数理解析研究所講究録
第2073巻 2018年 44-54

GENERALIZED KARCHER EQUATION, RELATIVE OPERATOR ENTROPY AND THE ANDO-HIAI INEQUALITY

TAKEAKI YAMAZAKI

ABSTRACT. In this paper, we shall give a concrete relation between generalized Karcher equation and operator means as its solution. Next, we shall show two types of the Ando-Hiai inequalities for the solution of the generalized Karcher equations. In this discussion, we also give properties of relative operator entropy.

1. INTRODUCTION

The theory of operator means was firstly considered in [22]. In that paper, the operator geometric mean has been defined. Then the axiom of operator means of two-operators was defined in [14]. However, this axiom cannot be extended over more than three operators, especially, many people attempted to define operator geometric mean of n-operators with natural properties. For this problem, the first solution was given in [2]. In that paper, a geometric mean of n-positive definite matrices was defined, and it has10 nice properties, for instance, operator monotonicity. Since then operator geometric means has been discussed in many papers, for example, [5, 12, 13]. Especially, we pay attention to a geometric mean of n-positive definite matrices which is defined in [3]. It was defined by using the property that the set of all positive definite matrices is a Riemannian manifold with non-positive curvature. Then it was shown that the geometric mean can be defined by a solution of a matrix equation in [19].

For bounded linear operators on a Hilbert space case, although, we can not define the geometric mean of n-operators by the same way to [3], it can be defined as a solution of the same operator equation to [19] in [16]. This operator equation is called the Karcher equation, and the geometric mean is called the Karcher mean. It is shown in [4, 15] that the Karcher mean satisfies all 10 properties stated in [2]. Moreover the Karcher mean satisfies the Ando-Hiai inequality – a one of the most important operator inequality in the operator theory – [16, 17, 24], and the geometric mean which satisfies the Ando-Hiai inequality should be the Karcher mean [24]. Hence a lot of people study the Karcher mean.

As an extension of the Karcher mean, the power mean is defined in [17]. It interpolates the arithmetic, geometric (Karcher) and harmonic means, and it is defined by a solution of an operator equation. It is known that some operator inequalities relating to the power mean has been obtained [18]. In a recent year, Pálfia [20] generalized the Karcher mean by generalizing the Karcher equation. Then he obtained various kind

2010 *Mathematics Subject Classification.* Primary 47A63. Secondary 47A64.
Key words and phrases. Positive definite operator; operator mean; operator monotone function; power mean; Karcher equation; Karcher mean; relative operator entropy; Tsalise relative operator entropy; the Ando-Hiai inequality.

of operator means of n-operators. We can obtain the Karcher and power means as special cases of the new operator means. But we have not known any concrete relation between the generalized Karcher equation and operator means, i.e., we do not know which operator mean can be obtained from a given generalized Karcher equation.

In this report, we shall give a concrete relation between the generalized Karcher equation and operator means. In fact, we will give an inverse function of a representing function of an operator mean which is derived from a given generalized Karcher equation. In this discussion, representing function of relative operator entropy is very important. Next we shall give the Ando-Hiai type operator inequalities. Here we shall show two-types of Ando-Hiai inequalities, and we shall give a property of relative operator entropies. For the first type of the Ando-Hiai inequality, we shall give an Ando-Hiai type operator inequality for a given operator mean. The second one discusses an equivalence relation that the Ando-Hiai type inequality holds. The Ando-Hiai inequality was shown in [1], firstly. Then it has been extended to the Karcher and the power means in [16, 17, 18, 24]. On the other hand, the second type was firstly considered in [23]. He considered an arbitrary operator mean of 2-operators. In this report, we shall generalize these results into several operator means of n-operators which are derived from the generalized Karcher equation. At the same time, we shall study properties of relative operator entropies.

This report consists as follows: In Section 2, we shall introduce some basic notations, definitions and theorems. In Section 3, we shall obtain a relation among the generalized Karcher equation, relative operator entropy and operator means. In Section 4, we shall show the Ando-Hiai type inequalities for operator means which are derived from the solution of the generalized Karcher Equation.

2. Preliminaries

In what follows let $\mathcal{H}$ be a Hilbert space, and $\mathcal{B}(\mathcal{H})$ be a set of all bounded linear operators on $\mathcal{H}$. An operator $A \in \mathcal{B}(\mathcal{H})$ is positive definite (resp. positive semi-definite) if $\langle Ax, x\rangle > 0$ (resp. $\langle Ax, x\rangle \geq 0$) holds for all non-zero $x \in \mathcal{H}$. If A is positive semi-definite, we denote $A \geq 0$. Let $\mathcal{PS}, \mathcal{P} \subset \mathcal{B}(\mathcal{H})$ be the sets of all positive semi-definite operators and positive definite operators, respectively. For self-adjoint operators A and B, $A \geq B$ is defined by $A - B \geq 0$. A real-valued function f defined on an interval I satisfying

$$B \leq A \implies f(B) \leq f(A)$$

for all self-adjoint operators $A, B \in \mathcal{B}(\mathcal{H})$ such that $\sigma(A), \sigma(B) \in I$ is called an operator monotone function, where $\sigma(X)$ means the spectrum of $X \in \mathcal{B}(\mathcal{H})$.

2.1. Operator mean.

Definition 1 (Operator mean, [14]). Let $\sigma : \mathcal{PS}^2 \to \mathcal{PS}$ be a binary operation. If σ satisfies the following four conditions, σ is called an operator mean.

(1) If $A \leq C$ and $B \leq D$, then $\sigma(A, B) \leq \sigma(C, D)$,
(2) $X^*\sigma(A, B)X \leq \sigma(X^*AX, X^*BX)$ for all $X \in B(\mathcal{H})$,
(3) σ is upper semi-continuous on $\mathcal{PS}^2$,
(4) $\sigma(I, I) = I$, where I means the identity operator.

We notice that if X is invertible in (2), then equality holds.

Theorem A ([14]). *Let σ be an operator mean. Then there exists an operator monotone function f on $(0,\infty)$ such that $f(1)=1$ and*

$$\sigma(A,B)=A^{\frac{1}{2}}f(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})A^{\frac{1}{2}}$$

for all $A\in\mathcal{P}$ and $B\in\mathcal{PS}$. A function f is called a representing function of an operator mean σ.

Especially, if the assumption $f(1)=1$ is removed, then $\sigma(A,B)$ is called solidarity [7] or perspective[6]. Let $\varepsilon>0$ be a real number. Then we have $A_\varepsilon=A+\varepsilon I, B_\varepsilon=B+\varepsilon I\in\mathcal{P}$ for $A,B\in\mathcal{PS}$, and we can define an operator mean $\sigma(A,B)$ by $\sigma(A,B)=\lim_{\varepsilon\searrow 0}\sigma(A_\varepsilon,B_\varepsilon)$. We note that for an operator mean σ with a representing function f, $f'(1)=\lambda\in[0,1]$ (cf. [10, 21]), and we call σ a λ-weighted operator mean. Typical examples of operator means are the weighted geometric and weighted power means. These representing functions are $f(x)=x^\lambda$ and $f(x)=[1-\lambda+\lambda x^t]^{\frac{1}{t}}$, respectively, where $\lambda\in[0,1]$ and $t\in[-1,1]$ (in the case $t=0$, we consider $t\to 0$).The weighted power mean interpolates the arithmetic, geometric and harmonic means by putting $t=1,0,-1$, respectively. In what follows, the λ-weighted geometric and λ-weighted power means of $A,B\in\mathcal{PS}$ are denoted by $A\sharp_\lambda B$ and $P_t(\lambda;A,B)$, respectively, i.e.,

$$A\sharp_\lambda B=A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^\lambda A^{\frac{1}{2}},$$

$$P_t(\lambda;A,B)=A^{\frac{1}{2}}\left[1-\lambda+\lambda(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^t\right]^{\frac{1}{t}}A^{\frac{1}{2}}.$$

2.2. **The Karcher and the power means.** Geometric and power means of two-operators can be extended over more than 3-operators via the solution of operator equations as follows. Let n be a natural number, and let Δ_n be a set of all probability vectors, i.e.,

$$\Delta_n=\{\omega=(w_1,...,w_n)\in(0,1)^n|\ \sum_{i=1}^n w_i=1\}.$$

Definition 2 (The Karcher mean, [3, 16, 19]). Let $\mathbb{A}=(A_1,...,A_n)\in\mathcal{P}^n$ and $\omega=(w_1,...,w_n)\in\Delta_n$. Then the weighted Karcher mean $\Lambda(\omega;\mathbb{A})$ is defined by a unique positive solution of the following operator equation;

$$\sum_{i=1}^n w_i\log(X^{\frac{-1}{2}}A_iX^{\frac{-1}{2}})=0.$$

The Karchar mean of 2-operators coincides with the geometric mean of 2-operators, i.e., for each $\lambda\in[0,1]$, the solution of

$$(1-\lambda)\log(X^{\frac{-1}{2}}AX^{\frac{-1}{2}})+\lambda\log(X^{\frac{-1}{2}}BX^{\frac{-1}{2}})=0$$

is $X=A\sharp_\lambda B=A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^\lambda A^{\frac{1}{2}}$, easily. We can consider the Karcher mean as a geometric mean of n-operators. Properties of the Karcher mean are introduced in [16], for example.

The following power mean is an extension of the Karcher mean which interpolates the arithmetic, harmonic and the Karcher (geometric) means.

Definition 3 (The power mean, [16, 17]). Let $\mathbb{A} = (A_1, ..., A_n) \in \mathcal{P}^n$ and $\omega = (w_1, ..., w_n) \in \Delta$. Then for $t \in [-1, 1]$, the weighted power mean $P_t(\omega; \mathbb{A})$ is defined by a unique positive solution of the following operator equation;

$$\sum_{i=1}^{n} w_i (X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}})^t = I.$$

In fact, put $t = 1$ and $t = -1$, then the arithmetic and harmonic means are easily obtained, respectively. Also let $t \to 0$, we have the Karcher mean [16, 17]. Properties of the power mean are introduced in [16, 17].

Recently, the above operator equations are generalized as follows. Let $\mathcal{M}$ be a set of all operator monotone functions, and let

$$\mathcal{L} = \{g \in \mathcal{M} |\ g(1) = 0 \text{ and } g'(1) = 1\}.$$

Definition 4 (Generalized Karcher Equation (GKE), [20]). Let $g \in \mathcal{L}$, $\mathbb{A} = (A_1, ..., A_n) \in \mathcal{P}^n$ and $\omega = (w_1, ..., w_n) \in \Delta_n$. Then the following operator equation is called the Generalized Karcher equation (GKE).

$$\sum_{i=1}^{n} w_i g(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}) = 0. \tag{2.1}$$

Theorem B ([20]). *Any GKE has a unique solution $X \in \mathcal{P}$.*

The Karcher and the power means can be obtained by putting $g(x) = \log x$ and $g(x) = \frac{x^t - 1}{t}$ in (2.1), respectively. In what follows $\sigma_g(\omega; \mathbb{A})$ (or σ_g, simply) denotes the solution X of (2.1). Properties of σ_g are obtained in [20], here we state some of them as follows.

Theorem C ([20]). *Let $g \in \mathcal{L}$ and $\mathbb{A} = (A_1, ..., A_n), \mathbb{B} = (B_1, ..., B_n) \in \mathcal{P}^n$. Then σ_g satisfies the following properties.*

(1) *$\sigma_g(\omega; \mathbb{A}) \le \sigma_g(\omega; \mathbb{B})$ holds if $A_i \le B_i$ for all $i = 1, ..., n$,*
(2) *$X^* \sigma_g(\omega; \mathbb{A}) X = \sigma_g(\omega; X^* \mathbb{A} X)$ for all invertible $X \in B(\mathcal{H})$, where $X^* \mathbb{A} X = (X^* A_1 X, ..., X^* A_n X)$,*
(3) *σ_g is continuous on each operator $\mathcal{P}$, with respect to the Thompson metric,*
(4) *$\sigma_g(\omega; \mathbb{I}) = I$, where $\mathbb{I} = (I, ..., I)$.*

Moreover, $\sigma_g((1 - w, w); A, B)$ will be a w-weighted operator mean.

More generalization is discussed in [11, 20].

2.3. **Ando-Hiai inequality.** The Ando-Hiai inequality is one of the most important inequalities in the operator theory.

Theorem D (The Ando-Hiai inequality [1]). *Let $A, B \in \mathcal{PS}$ and $\lambda \in [0, 1]$. If $A \natural_\lambda B \le I$ holds, then $A^r \natural_\lambda B^r \le I$ holds for all $r \ge 1$.*

The Ando-Hiai inequality has been extended into the following two-types.

Theorem E (Extension of the Ando-Hiai inequality 1, [16, 17, 18, 24]). *Let $\mathbb{A} = (A_1, ..., A_n) \in \mathcal{P}^n$, $\omega \in \Delta_n$ and $t \in (0, 1]$. Then the following hold.*

(1) *$\Lambda(\omega; \mathbb{A}) \le I$ implies $\Lambda(\omega; \mathbb{A}^r) \le I$ for all $r \ge 1$,*

(2) $P_t(\omega;\mathbb{A}) \le I$ *implies* $P_{\frac{t}{r}}(\omega;\mathbb{A}^r) \le I$ *for all* $r \ge 1$,
(3) $P_{-t}(\omega;\mathbb{A}) \ge I$ *implies* $P_{-\frac{t}{r}}(\omega;\mathbb{A}^r) \ge I$ *for all* $r \ge 1$,

where $\mathbb{A}^r = (A_1^r, ..., A_n^r)$.

We remark that opposite inequalities of Theorems D and E (1) hold because $\Lambda(\omega;\mathbb{A})^{-1} = \Lambda(\omega;\mathbb{A}^{-1})$ holds for all $\mathbb{A} \in \mathcal{P}^n$ and $w \in \Delta_n$, where $\mathbb{A}^{-1} = (A_1^{-1}, ..., A_n^{-1})$. Moreover, the Karcher mean characterizes the property in Theorem E (1) [24].

We notice for Theorem E (2) and (3). Different power means appear in each statement, more precisely, there are power means with different parameters. Relating to the fact, the Ando-Hiai inequality has been extended to the following another form.

Theorem F (Extension of the Ando-Hiai inequality 2, [23]). *Let σ be an operator mean with a representing function f. Then the following are equivalent.*

(1) $f(x^r) \le f(x)^r$ *holds for all* $x \in (0,\infty)$ *and* $r \ge 1$,
(2) $\sigma(A,B) \le I$ *implies* $\sigma(A^r,B^r) \le I$ *for all* $A, B \in \mathcal{PS}$ *and* $r \ge 1$.

3. Relations among generalized Karcher equation, relative operator entropy and operator means

In this section, we shall give a relation between GKE and operator means. First, we shall give a concrete form of an inverse function of a representing function of an operator mean derived from GKE. Before introducing results, we notice as follows. A representing function of an operator mean is defined for only operator means of two operators. In this report, we usually treat operator means of n-operators, and as a special case, we can treat operator means of two-operators. Here we shall use a representing function of an operator mean which is defined by an operator mean of two-operators. More precisely, let $\sigma_g(\omega;\mathbb{A})$ be a solution of (2.1). Then for $\lambda \in (0,1)$, its representing function f is defined by

$$f_\lambda(x) = \sigma_g((1-\lambda,\lambda);1,x),$$

i.e., $f_\lambda(x)$ satisfies the following GKE:

$$(1-\lambda)g\left(\frac{1}{f_\lambda(x)}\right) + \lambda g\left(\frac{x}{f_\lambda(x)}\right) = 0 \qquad (3.1)$$

for all $x > 0$. We note that $f_1(x) = x$ and $f_0(x) = 1$ by (3.1). Hence we can define f_λ for all $\lambda \in [0,1]$.

Proposition 1 (see also [21]). *Let $g \in \mathcal{L}$. Then for each $\lambda \in (0,1)$, the inverse of f_λ in* (3.1) *is given by*

$$f_\lambda^{-1}(x) = xg^{-1}\left(-\frac{1-\lambda}{\lambda}g(x)\right).$$

Proof. Let σ_g be an operator mean derived from the GKE. Then for each $\lambda \in (0,1)$, $y = f_\lambda(x)$ satisfies the following equation

$$(1-\lambda)g\left(\frac{1}{y}\right) + \lambda g\left(\frac{x}{y}\right) = 0.$$

It is equivalent to

$$g\left(\frac{x}{y}\right) = -\frac{1-\lambda}{\lambda} g\left(\frac{1}{y}\right),$$

and thus

$$f_\lambda^{-1}(y) = x = y g^{-1}\left(-\frac{1-\lambda}{\lambda} g\left(\frac{1}{y}\right)\right).$$

The proof is completed. □

Proposition 2. *Let $g \in \mathcal{L}$. Then for each $\lambda \in [0,1]$, f_λ in* (3.1) *is differentiable on $\lambda \in (0,1)$ and*

$$\left.\frac{\partial}{\partial\lambda} f_\lambda(x)\right|_{\lambda=0} = \lim_{\lambda\searrow 0} \frac{\partial}{\partial\lambda} f_\lambda(x) = g(x).$$

For the Karcher mean case, $f_\lambda(x) = x^\lambda$ and $g(x) = \frac{\partial}{\partial\lambda} x^\lambda|_{\lambda=0} = \log x$, and $g(x) = \log x$ is a representing function of the relative operator entropy [8]. In fact let $A, B \in \mathcal{P}$. Then the relative operator entropy $S(A|B)$ is defined by

$$S(A|B) = A^{\frac{1}{2}} \log(A^{\frac{-1}{2}} B A^{\frac{-1}{2}}) A^{\frac{1}{2}}.$$

For the power mean case, $f_\lambda(x) = [1-\lambda+\lambda x^t]^{\frac{1}{t}}$ and $g(x) = \frac{\partial}{\partial\lambda}[1-\lambda+\lambda x^t]^{\frac{1}{t}}|_{\lambda=0} = \frac{x^t-1}{t}$, and $g(x) = \frac{x^t-1}{t}$ is a representing function of the Tsallis relative operator entropy [25]. In fact let $A, B \in \mathcal{P}$. Then the Tsallis relative operator entropy $T_t(A|B)$ is defined by

$$T(A|B) = A^{\frac{1}{2}} \frac{(A^{\frac{-1}{2}} B A^{\frac{-1}{2}})^t - I}{t} A^{\frac{1}{2}} = \frac{A\natural_t B - A}{t}.$$

So relative operator entropy is closely related to the GKE and operator means.

Proof of Proposition 2. First of all, $g \in \mathcal{L}$ is a differentiable function since g is an operator monotone function. By $g \in \mathcal{L}$ and Proposition 1, the representing function $f_\lambda(x)$ is differentiable on $\lambda \in (0,1)$, and it satisfies (3.1). By differentiating (3.1) both side on λ, we have

$$\begin{aligned}-g\left(\frac{1}{f_\lambda(x)}\right) + (1-\lambda) g'\left(\frac{1}{f_\lambda(x)}\right)\left(-\frac{1}{f_\lambda(x)^2}\right)\frac{\partial}{\partial\lambda} f_\lambda(x)& \\ + g\left(\frac{x}{f_\lambda(x)}\right) + \lambda g'\left(\frac{x}{f_\lambda(x)}\right)\left(\frac{-x}{f_\lambda(x)^2}\right)\frac{\partial}{\partial\lambda} f_\lambda(x) &= 0.\end{aligned}$$

Here we take a limit $\lambda \searrow 0$, by $f_\lambda(x) \to 1$, $g(1) = 0$ and $g'(1) = 1$, we have

$$\left(-\left.\frac{\partial}{\partial\lambda} f_\lambda(x)\right|_{\lambda=0}\right) + g(x) = 0.$$

Hence we have

$$\left.\frac{\partial}{\partial\lambda} f_\lambda(x)\right|_{\lambda=0} = g(x).$$

□

4. The Ando-Hiai inequalities for the solution of the GKE

In this section, we shall show extensions of Ando-Hiai inequalities. To prove them, the following result is very important.

Theorem 3. *Let $g \in \mathcal{L}$, $\mathbb{A} = (A_1, ..., A_n) \in \mathcal{P}^n$ and $\omega = (w_1, ..., w_n) \in \Delta_n$. Then the following hold.*

(1) $\sum_{i=1}^n w_i g(A_i) \geq 0$ *implies* $\sigma_g(\omega; \mathbb{A}) \geq I$, *and*
(2) $\sum_{i=1}^n w_i g(A_i) \leq 0$ *implies* $\sigma_g(\omega; \mathbb{A}) \leq I$.

To prove Theorem 3, we shall prepare the following property of σ_g.

Lemma 4. *Let $g \in \mathcal{L}$, $\omega = (w_1, ..., w_n) \in \Delta_n$ and $\mathbb{A} = (A_1, ..., A_n) \in \mathcal{P}^n$. Then*

$$\left[\sum_{i=1}^n w_i A_i^{-1}\right]^{-1} \leq \sigma_g(\omega; \mathbb{A}) \leq \sum_{i=1}^n w_i A_i.$$

Proof. We note that for each $g \in \mathcal{L}$,

$$1 - x^{-1} \leq g(x) \leq x - 1$$

holds for all $x \in (0, \infty)$ [20, (18)]. Let $X = \sigma_g(\omega; \mathbb{A})$. Then we have

$$0 = \sum_{i=1}^n w_i g(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}) \leq \sum_{i=1}^n w_i (X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}} - I),$$

i.e., $X \leq \sum_{k=1}^n w_i A_i$. The latter part can be shown by the same way by using $g(x) \geq 1 - x^{-1}$. □

Proof of Theorem 3. Proof of (1). Assume that $\sum_{i=1}^n w_i g(A_i) \geq 0$ holds. Since an operator monotone function g satisfies $g(1) = 0$, there exists $X \leq I$ such that

$$\sum_{i=1}^n \frac{w_i}{2} g(A_i) + \frac{1}{2} g(X) = 0.$$

Hence we have

$$I = \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X)\right),$$

where $(\frac{\omega}{2}, \frac{1}{2}) = (\frac{w_1}{2}, ..., \frac{w_n}{2}, \frac{1}{2}) \in \Delta_{n+1}$ and $(\mathbb{A}, X) = (A_1, ..., A_n, X) \in \mathcal{P}^{n+1}$. Here we define an operator sequence $\{X_k\} \subset \mathcal{P}$ by

$$X_0 = I, \ X_{k+1} = \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X_k)\right).$$

Then

$$\begin{aligned} X_0 = I &= \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X)\right) \\ &\leq \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, I)\right) = X_1 \\ &\leq \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X_1)\right) = X_2 \leq \cdots \leq X_n, \end{aligned}$$

where the inequalities hold by operator monotonicity of σ_g, i.e., Theorem C (1). By Lemma 4, we have

$$X_k \leq \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X_k)\right) \leq \sum_{i=1}^{n} \frac{w_i}{2} A_i + \frac{1}{2} X_k,$$

and we have $X_k \leq \sum_{i=1}^{n} w_i A_i$ for all $k = 1, 2, \ldots$. Hence there exists a unique limit point $\lim_{k\to\infty} X_k = X_\infty \in \mathcal{P}$. It satisfies

$$X_\infty = \sigma_g\left(\left(\frac{\omega}{2}, \frac{1}{2}\right); (\mathbb{A}, X_\infty)\right),$$

and then we have

$$\sum_{i=1}^{n} w_i g(X_\infty^{\frac{-1}{2}} A_i X_\infty^{\frac{-1}{2}}) = 0,$$

that is,

$$I \leq X_\infty = \sigma_g(\omega; \mathbb{A}).$$

Proof of (2) is shown by the same way and using $[\sum_{i=1}^{n} w_i A_i^{-1}]^{-1} \leq \sigma_g(\omega; \mathbb{A})$. □

Using Theorem 3, we can get an elementary property of the solution of the GKE.

Theorem 5. *Let $f, g \in \mathcal{L}$. Then $g(x) \leq f(x)$ holds for all $x \in (0, \infty)$ if and only if $\sigma_g(\omega; \mathbb{A}) \leq \sigma_f(\omega; \mathbb{A})$ holds for all $\omega \in \Delta_n$ and $\mathbb{A} \in \mathcal{P}^n$.*

Proof. Proof of ($\Longrightarrow$). Let $\omega = (w_1, \ldots, w_n) \in \Delta_n$, $\mathbb{A} = (A_1, \ldots, A_n) \in \mathcal{P}^n$ and $X = \sigma_g(\omega; \mathbb{A})$. Assume that $g(x) \leq f(x)$ holds for all $x \in (0, \infty)$. Then

$$0 = \sum_{i=1}^{n} w_i g(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}) \leq \sum_{i=1}^{n} w_i f(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}).$$

By Theorem 3, we have $I \leq \sigma_f(\omega; X^{\frac{-1}{2}} \mathbb{A} X^{\frac{-1}{2}}) = X^{\frac{-1}{2}} \sigma_f(\omega; \mathbb{A}) X^{\frac{-1}{2}}$, i.e.,

$$\sigma_g(\omega; \mathbb{A}) = X \leq \sigma_f(\omega; \mathbb{A}).$$

Proof of ($\Longleftarrow$). It is enough to consider the two-operators case. For $\lambda \in [0, 1]$, let $r_{g,\lambda}$ and $r_{f,\lambda}$ are the representing functions of $\sigma_g(1-\lambda, \lambda; A, B)$ and $\sigma_f(1-\lambda, \lambda; A, B)$, respectively. Then $r_{g,\lambda}(x) \leq r_{f,\lambda}(x)$ holds for all $x \in (0, \infty)$ and $\lambda \in [0, 1]$, and we have

$$\frac{r_{g,\lambda}(x) - 1}{\lambda} \leq \frac{r_{f,\lambda}(x) - 1}{\lambda}$$

holds for all $x \in (0, \infty)$ and $\lambda \in (0, 1]$. Let $\lambda \searrow 0$, we have $g(x) \leq f(x)$ by Proposition 2. □

Here, we shall show extensions of the Ando-Hiai inequality.

Theorem 6 (Extension of the Ando-Hiai inequality, 1). *Let $g \in \mathcal{L}$, $\mathbb{A} \in \mathcal{P}^n$ and $\omega \in \Delta_n$. If $\sigma_g(\omega; \mathbb{A}) \leq I$ holds, then $\sigma_{g_p}(\omega; \mathbb{A}^p) \leq I$holds for all $p \geq 1$, where $g_p(x) = pg(x^{1/p})$. Moreover the representing function of σ_{g_p} is $f_{p,\lambda}(x) = f_\lambda(x^{1/p})^p$, where f_λ is a representing function of λ-weighted operator mean σ_g.*

We notice that $g_p(x^{1/p}) \in \mathcal{L}$ for all $p \geq 1$.

By putting $g(x) = \log x$ in Theorem 6, σ_g coincides with the Karcher mean. Then we have Theorem E (1). Moreover put $g(x) = \frac{x^t - 1}{t}$ in Theorem 6, σ_g coincides with the power mean. Then we have Theorem E (2).

Proof of Theorem 6. Let $X = \sigma_g(\omega; \mathbb{A}) \leq I$. For $p \in [1, 2]$, we have

$$0 = \sum_{i=1}^{n} w_i g(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}) = \sum_{i=1}^{n} w_i g((X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}})^{\frac{p}{p}}) \geq \sum_{i=1}^{n} w_i g((X^{\frac{-1}{2}} A_i^p X^{\frac{-1}{2}})^{\frac{1}{p}}),$$

where the last inequality holds by the Hansen's inequality [9].

Hence

$$0 \geq \sum_{i=1}^{n} w_i p g((X^{\frac{-1}{2}} A_i^p X^{\frac{-1}{2}})^{\frac{1}{p}}) = \sum_{i=1}^{n} w_i g_p(X^{\frac{-1}{2}} A_i^p X^{\frac{-1}{2}}),$$

and we have $\sigma_{g_p}(\omega; X^{\frac{-1}{2}} \mathbb{A}^p X^{\frac{-1}{2}}) \leq I$ by Theorem 3, i.e.,

$$\sigma_{g_p}(\omega; \mathbb{A}^p) \leq X \leq I$$

for $p \in [1, 2]$. Applying the same way to $\sigma_{g_p}(\omega; \mathbb{A}^p) \leq I$, we have $\sigma_{g_{pp'}}(\omega; \mathbb{A}^{pp'}) \leq I$ for $p' \in [1, 2]$ and $pp' \in [1, 4]$. Repeating this method, we have $\sigma_{g_p}(\omega; \mathbb{A}^p) \leq I$ for all $p \geq 1$.

Let f_λ be a representing function of a λ-weighted operator mean σ_g, and let $f_{p,\lambda}$ be a representing function of σ_{g_p}. We note that the inverse function of $g_p(x) = pg(x^{1/p})$ is $\{g^{-1}(\frac{x}{p})\}^p$. Hence by Proposition 1, we have

$$\begin{aligned}
f_{p,\lambda}^{-1}(x) &= x g_p^{-1}\left(-\frac{1-\lambda}{\lambda} g_p\left(\frac{1}{x}\right)\right) \\
&= x\left\{g^{-1}\left(-\frac{1-\lambda}{p\lambda} \cdot pg\left(\frac{1}{x^{1/p}}\right)\right)\right\}^p \\
&= \left\{x^{\frac{1}{p}} g^{-1}\left(-\frac{1-\lambda}{\lambda} g\left(\frac{1}{x^{1/p}}\right)\right)\right\}^p \\
&= f_\lambda^{-1}(x^{1/p})^p.
\end{aligned}$$

Therefore $f_{p,\lambda}(x) = f_\lambda(x^{1/p})^p$. □

We can prove the opposite inequalities in Theorem 6 by the same way.

Theorem 7 (Extension of the Ando-Hiai inequality, 2). *Let $g \in \mathcal{L}$. Assume f_λ is a representing function of an operator mean $\sigma_g(1-\lambda, \lambda; A, B)$. Then the following are equivalent.*

(1) *$f_\lambda(x)^p \leq f_\lambda(x^p)$ holds for all $p \geq 1$, $\lambda \in [0, 1]$ and $x \in (0, \infty)$,*
(2) *$pg(x) \leq g(x^p)$ for all $p \geq 1$ and $x \in (0, \infty)$,*
(3) *$\sigma_g(\omega; \mathbb{A}) \geq I$ implies $\sigma_g(\omega; \mathbb{A}^p) \geq I$ for all $\omega \in \Delta_n$, $\mathbb{A} \in \mathcal{P}^n$ and $p \geq 1$.*

For the two-operators case, Theorem 7 coincides with the opposite inequality of Theorem F (it was shown in [23]). Moreover, we can obtain a property of relative operator entropy in the above theorem but it is not given in Theorem F.

Proof. Proof of (1) $\Longrightarrow$ (2). Since $1+p(x-1) \leq x^p$ holds for all $p \geq 1$ and $x \in (0,\infty)$, we have

$$p\left(\frac{f_\lambda(x)-1}{\lambda}\right) \leq \frac{f_\lambda(x)^p-1}{\lambda} \leq \frac{f_\lambda(x^p)-1}{\lambda}$$

holds for all $p \geq 1$, $\lambda \in (0,1]$ and $x \in (0,\infty)$ by the assumption. By letting $\lambda \searrow 0$, we have $pg(x) \leq g(x^p)$ by Proposition 2.

Proof of (2) $\Longrightarrow$ (3). Let $X = \sigma_g(\omega; \mathbb{A}) \geq I$. For $p \in [1,2]$,

$$\begin{aligned} 0 &= \sum_{i=1}^{n} w_i p g(X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}}) \\ &\leq \sum_{i=1}^{n} w_i g((X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}})^p) \qquad \text{(by (2))} \\ &\leq \sum_{i=1}^{n} w_i g(X^{\frac{-1}{2}} A_i^p X^{\frac{-1}{2}}), \end{aligned}$$

where the last inequality holds by the Hansen's inequality [9]. Hence by Theorem 3, we have $I \leq \sigma_g(\omega; X^{\frac{-1}{2}} \mathbb{A}^p X^{\frac{-1}{2}})$, i.e.,

$$I \leq X \leq \sigma_g(\omega; \mathbb{A}^p).$$

Applying the same way to $I \leq \sigma_g(\omega; \mathbb{A}^p)$, we have $I \leq \sigma_g(\omega; \mathbb{A}^{pp'})$ for $p' \in [1,2]$ and $pp' \in [1,4]$. Repeating this method, we have $I \leq \sigma_g(\omega; \mathbb{A}^p)$ for all $p \geq 1$.

Proof of (3) $\Longrightarrow$ (1) is shown in [23]. □

By the similar way, we have the following result.

Theorem 7'. *Let $g \in \mathcal{L}$. Assume f_λ is a representing function of an operator mean $\sigma_g(1-\lambda, \lambda; A, B)$. Then the following are equivalent.*

(1) *$f_\lambda(x)^p \geq f_\lambda(x^p)$ holds for all $p \geq 1$, $\lambda \in [0,1]$ and $x \in (0,\infty)$,*
(2) *$pg(x) \geq g(x^p)$ for all $p \geq 1$ and $x \in (0,\infty)$,*
(3) *$\sigma_g(\omega; \mathbb{A}) \leq I$ implies $\sigma_g(\omega; \mathbb{A}^p) \leq I$ for all $\omega \in \Delta$, $\mathbb{A} \in \mathcal{P}^n$ and $p \geq 1$.*

It is just an extension of Theorem F.

References

[1] T. Ando and F. Hiai, *Log majorization and complementary Golden- Thompson type inequalities*, Linear Algebra Appl., **197/198** (1994), 113–131.

[2] T. Ando, C.-K. Li and R. Mathias, *Geometric means*, Linear Algebra Appl., **385** (2004), 305–334.

[3] R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means*, Linear Algebra Appl., **413** (2006), 594–618.

[4] R. Bhatia and R. L. Karandikar, *Monotonicity of the matrix geometric mean*, Math. Ann., **353** (2012), 1453–1467.

[5] D. A. Bini, B. Meini and F. Poloni, *An effective matrix geometric mean satisfying the Ando-Li-Mathias properties*, Math. Comp., **79** (2010), 437–452.

[6] E. Effros and F. Hansen, *Non-commutative perspectives*, Ann. Funct. Anal., 5 (2014), 74–79.

[7] J. I. Fujii, M. Fujii and Y. Seo, *An extension of the Kubo-Ando theory: solidarities*, Math. Japon., **35** (1990), 387–396.

[8] J. I. Fujii, *Operator means and the relative operator entropy*, Operator theory and complex analysis (Sapporo, 1991), 161–172, Oper. Theory Adv. Appl., **59**, Birkhäuser, Basel, 1992.
[9] F. Hansen, *An operator inequality*, Math. Ann., **246** (1979/80), 249–250.
[10] F. Hansen, *The fast track to Löwner's theorem*, Linear Algebra Appl., **438** (2013), 4557–4571.
[11] F. Hiai, *Operator means deformed by a fixed point method*, arXiv:1711.10170.
[12] S. Izumino and N. Nakamura, *Geometric means of positive operators. II*, Sci. Math. Jpn., **69** (2009), 35–44.
[13] S. Izumino and N. Nakamura, *Weighted geometric means of positive operators*, Kyungpook Math. J., **50** (2010), 213–228.
[14] F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann., **246** (1979/80), 205–224.
[15] J. Lawson and Y.Lim, *Monotonic properties of the least squares mean*, Math. Ann., **351** (2011), 267–279.
[16] J. Lawson and Y. Lim, *Karcher means and Karcher equations of positive definite operators*, Trans. Amer. Math. Soc. Ser. B, **1** (2014), 1–22.
[17] Y. Lim and M. Pálfia, *Matrix power means and the Karcher mean*, J. Funct. Anal., **262** (2012), 1498–1514.
[18] Y. Lim and T. Yamazaki, *On some inequalities for the matrix power and Karcher means*, Linear Algebra Appl., **438** (2013), 1293–1304.
[19] M. Moakher, *A differential geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl., **26** (2005), 735–747.
[20] M. Pálfia, *Operator means of probability measures and generalized Karcher equations*, Adv. Math., **289** (2016), 951–1007.
[21] M. Pálfia, *Semigroups of operator means and generalized Karcher equations*, arXiv:1208.5603.
[22] Pusz and Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys., **8** (1975), 159–170.
[23] S. Wada, *Some ways of constructing Furuta-type inequalities*, Linear Algebra Appl., **457** (2014), 276–286.
[24] T. Yamazaki, *The Riemannian mean and matrix inequalities related to the Ando-Hiai inequality and chaotic order*, Oper. Matrices, **6** (2012), 577–588.
[25] K. Yanagi, K. Kuriyama and S. Furuichi, *Generalized Shannon inequalities based on Tsallis relative operator entropy*, Linear Algebra Appl., **394** (2005), 109–118.

Department of Electrical, Electronic and Computer Engineering, Toyo University, Kawagoe-Shi, Saitama, 350-8585, Japan.
E-mail address: t-yamazaki@toyo.jp