \cdot \left(1 - \frac{k-1}{k}\lambda\right)^{r} / \left(\frac{k-1}{k}\lambda\right) = \frac{K_{i}}{k-1} \left(1 - \frac{k-1}{k}\lambda\right)^{r}$$

Hence $d(A_i^{(r+1)}, A_i^{(s+1)}) \to 0$ as $r(< s) \to \infty$, so that $\{A_i^{(r)}\}$ is convergent. From (2.5), we easily see that all $\{A_i^{(r)}\}$ have the same limit, which guarantees the desired geometric mean to be defined.

It is not difficult to see that the geometric mean $G_{\lambda}(A_1, ..., A_{k+1})$ satisfies all properties P1-P10. For example, to see P3, let $\pi(A_1, A_2, ..., A_{k+1}) = (A_{\pi(1)}, ..., A_{\pi(k+1)})$ be a permutation of $(A_1, A_2, ..., A_{k+1})$, and let

$$B_i^{(1)} = A_{\pi(i)}^{(1)} = A_{\pi(i)}, \quad B_i^{(r+1)} = B_i^{(r)} \sharp_\lambda G((B_j)_{j \neq i}^{(r)})$$

for $i = 1, ..., k+1, \ r \ge 1$

Then we see that $B_i^{(r)} = A_{\pi(i)}^{(r)}$. In fact, assuming that $B_i^{(r)} = A_{\pi(i)}^{(r)}$ (i = 1, ..., k + 1), we have

$$B_i^{(r+1)} = A_{\pi(i)}^{(r)} \sharp_\lambda G((A_{\pi(j)})_{j \neq i}) = A_{\pi(i)}^{(r+1)}.$$

Hence $\{B_i^{(r)}\}\$ and $\{A_{\pi(i)}^{(r)}\}\$ coincide, so that they converge to the same limit, which is desired.

For the inequality (2.4), let the sequences $\{B_1^{(r)}\}, ..., \{B_{k+1}^{(r)}\}$ be defined corresponding to $B_1, ..., B_{k+1}$, similarly as (2.3) for $A_1, ..., A_{k+1}$. Then for each *i*, from (2.1) and the

assumption (2.2), we have

$$\begin{split} d(A_i^{(r+1)}, B_i^{(r+1)}) &= d(A_i^{(r)} \sharp_\lambda G((A_j^{(r)})_{j \neq i}), B_i^{(r)} \sharp_\lambda G((B_j^{(r)})_{j \neq i})) \\ &\leq (1 - \lambda) d(A_i^{(r)}, B_i^{(r)}) + \lambda d(G((A_j^{(r)})_{j \neq i}), G((B_j^{(r)})_{j \neq i})) \\ &\leq (1 - \lambda) d(A_i^{(r)}, B_i^{(r)}) + \lambda \cdot \frac{1}{k} \sum_{j=1, j \neq i}^{k+1} d(A_j^{(r)}, B_j^{(r)}) \\ &= \left(1 - \frac{k+1}{k}\lambda\right) d(A_i^{(r)}, B_i^{(r)}) + \frac{\lambda}{k} \sum_{j=1}^{k+1} d(A_j^{(r)}, B_j^{(r)}). \end{split}$$

Summing up all $d(A_i^{(r+1)}, B_i^{(r+1)})$ with respect to i, we have

$$\sigma_{r+1} := \sum_{i=1}^{k+1} d(A_i^{(r+1)}, B_i^{(r+1)})$$

$$\leq \left(1 - \frac{k+1}{k}\lambda\right) \sum_{i=1}^{k+1} d(A_i^{(r)}, B_i^{(r)}) + \frac{k+1}{k}\lambda \sum_{j=1}^{k+1} d(A_j^{(r)}, B_j^{(r)})$$

$$= \sum_{i=1}^{k+1} d(A_i^{(r)}, B_i^{(r)}) \ (= \sigma_r).$$

Hence $\sigma_{r+1} \leq \sigma_r \leq \cdots \leq \sigma_1$, that is, $\sigma_{r+1} \leq \sum_{i=1}^{k+1} d(A_i, B_i)$. Taking the limit as $r \to \infty$, we have the desired inequality since $\sigma_{r+1} \to (k+1)d(G_\lambda(A_1, ..., A_{k+1}), G_\lambda(B_1, ..., B_{k+1}))$.

Example 2.2 Let

$$A_{1} = \begin{bmatrix} 10 & 1 \\ 1 & 0.2 \end{bmatrix}, \quad A_{2} = \begin{bmatrix} 4.1 & 4.9 \\ 4.9 & 6.1 \end{bmatrix} \text{ and } A_{3} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then by numerical computation we have, (discarded less than 10^{-10} ,)

$$G_{1/3} = \begin{bmatrix} 1.6472832734 & 0.6138234917 \\ 0.6138234917 & 0.8357878097 \end{bmatrix} (= A_1^{(r)} = A_2^{(r)} = A_3^{(r)} \text{ for } r \ge 39),$$

$$G_{1/2} = \begin{bmatrix} 1.6499095763 & 0.6157374707 \\ 0.6157374707 & 0.8358837675 \end{bmatrix} (= A_1^{(r)} = A_2^{(r)} = A_3^{(r)} \text{ for } r \ge 20),$$

$$G_{2/3} = \begin{bmatrix} 1.6600838645 & 0.6231334993 \\ 0.6231334993 & 0.8362802552 \end{bmatrix} (= A_1^{(r)} = A_2^{(r)} = A_3^{(r)} \text{ for } r \ge 4)$$

and

$$G_1 = \begin{bmatrix} 1.6970826618 & 0.6497880663\\ 0.6497880663 & 0.8380408114 \end{bmatrix} (= A_1^{(r)} = A_2^{(r)} = A_3^{(r)} \text{ for } r \ge 39).$$

Now for more convenient expression, denote by $(G, \lambda) = (G, \lambda)(A_1, ..., A_{k+1})$ the geometric mean constructed as in Theorem 2.1. Then successively we can define

$$(G, \lambda_1, ..., \lambda_\ell) = ((G, \lambda_1, ..., \lambda_{\ell-1}), \lambda_\ell).$$

Let $G = \sharp(A_1, A_2) = A_1 \sharp A_2$. Then $(\sharp, 1, ..., 1)$ is the geometric mean (of k operators) given by Ando-Li-Mathias in [2], and $(\sharp; \frac{2}{3}, ..., \frac{k-1}{k})$ is one given in [9].

Example 2.3. Let

$$A_{1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad A_{2} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad A_{3} = \begin{bmatrix} 3 & \sqrt{2} \\ \sqrt{2} & 1 \end{bmatrix} \text{ and } A_{4} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then by numerical computation, we obtain, (discarded less than 10⁻¹⁰,) for $r \ge 4$,
 $(\sharp; \frac{2}{3}, \frac{3}{4})(A_{1}, A_{2}, A_{3}, A_{4}) = \begin{bmatrix} 1.4126934750 & 0.7066270669 \\ 0.7066270669 & 1.0331915013 \end{bmatrix} (= A_{1}^{(r)} = A_{2}^{(r)} = A_{3}^{(r)} = A_{4}^{(r)}).$

3. YAMAZAKI'S FORMURA FOR THE GEOMETRIC MEAN OF THREE MATRICES

Let A_1, A_2 and A_3 be 2×2 (positive definite) matrices. Write

$$G_3 = (\sharp, 1)(A_1, A_2, A_3) \ (G_2 = A_1 \sharp A_2).$$

Related to G_3 , refining a result in [2], Yamazaki [11] presented the following formula : If det $A_i = 1$ for i = 1, 2, 3, then

$$G_3 = \frac{A_1 + A_2 + A_3}{\sqrt{\det(A_1 + A_2 + A_3)}}$$

under the trace condition:

(3.1)
$$\tau(A_i^{-1}A_j) = c \text{ (a constant) for } i, j = 1, 2, 3, i \neq j.$$

Here $\tau(A)$ is the trace of A. (In [11], Yamazaki further presented the similar formula as above for more than three matrices.)

With respect to the above result, we show the following fact, which implies that every parametrized geometric mean $\tilde{G}_3 = (\sharp, \lambda)(A_1, A_2, A_3)$ coincides, that is, \tilde{G}_3 does not depend on λ under the condition (3.1):

Theorem 3.1. Let A_1, A_2, A_3 be 2×2 matrices with det $A_i = 1$ (i = 1, 2, 3). Then under the condition (3.1)

(3.2)
$$\tilde{G}_3 = (\sharp, \lambda)(A_1, A_2, A_3) = \frac{A_1 + A_2 + A_3}{\sqrt{\det(A_1 + A_2 + A_3)}}$$
 for any $\lambda \in (0, 1]$.

Before we prove the theorem, we want to provide a lemma:

Lemma 3.2. Let A, B be 2×2 matrices such that det $A = \det B = 1$, and let $\lambda \in (0, 1]$. Then

(3.3)
$$A\sharp_{\lambda}B = \phi_{\lambda}(\tau(A^{-1}B))A + \psi_{\lambda}(\tau(A^{-1}B))B.$$

where

$$\phi_{\lambda}(t) = \frac{1}{\sqrt{t^2 - 4}} \left\{ \left(\frac{t + \sqrt{t^2 - 4}}{2} \right)^{1 - \lambda} - \left(\frac{t - \sqrt{t^2 - 4}}{2} \right)^{1 - \lambda} \right\},\$$

$$\psi_{\lambda}(t) = \frac{1}{\sqrt{t^2 - 4}} \left\{ \left(\frac{t + \sqrt{t^2 - 4}}{2} \right)^{\lambda} - \left(\frac{t - \sqrt{t^2 - 4}}{2} \right)^{\lambda} \right\} \quad for \ t > 2$$

and

$$\phi_{\lambda}(2) = \lim_{t \to 2} \phi_{\lambda}(t) = 1 - \lambda, \ \psi_{\lambda}(2) = \lim_{t \to 2} \psi_{\lambda}(t) = \lambda.$$

In particular, (for $\lambda = 1/2$) [2, Proposition 2.1]

(3.4)
$$A \sharp B = \frac{A+B}{\sqrt{\tau(A^{-1}B)+2}} \left(= \frac{A+B}{\sqrt{\det(A+B)}} \right)$$

Proof. Let C be a 2×2 matrix such that $C \neq I_2$ and det C = 1. (I_2 is the identity matrix.) Then we can see $\sigma(C) = \{p, p^{-1}\}$ (the spectrum of C) with some p > 1. Hence from [2, p. 310], we have

(3.5)
$$C^{\lambda} = \frac{p^{1-\lambda} - p^{-(1-\lambda)}}{p - p^{-1}} I_2 + \frac{p^{\lambda} - p^{-\lambda}}{p - p^{-1}} C.$$

It follows from $\tau(C) = p + p^{-1}$ that $p^{\pm 1} = \frac{\tau(C) \pm \sqrt{\tau(C)^2 - 4}}{2}$. Hence

(3.6)
$$C^{\lambda} = \phi_{\lambda}(\tau(C))I_2 + \psi_{\lambda}(\tau(C))C.$$

Now for A, B we may only consider the case $A \neq B$. Then note that $A^{-1/2}BA^{-1/2} \neq I_2$, $\det(A^{-1/2}BA^{-1/2}) = \det(A^{-1}B) = 1$, and that $\tau(A^{-1/2}BA^{-1/2}) = \tau(A^{-1}B)$ since $\sigma(A^{-1/2}BA^{-1/2}) = \sigma(A^{-1}B)$. Hence, replacing C by $A^{-1/2}BA^{-1/2}$ in (3.6), we have

$$(A^{-1/2}BA^{-1/2})^{\lambda} = \phi_{\lambda}(\tau(A^{-1}B))I_2 + \psi_{\lambda}(\tau(A^{-1}B))A^{-1/2}BA^{-1/2}$$

Multiplying by $A^{1/2}$ both sides of the above identity from the left and the right, we now obtain the identity(3.3). For $\lambda = 1/2$, we can see that $\phi_{\lambda}(t) = \psi_{\lambda}(t) = \frac{1}{\sqrt{t+2}}$. We can also see that

$$\det(A+B) = \det A^{-1} \det(A+B) = \det(I_2 + A^{-1}B) = \tau(A^{-1}B) + 2$$

This implies the desired identity (3.4).

Proof of Theorem 3.1. What we have to prove is, for all $r \ge 1$, and i = 1, 2, 3

(3.7)
$$A_{i}^{(r)} = \alpha_{r}A_{i} + \beta_{r}\sum_{j=1, j\neq i}^{3} A_{j}$$

holds for some scalars α_r and β_r (each of which does not depend on *i*): In fact, if (3.7) holds, then by the similar argument as in [2, p.329-330], or, by direct computation, we have

$$\lim_{r\to\infty}\alpha_r=\lim_{r\to\infty}\beta_r=\alpha$$

for some α , so that $\tilde{G}_3 = \alpha(A_1 + A_2 + A_3)$, or

$$\alpha = \frac{1}{\sqrt{\det(A_1 + A_2 + A_3)}},$$

which is desired.

Now our task is to show (3.7) (by induction on r). For r = 1, this is clear, since we can put $\alpha_1 = 1$ and $\beta_1 = 0$. Assume that (3.7) holds (for r). Then we have to show that the identity holds for r + 1 instead of r, that is,

(3.8)
$$A_i^{(r+1)} = \alpha_{r+1}A_i + \beta_{r+1}\sum_{j=1, j\neq i}^3 A_j$$

for some scalars α_{r+1} and β_{r+1} . We devide the proof into four steps.

Step 1. $c_r := \tau((A_i^{(r)})^{-1}A_j^{(r)})$ (for $i \neq j$) does not depend on i, j, or, more precisely, (3.9) $c_r = (4+2c)\alpha_r\beta_r + (2+3c)\beta_r^2 + c\alpha_r^2.$ To see this fact, first note that det $A_{\ell} = 1$ (for $\ell = 1, 2, 3$) and det $A_i^{(r)} = 1$ (by P9), so that the inverse operation on 2×2 matrices assures

$$(A_i^{(r)})^{-1} = \left(\alpha_r A_i + \beta_r \sum_{\ell=1, \ell \neq i}^3 A_\ell\right)^{-1} = \alpha_r A_i^{-1} + \beta_r \sum_{\ell=1, \ell \neq i}^3 A_\ell^{-1}.$$

Hence, (for example,) for i = 1, j = 2,

$$(A_1^{(r)})^{-1}A_2^{(r)} = (\alpha_r A_1^{-1} + \beta_r A_2^{-1} + \beta_r A_3^{-1})(\alpha_r A_2 + \beta_r A_1 + \beta_r A_3)$$

= $\alpha_r \beta_r (2I_2 + A_1^{-1}A_3 + A_3^{-1}A_2) + \beta_r^2 (I_2 + A_2^{-1}A_1 + A_3^{-1}A_1 + A_2^{-1}A_3) + \alpha_r^2 A_1^{-1}A_2.$

¿From this and the trace condition, we have

$$\tau((A_1^{(r)})^{-1}A_2^{(r)}) = (4+2c)\alpha_r\beta_r + (2+3c)\beta_r^2 + c\alpha_r^2 = c_r.$$

For other pairs of i, j we can obtain the same value c_r .

Step 2. Let

$$d_r = \frac{2\beta_r}{\sqrt{2+c_r}}$$
 and $e_r = \frac{\beta_r + \alpha_r}{\sqrt{2+c_r}}$.

Then from (3.4), (3.7) and Step 1, we have

$$G((A_j^{(r)})_{j\neq 3}) = A_1^{(r)} \sharp A_2^{(r)} = \frac{1}{\sqrt{\tau((A_1^{(r)})^{-1}A_2^{(r)}) + 2}} (A_1^{(r)} + A_2^{(r)})$$
$$= \frac{1}{\sqrt{c_r + 2}} [2\beta_r A_3 + (\beta_r + \alpha_r)(A_1 + A_2)]$$
$$= d_r A_3 + e_r (A_1 + A_2).$$

Similarly, we have $G((A_j^{(r)})_{j\neq 1}) = d_r A_1 + e_r (A_2 + A_3)$ and $G((A_j^{(r)})_{j\neq 2}) = d_r A_2 + e_r (A_3 + A_1)$.

Step 3. We have to show that $f_r := \tau((A_i^{(r)})^{-1}G((A_j^{(r)})_{j\neq i}))$ is independent from *i*. For i = 1, by (3.7) and Step 2, we have

$$\begin{split} f_r &= \tau \left(\left(\alpha_r A_1 + \beta_r (A_2 + A_3) \right)^{-1} \left(d_r A_1 + e_r (A_2 + A_3) \right) \right) \\ &= \tau \left(\left(\alpha_r A_1^{-1} + \beta_r (A_2^{-1} + A_3^{-1}) \right) \left(d_r A_1 + e_r (A_2 + A_3) \right) \right) \\ &= \tau \left((\alpha_r d_r + 2\beta_r e_r) I + \alpha_r e_r (A_1^{-1} A_2 + A_1^{-1} A_3) \\ &+ \beta_r d_r (A_2^{-1} A_1 + A_3^{-1} A_1) + \beta_r e_r (A_2^{-1} A_3 + A_3^{-1} A_2) \right) \\ &= 2\alpha_r d_r + (4 + 2c) \beta_r e_r + 2c\alpha_r e_r + 2c\beta_r d_r. \end{split}$$

For i = 2 and 3, we can obtain the same value f_r . This is what we want in this step. Now in the final step, we show (3.8):

Step 4. By (3.3), Steps 2 and 3, we have

$$\begin{split} A_i^{(r+1)} &= A_i^{(r)} \sharp_{\lambda} G((A_j^{(r)})_{j \neq i}) \\ &= \phi_{\lambda} \left(\tau((A_i^{(r)})^{-1} G((A_j^{(r)})_{j \neq i})) \right) A_i^{(r)} + \psi_{\lambda} \left(\tau((A_i^{(r)})^{-1} G((A_j^{(r)})_{j \neq i})) \right) G((A_j^{(r)})_{j \neq i}) \\ &= \phi_{\lambda}(f_r) \left(\alpha_r A_i + \beta_r \sum_{j=1, j \neq i}^3 A_j \right) + \psi_{\lambda}(f_r) \left(d_r A_i + e_r \sum_{j=1, j \neq i}^3 A_j \right) \\ &= \{ \alpha_r \phi_{\lambda}(f_r) + d_r \psi_{\lambda}(f_r) \} A_i + \{ \beta_r \phi_{\lambda}(f_r) + e_r \psi_{\lambda}(f_r) \} \sum_{j=1, j \neq i}^3 A_j. \end{split}$$

Putting

$$\alpha_{r+1} = \alpha_r \phi_{\lambda}(f_r) + d_r \psi_{\lambda}(f_r) \text{ and } \beta_{r+1} = \beta_r \phi_{\lambda}(f_r) + e_r \psi_{\lambda}(f_r),$$

in the desired (2.8)

we obtain the desired (3.8).

Remark 3.4. A nontrivial example of a triple of 2×2 matrices with their determinants 1, satisfying the trace condition (3.1) is given as follows: Let c > 2, and let a, b be positive numbers such that a(c-a) > 1, b(c-b) > 1, respectively, and further the identity

(3.10)
$$(c+2)(a^2+b^2+1) - 2c(ab+a+b) = 0$$

holds (In fact, the parameters a = 1, b = 2 and c = 3 satisfy these conditions). In case $(a + b - 1)c - 2ab \ge 0$, we put

$$A_1 = \begin{bmatrix} a & \sqrt{a(c-a)-1} \\ \sqrt{a(c-a)-1} & c-a \end{bmatrix}, \quad A_2 = \begin{bmatrix} b & \sqrt{b(c-b)-1} \\ \sqrt{b(c-b)-1} & c-b \end{bmatrix}$$

and $A_3 = I_2$. Then it is easy to see that det $A_i = 1$ for all *i*. For the trace condition, it

and $A_3 = I_2$. Then it is easy to see that det $A_i = 1$ for all *i*. For the trace condition, it is also easy to see $\tau(A_1^{-1}A_3) = \tau(A_2^{-1}A_3) = c$. So it suffices to show that $\tau(A_1^{-1}A_2) = c$. Since

$$\tau(A_1^{-1}A_2) = (a+b)c - 2ab - 2\sqrt{a(c-a) - 1}\sqrt{b(c-b) - 1}$$

we have to show that

$$(a+b)c - 2ab - 2\sqrt{a(c-a) - 1}\sqrt{b(c-b) - 1} - c \ (=\tau(A_1^{-1}A_2) - c) = 0.$$

or, equivalently,

$$\{(a+b-1)c-2ab\}^2 - \left(2\sqrt{a(c-a)-1}\sqrt{b(c-b)-1}\right)^2 = 0.$$

Write Q the left side of the above identity. Then

$$Q = \{(a+b-1)^2 - 4ab\}c^2 + 4(ab+a+b)c - 4(a^2+b^2+1)$$
$$= (c-2)\{(c+2)(a^2+b^2+1) - 2c(ab+a+b)\} = 0$$

by the assumption (3.10). Hence A_1, A_2 and A_3 are desired matrices.

In case $(a+b-1)c - 2ab \le 0$, we define A'_1 by

$$A'_1 = \left[\begin{array}{cc} a & -\sqrt{a(c-a)-1} \\ -\sqrt{a(c-a)-1} & c-a \end{array}\right].$$

Then A'_1, A_2 and A_3 become a desired triple of matrices.

Remark 3.5. By the similar argument as in Theorem 3.1, we can show the similar fact as (3.2) of the geometric mean for more than three (2×2) matrices, though the trace condition then becomes very restrictive.

SAICHI IZUMINO AND NOBORU NAKAMURA

References

- [1] T. ANDO, Topics on operator inequalities, Hokkaido Univ. Lecture Note, 1978.
- [2] T. ANDO, C.-K. LI and R. MATHIAS, Geometric means, Linear Algebra Appl., 385 (2004), 305-334.
- [3] E. ANDRUCHOW, G. CORACH and D. STOJANOFF, Geometrical significance of the Löwner-Heinz inequality, Proc. Amer. Math. Soc., 128 (1999), 1031-1037.
- G. CORACH, H. PORTA and L. RECHT, Convexity of the geodesic distance on spaces of positive operators, Illinois J. Math., 38 (1994), 87-94.
- [5] J.I. FUJII, M. FUJII, M. NAKAMURA, J. PEČARIĆ and Y. SEO, A reverse inequality for the weighted geometric mean due to Lawson-Lim, Linear Algebra Appl., 427 (2007), 272-284.
- [6] F. KUBO and T. ANDO, Means of positive linear operators, Math. Ann., 246 (1980), 205-224.
- J. LAWSON and Y. LIM, A general framework for extending means to higher orders, preprint. http: //arxiv.org/PS_cache/math/pdf/0612/0612293v1.pdf.
- [8] N. NAKAMURA, Geometric operator mean induced from the Riccati equation, Sci. Math. Japon., 66 (2007), 83-87.
- [9] N. NAKAMURA, Geometric means of positive operators, Kyungpook Math. J., to appear.
- [10] R. D. NUSSBAUM, Hilbert's projective metric and iterated nonlinear maps, Mem. Amer. Math. Soc., 75 (391) (1988).
- [11] T. YAMAZAKI, private communication.

* UNIVERSITY OF TOYAMA, GOFUKU, TOYAMA 930-8555, JAPAN *E-mail address*: s-izumino@@h5.dion.ne.jp

** FUJIKOSHI-KOGYO SENIOR HIGHSCHOOL, HIGASHI-ISHIGANE 7-5, TOYAMA, 930-0964, JAPAN *E-mail address*: nakamura@fujikoshiko.ac.jp