Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 7, Issue 1, Article 36, 2006

CHARACTERIZATIONS OF TRACIAL PROPERTY VIA INEQUALITIES

TAKASHI SANO AND TAKESHI YATSU

DEPARTMENT OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
YAMAGATA UNIVERSITY
YAMAGATA 990-8560, JAPAN.
sano@sci.kj.yamagata-u.ac.jp

Received 13 October, 2005; accepted 08 December, 2005
Communicated by F. Hansen

Dedicated to Professor Marie Choda on the occasion of her 65th birthday.

ABSTRACT. In this article, we give characterizations of a tracial property for a positive linear functional via inequalities; we have necessary and sufficient conditions for a faithful positive linear functional φ to be a positive scalar multiple of the trace by inequalities: for a non matrix monotone, increasing function f,

$$X \leqq Y \Rightarrow \varphi(f(X)) \leqq \varphi(f(Y))$$

is considered. Also for a non matrix convex, convex function f,

$$\varphi\left(f\left(\frac{X+Y}{2}\right)\right) \leqq \varphi\left(\frac{f(X)+f(Y)}{2}\right)$$

is studied. We also show that suppose

$$0 \leqq \varphi\left(p_{m,k}\left(X,Y\right)\right)$$

for all $X, Y \geqq O$, then φ should be a positive scalar multiple of the trace. Here, $p_{m,k}(X,Y)$ is the coefficient of t^k in the polynomial $(X+tY)^m$ and $1 \leqq k \leqq m-1$.

Key words and phrases: Trace; Inequality; Non matrix monotone function of order 2; Non matrix convex function of order 2; Bessis-Moussa-Villani conjecture.

2000 *Mathematics Subject Classification.* 15A42, 47A63, 15A60, 47A30.

1. INTRODUCTION

In operator theory, matrix monotone functions and matrix convex ones have played a significant role, for instance, see [3, 4, 1, 2]. A real-valued continuous function f on an interval I

ISSN (electronic): 1443-5756

298-05

($\subseteqq \mathbb{R}$) is called matrix monotone of order n if $X \leqq Y$ implies $f(X) \leqq f(Y)$ for all $n \times n$ Hermitian matrices X and Y with eigenvalues in I. If f is matrix monotone of all orders, f is said to be matrix monotone or operator monotone. When f is matrix monotone of order $n,$ for a positive linear functional φ on $n \times n$ matrices, we have

$$X \leqq Y \Rightarrow \varphi(f(X)) \leqq \varphi(f(Y))$$

for $n\times n$ Hermitian matrices X and Y. Also, for an increasing function f and Hermitian matrices X and Y with $X \leqq Y$,

$$\mathrm{Tr}(f(X)) \leqq \mathrm{Tr}(f(Y))$$

holds in which Tr is the standard trace on matrices (for more details, see the argument at the beginning of Section 2).

A real-valued continuous function f on an interval I ($\subseteqq \mathbb{R}$) is called matrix convex of order n if the inequality

$$f\left(\frac{X+Y}{2}\right) \leqq \frac{f(X)+f(Y)}{2}$$

is satisfied for all $n \times n$ Hermitian matrices X and Y with eigenvalues in I. If f is matrix convex of all orders, f is said to be matrix convex or operator convex. When f is matrix convex of order $n,$ for a positive linear functional φ on $n \times n$ matrices,

$$\varphi\left(f\left(\frac{X+Y}{2}\right)\right) \leqq \varphi\left(\frac{f(X)+f(Y)}{2}\right)$$

holds for $n \times n$ Hermitian matrices X and Y with eigenvalues in I. And for a convex function f and Hermitian matrices X and Y, we have

$$\mathrm{Tr}\left(f\left(\frac{X+Y}{2}\right)\right) \leqq \mathrm{Tr}\left(\frac{f(X)+f(Y)}{2}\right)$$

(see basic facts on Jensen's inequalities explained before the proof of Theorem 3.3).

In this article, we give characterizations of tracial properties for positive linear functionals via inequalities; we have necessary and sufficient conditions for a faithful positive linear functional φ to be tracial by inequalities: for a non matrix monotone, increasing function f,

$$X \leqq Y \Rightarrow \varphi(f(X)) \leqq \varphi(f(Y))$$

is considered. Also for a non matrix convex, convex function f,

$$\varphi\left(f\left(\frac{X+Y}{2}\right)\right) \leqq \varphi\left(\frac{f(X)+f(Y)}{2}\right)$$

is studied. We have a criterion of non matrix monotonicity of order 2 or non matrix convexity of order 2. We show a necessary and sufficient condition for the function

$$X \mapsto \{\mathrm{Tr}(|X|^p C)\}^{\frac{1}{p}}$$

$(p > 2)$ to be a norm; the function is essentially the Schatten p-norm.

We also observe an inequality given by a coefficient of a certain polynomial: let $p_{m,k}(X,Y)$ be the coefficient of t^k in the polynomial $(X+tY)^m$ for $X,Y \in M_n(\mathbb{C})$, $m \in \mathbb{N}$, and $t \in \mathbb{C}$ and $1 \leqq k \leqq m-1$. Suppose that

$$0 \leqq \varphi\left(p_{m,k}\left(X,Y\right)\right)$$

for all $X,Y \geqq O$. Then φ should be a positive scalar multiple of the trace (see the remark of Proposition 3.1 about the BMV conjecture).

We remark that divided differences are useful in this article: we refer the reader to [4, 2, 6].

We would like to express our sincere gratitude to Professor Tsuyoshi Ando for reading the previous manuscripts and for fruitful comments. We would like to thank the members of Tohoku-Seminar for valuable advice, especially Professor Sin-ei Takahashi for useful comments on Proposition 2.1 and Professor Fumio Hiai for pointing out the BMV conjecture to us. We are also grateful to the editor and the referee for careful reading of the manuscripts and for helpful comments.

2. INEQUALITIES OF NON MATRIX MONOTONE FUNCTIONS

Let $M_n(\mathbb{C})$ be the set of all complex n-square matrices and let φ be a faithful positive linear functional on $M_n(\mathbb{C})$. Let f be an increasing function on $I=(a,b)$. For Hermitian matrices $X,Y \in M_n(\mathbb{C})$ with $a1 < X \leqq Y < b1$,

$$\lambda_i(X) \leqq \lambda_i(Y)$$

for $i=1,2,\dots,n$ in which λ_i is the i-th eigenvalue with $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_n$. Since f is increasing,

$$\lambda_i(f(X)) = f(\lambda_i(X)) \leqq f(\lambda_i(Y)) = \lambda_i(f(Y)).$$

Hence, it follows that

$$\mathrm{Tr}(f(X)) = \sum_{i=1}^{n} \lambda_i(f(X)) \leqq \sum_{i=1}^{n} \lambda_i(f(Y)) = \mathrm{Tr}(f(Y)).$$

Let us study the following inequality for a strictly increasing, differentiable function f on $I=(a,b)$ and $\varepsilon \in [0,1]$:

$$f'(\lambda)\alpha^2(1+\varepsilon) - 2\alpha\sqrt{1-\alpha^2}\,\frac{f(\lambda)-f(\mu)}{\lambda-\mu}(1-\varepsilon) + f'(\mu)\left(1-\alpha^2\right)(1+\varepsilon) \geqq 0$$

for all $\mu,\lambda \in I$ $(\mu<\lambda)$ and all $\alpha \in [0,1]$. By considering $0<\alpha<1$, we have the equivalent inequality

$$\frac{\alpha}{\sqrt{1-\alpha^2}}f'(\lambda) + \frac{\sqrt{1-\alpha^2}}{\alpha}f'(\mu) \geqq 2\,\frac{1-\varepsilon}{1+\varepsilon}\cdot\frac{f(\lambda)-f(\mu)}{\lambda-\mu}$$

for all $\mu,\lambda \in I$ $(\mu<\lambda)$ and all $\alpha \in (0,1)$. Let

$$t := \frac{\alpha}{\sqrt{1-\alpha^2}}, \quad \delta := \frac{1-\varepsilon}{1+\varepsilon}.$$

Then notice that

$$0 < \alpha < 1 \Leftrightarrow 0 < t < \infty, \quad 0 \leqq \varepsilon \leqq 1 \Leftrightarrow 0 \leqq \delta \leqq 1,$$

and $\varepsilon = 1$ if and only if $\delta = 0$. The corresponding inequality is described as

$$\frac{1}{2}\left(tf'(\lambda) + \frac{1}{t}f'(\mu)\right) \geqq \delta\, \frac{f(\lambda) - f(\mu)}{\lambda - \mu}$$

for all $0 < t < \infty$ and all $\mu, \lambda \in I$ $(\mu < \lambda)$. Hence, by considering arithmetic-geometric mean inequality in the left-hand side, we have

$$\sqrt{f'(\lambda)f'(\mu)} \geqq \delta\, \frac{f(\lambda) - f(\mu)}{\lambda - \mu}.$$

In this case, the condition $\varepsilon = 1$ or $\delta = 0$ is given by

$$\inf_{\lambda > \mu} \frac{\sqrt{f'(\lambda)f'(\mu)}}{\frac{f(\lambda)-f(\mu)}{\lambda-\mu}} = 0.$$

We summarize our observation as follows:

Proposition 2.1. *Let f be a strictly increasing, continuously differentiable function on $I = (a, b)$ and $\varepsilon \in [0, 1]$. Suppose that*

$$\inf_{\lambda > \mu} \frac{\sqrt{f'(\lambda)f'(\mu)}}{\frac{f(\lambda)-f(\mu)}{\lambda-\mu}} = 0.$$

Then the inequality

$$f'(\lambda)\alpha^2(1+\varepsilon) - 2\alpha\sqrt{1-\alpha^2}\, \frac{f(\lambda) - f(\mu)}{\lambda - \mu}(1-\varepsilon) + f'(\mu)(1-\alpha^2)(1+\varepsilon) \geqq 0$$

holds for all $\mu, \lambda \in I$ $(\mu < \lambda)$ and all $\alpha \in [0, 1]$ if and only if $\varepsilon = 1$.

The following are examples:

$$x^p (p > 1) \ \ on\ (0, a), \qquad x^p (p > 1) \ \ on\ (a, \infty),$$
$$e^x \ \ on\ (a, \infty), \qquad e^x \ \ on\ (-\infty, a)$$

for a constant a.

By direct computations, it is easy to see that each example satisfies the condition so details are left to the reader.

Theorem 2.2. *Let φ be a faithful positive linear functional on $M_n(\mathbb{C})$ and let f be a function as in Proposition 2.1. Then*

$$\varphi(f(X)) \leqq \varphi(f(Y)) \quad \textit{whenever} \quad aI < X \leqq Y < bI \tag{2.1}$$

if and only if φ is a positive scalar multiple of the trace.

Proof. At the beginning of this section it was explained that if φ is a positive scalar multiple of the trace then the inequality (2.1) holds. We show the converse: since there is uniquely a positive definite matrix D such that $\varphi(X) = \mathrm{Tr}(XD)$ for $X \in M_n(\mathbb{C})$, we have to prove that D is a positive scalar multiple of the identity matrix. Taking into consideration

$$\varphi\left(V^* \cdot V\right) = \mathrm{Tr}\left(\cdot\, VDV^*\right)$$

for all unitary V and that VDV^* is diagonal for a unitary V, we assume that D is a diagonal matrix $\mathrm{diag}(d_1, \ldots, d_n)$. To show that $d_i = d_j$ for any pair of d_i, d_j $(i \neq j)$, we consider matrices $X = (x_{kl})$ with x_{kl} zero except for $(k,l) = (i,i), (i,j), (j,i), (j,j)$. Hence, it suffices to consider the case $n = 2$ so that we suppose

$$D = \mathrm{diag}(\varepsilon, 1)$$

for a number ε $(0 < \varepsilon \leqq 1)$. We show that $\varepsilon = 1$.

Let

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad A_{\lambda,\mu} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \quad P_\alpha = \begin{pmatrix} \alpha^2 & \alpha\sqrt{1-\alpha^2} \\ \alpha\sqrt{1-\alpha^2} & 1-\alpha^2 \end{pmatrix}$$

for $\lambda, \mu (\lambda \neq \mu, a < \lambda, \mu < b)$, α $(0 \leqq \alpha \leqq 1)$. For all $t > 0$,

$$UA_{\lambda,\mu}U^* \leqq U(A_{\lambda,\mu} + tP_\alpha)U^*,$$

and $a1 < A_{\lambda,\mu} + tP_\alpha < b1$ for small $t > 0$. Then, by assumption

$$\mathrm{Tr}(Uf(A_{\lambda,\mu})U^*D) \leqq \mathrm{Tr}(Uf(A_{\lambda,\mu} + tP_\alpha)U^*D).$$

This implies that

$$\frac{d}{dt}\mathrm{Tr}(Uf(A_{\lambda,\mu} + tP_\alpha)U^*D)\bigg|_{t=0} \geqq 0.$$

Also, the standard fact, see [2, page 124] for instance, yields

$$\begin{aligned}
\frac{d}{dt}\bigg|_{t=0} f(A_{\lambda,\mu} + tP_\alpha) &= \begin{pmatrix} f'(\lambda) & \frac{f(\lambda)-f(\mu)}{\lambda-\mu} \\ \frac{f(\lambda)-f(\mu)}{\lambda-\mu} & f'(\mu) \end{pmatrix} \circ P_\alpha \\
&= \begin{pmatrix} f'(\lambda)\alpha^2 & \frac{f(\lambda)-f(\mu)}{\lambda-\mu}\alpha\sqrt{1-\alpha^2} \\ \frac{f(\lambda)-f(\mu)}{\lambda-\mu}\alpha\sqrt{1-\alpha^2} & f'(\mu)(1-\alpha^2) \end{pmatrix},
\end{aligned}$$

where $\circ$ stands for the Hadamard (i.e., entry-wise) product. Hence, it follows that

$$\begin{aligned}
0 &\leqq \frac{d}{dt}\mathrm{Tr}(f(A_{\lambda,\mu} + tP_\alpha)U^*DU)\bigg|_{t=0} \\
&= \mathrm{Tr}\left(\frac{d}{dt}\bigg|_{t=0} f(A_{\lambda,\mu} + tP_\alpha)\right) \cdot U^*DU) \\
&= \mathrm{Tr}\left(\begin{pmatrix} f'(\lambda)\alpha^2 & \frac{f(\lambda)-f(\mu)}{\lambda-\mu}\alpha\sqrt{1-\alpha^2} \\ \frac{f(\lambda)-f(\mu)}{\lambda-\mu}\alpha\sqrt{1-\alpha^2} & f'(\mu)(1-\alpha^2) \end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix} 1+\varepsilon & -(1-\varepsilon) \\ -(1-\varepsilon) & 1+\varepsilon \end{pmatrix}\right) \\
&= \frac{1}{2}\left\{f'(\lambda)\alpha^2(1+\varepsilon) - 2\alpha\sqrt{1-\alpha^2}\,\frac{f(\lambda)-f(\mu)}{\lambda-\mu}(1-\varepsilon) + f'(\mu)(1-\alpha^2)(1+\varepsilon)\right\}
\end{aligned}$$

for all α, λ, μ. Therefore, thanks to Proposition 2.1, the proof is completed. □

In the proof of Theorem 2.2, a criterion of non matrix monotonicity of order 2 is obtained:

Corollary 2.3. *Let f be a function as in Proposition 2.1. Then f is not matrix monotone of order 2.*

As a corollary of Theorem 2.2, we have

Theorem 2.4. *Let $C \in M_n(\mathbb{C})$ be a positive definite matrix and let $p > 2$ be given. Then the function*

$$\mu(X) := \mathrm{Tr}(|X|^p C)^{\frac{1}{p}} \quad (X \in M_n(\mathbb{C}))$$

is a norm if and only if C is a positive scalar multiple of the identity matrix.

Proof. Suppose that μ is a norm. Then, by definition

$$\mu(UX) = \mu(X)$$

for all unitary U. For positive semidefinite matrices X, Y with $X \leqq Y$, there is a contraction V such that $X^{\frac{1}{2}} = VY^{\frac{1}{2}}$ and V is a convex combination of unitary matrices: $V = \sum_{i=1}^{N} \lambda_i U_i$, where $\lambda_i \geqq 0, U_j$ is unitary $(j = 1, 2, \ldots, N)$ and $\sum_{i=1}^{N} \lambda_i = 1$. Hence, we have

$$\mu\left(X^{\frac{1}{2}}\right) = \mu\left(\left(\sum_{i=1}^{N} \lambda_i U_i\right) Y^{\frac{1}{2}}\right) \leqq \sum_{i=1}^{N} \lambda_i \mu\left(U_i Y^{\frac{1}{2}}\right) = \mu\left(Y^{\frac{1}{2}}\right)$$

since μ is a norm. Therefore, the faithful positive linear functional on $M_n(\mathbb{C})$ defined by

$$\varphi(\cdot) := \mathrm{Tr}(\cdot\, C)$$

satisfies

$$O \leqq X \leqq Y \Longrightarrow \varphi\left(X^{\frac{p}{2}}\right) \leqq \varphi\left(Y^{\frac{p}{2}}\right).$$

Since $\frac{p}{2} > 1$, it follows from Theorem 2.2 that C is a scalar multiple of the identity matrix; the proof is completed. □

3. INEQUALITIES OF NON MATRIX CONVEX FUNCTIONS

Let us start this section with the following observation of tracial properties:

Proposition 3.1. *Let φ be a faithful positive linear functional on $M_n(\mathbb{C})$. Let $p_{m,k}(X,Y)$ be the coefficient of t^k in the polynomial $(X + tY)^m$ for $X, Y \in M_n(\mathbb{C})$, $m \in \mathbb{N}$, and $t \in \mathbb{C}$ and $1 \leqq k \leqq m - 1$. Suppose that*

$$\varphi\left(p_{m,k}\left(X, Y\right)\right) \geqq 0$$

for all $X, Y \geqq O$. Then φ should be a positive scalar multiple of the trace.

Remark that the non-negativity of $\mathrm{Tr}\left(p_{m,k}\left(X,Y\right)\right)$ for all positive matrices X, Y is (equivalent to) the Bessis-Moussa-Villani conjecture (see [9, 7] and also [6]); it is known that it is the case if one of the following is satisfied:

(1) $k \leqq 2$ (or $m \leqq 5$),
(2) $n = 2$ (see Fact 5 in [8]),
(3) $n = 3$, $k = 6$ (see [7]).

Proof. Due to the same argument for $\varphi = \mathrm{Tr}(\cdot D)$ as in the proof of Theorem 2.2, it suffices to consider the case $n = 2$ so that we suppose

$$D = \mathrm{diag}(\varepsilon, 1)$$

for a number ε $(0 < \varepsilon \leqq 1)$. We show that $\varepsilon = 1$.

At first consider the case $k \geqq 2$; let

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad A = \begin{pmatrix} \alpha^2 & \alpha\sqrt{1-\alpha^2} \\ \alpha\sqrt{1-\alpha^2} & 1-\alpha^2 \end{pmatrix},$$
$$B = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$$

for a number α $(0 < \alpha < 1)$. Since $A^2 = A, p_{m,k}(A, B)$ is described as

$$AB^k + B^k A + \text{ the terms including } BAB + \text{the terms } AB^k A.$$

Notice that for a number λ, $A\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}A$ is

$$\begin{pmatrix} \alpha^4 + \alpha^2(1-\alpha^2)\lambda & \alpha^3\sqrt{1-\alpha^2} + \alpha(1-\alpha^2)\sqrt{1-\alpha^2}\lambda \\ \alpha^3\sqrt{1-\alpha^2} + \alpha(1-\alpha^2)\sqrt{1-\alpha^2}\lambda & \alpha^2(1-\alpha^2) + (1-\alpha^2)^2\lambda \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} \alpha^2 & \alpha\lambda\sqrt{1-\alpha^2} \\ \alpha\lambda\sqrt{1-\alpha^2} & \lambda^2(1-\alpha^2) \end{pmatrix},$$

and

$$AB^k = \begin{pmatrix} \alpha^2 & \alpha^{k+1}\sqrt{1-\alpha^2} \\ \alpha\sqrt{1-\alpha^2} & \alpha^k(1-\alpha^2) \end{pmatrix}.$$

Thus, we have for $l \geqq 2$

$$AB^l A = o(\alpha)\ (\alpha \to 0), \qquad BAB = o(\alpha)\ (\alpha \to 0),$$

where o is Landau's small o, and

$$\begin{aligned} 0 &\leqq \varphi(p_{m,k}(UAU^*, UBU^*)) \\ &= \varphi(Up_{m,k}(A, B)U^*) \\ &= \varphi\left(U\{AB^k + B^k A + o(\alpha)\}U^*\right) \\ &= \alpha\varphi\left(U\left\{\frac{1}{\alpha}(AB^k + B^k A) + o(1)\right\}U^*\right) \quad (\alpha \to 0). \end{aligned}$$

Dividing this inequality by $\alpha > 0$ and taking α as $\alpha \to 0$, we have

$$\begin{aligned} 0 &\leqq \varphi\left(U\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}U^*\right) \\ &= \varphi\left(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right) \\ &= \mathrm{Tr}\left(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\varepsilon & 0\\ 0 & 1\end{pmatrix}\right) = \varepsilon - 1, \end{aligned}$$

since

$$\frac{1}{\alpha}AB^k \to \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\,(\alpha \to 0).$$

Hence, $\varepsilon \geqq 1$ or $\varepsilon = 1$.

For the case $k = 1$, let

$$C = \begin{pmatrix}1 & 0\\ 0 & \alpha^2\end{pmatrix}\,(= B^2).$$

Then $p_{m,1}(A, C)$ is $AC + CA +$ the terms ACA. Notice that $ACA = AB^2A = o(\alpha)$ $(\alpha \to 0)$ and the preceding argument for $k \geqq 2$ works similarly. Therefore, the proof is completed. □

Remark 3.2. It follows from the proof that the inequality assumption for $0 \leqq X, Y \leqq 1$ is sufficient for the assertion.

Theorem 3.3. *Let φ be a faithful positive linear functional on $M_n(\mathbb{C})$ and let f be a twice continuously differentiable convex function on $[0, \infty)$ with*

$$f^{[2]}(0, 0, 0) = 0, \quad f^{[2]}(1, 0, 0) > 0,$$

where $f^{[2]}$ is the second divided difference of f. Then

$$\varphi\left(\frac{f(X) + f(Y)}{2}\right) \geqq \varphi\left(f\left(\frac{X+Y}{2}\right)\right) \tag{3.1}$$

holds for all $X, Y \geqq 0$ if and only if φ is a positive scalar multiple of the trace.

Also, $f(t) = t^p$ $(p > 2)$ on $[0, \infty)$ is such an example.

Before giving a proof, let us recall basic facts on matrix convex continuous functions: let f be a matrix convex continuous function of order n on an interval I. Then by definition,

$$\frac{f(X) + f(Y)}{2} \geqq f\left(\frac{X+Y}{2}\right)$$

holds for Hermitian matrices $X, Y \in M_n(\mathbb{C})$ with eigenvalues in I. This yields

$$\varphi\left(\frac{f(X) + f(Y)}{2}\right) \geqq \varphi\left(f\left(\frac{X+Y}{2}\right)\right)$$

for a positive linear functional φ on $M_n(\mathbb{C})$.

We also recall basic facts on Jensen's inequalities: for a convex continuous function f on I and Hermitian matrices $X, Y \in M_n(\mathbb{C})$ with eigenvalues in I,

$$\mathrm{Tr}\left(\frac{f(X) + f(Y)}{2}\right) \geqq \mathrm{Tr}\left(f\left(\frac{X+Y}{2}\right)\right) \tag{3.2}$$

is satisfied; this inequality is well-known, for instance, see [10, Proposition 3.1]: von Neumann observes the convexity $x \mapsto \mathrm{Tr}(f(x))$. E. H. Lieb gives a description of $\mathrm{Tr}(f(x))$ and B. Simon has further arguments. F. Hansen and G.K. Pedersen study generalizations; Jensen's operator inequality and Jensen's trace inequality. There are also many articles on these kinds of inequalities. See the introduction and references in [5] about Jensen's inequalities.

Proof. We have a proof of Jensen's inequality (3.2) for the reader's convenience: Ky Fan's maximum principle (for instance, see [2, page 35]) means that

$$\sum_{i=1}^{k} \lambda_i(X) + \sum_{i=1}^{k} \lambda_i(Y) \geqq \sum_{i=1}^{k} \lambda_i(X+Y)$$

for $k = 1, 2, \ldots, n-1$, and

$$\sum_{i=1}^{n} \lambda_i(X) + \sum_{i=1}^{n} \lambda_i(Y) = \sum_{i=1}^{n} \lambda_i(X+Y),$$

where λ_i is the i-th eigenvalue with $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_n$. In other words, $\left\{\lambda_i\left(\frac{X+Y}{2}\right)\right\}$ is majorized by $\left\{\frac{\lambda_i(X)+\lambda_i(Y)}{2}\right\}$. Then the majorization theory ([2, page 40]) says that

$$\sum_{i=1}^{n} f\left(\frac{\lambda_i(X)+\lambda_i(Y)}{2}\right) \geqq \sum_{i=1}^{n} f\left(\lambda_i\left(\frac{X+Y}{2}\right)\right).$$

Hence, from the convexity of f in the left-hand side, the above inequality (3.2) for Tr and f follows. Therefore, if φ is a positive scalar multiple of the trace, we have the inequality (3.1).

We show the converse: at first we have explicit calculations for the case $f(t) = t^m$ $(m \in \mathbb{N}, m \geqq 3)$ although we have a general treatment below: by assumption,

$$\varphi\left(\frac{X^m + (X+tY)^m}{2} - \left(\frac{X+(X+tY)}{2}\right)^m\right) \geqq 0$$

for $t > 0$ and $X, Y \geqq 0$. Since

$$\frac{X^m + (X+tY)^m}{2} - \left(X + \frac{1}{2}\,tY\right)^m = \frac{1}{4}\,p_{m,2}(X,Y)t^2 + o(t^2)\ (t \to 0),$$

$$0 \leqq t^2\{\varphi(p_{m,2}(X,Y)) + o(1)\} \qquad (t \to 0).$$

Thus, dividing this inequality by $t^2 > 0$ and taking t as $t \to 0$, we have

$$\varphi(p_{m,2}(X,Y)) \geqq 0.$$

Hence, in this case the assertion follows from Proposition 3.1.

Let us consider the general case: since

$$\varphi\left(\frac{f(X) + f(X+tY)}{2} - f\left(\frac{X+(X+tY)}{2}\right)\right) \geqq 0$$

for $t > 0$ and $X, Y \geqq 0$, the preceding argument yields similarly

$$\varphi\left(\left.\frac{d^2}{dt^2}\right|_{t=0} f(X+tY)\right) \geqq 0.$$

For the same matrices A, B, U as in the proof of Proposition 3.1,

$$\frac{1}{2}\frac{d^2}{dt^2}\bigg|_{t=0} f(A+tB)$$

is of the form

$$\begin{aligned}
&f^{[2]}(1,1,1)ABABA + f^{[2]}(1,1,0)ABAB(1-A)\\
&+ f^{[2]}(1,0,1)AB(1-A)BA + f^{[2]}(0,1,1)(1-A)BABA\\
&+ f^{[2]}(1,0,0)AB(1-A)B(1-A) + f^{[2]}(0,0,1)(1-A)B(1-A)BA\\
&+ f^{[2]}(0,1,0)(1-A)BAB(1-A) + f^{[2]}(0,0,0)(1-A)B(1-A)B(1-A)
\end{aligned}$$

(see the formula of the second divided difference in [2, page 129] and remark that $f^{[2]}$ is symmetric and A is an orthogonal projection: $A = 1\cdot A + 0\cdot(1-A)$). The order estimation for BAB, AB^2A and the assumption $f^{[2]}(0,0,0) = 0$ mean

$$\varphi(f^{[2]}(1,0,0)(AB^2 + B^2A) + o(\alpha)) \geqq 0 \quad (\alpha \to 0).$$

Hence, replacing A, B with UAU^*, UBU^*, we get

$$\varphi(U\{f^{[2]}(1,0,0)(AB^2 + B^2A) + o(\alpha)\}U^*) \geqq 0 \quad (\alpha \to 0).$$

Therefore, as in the proof of Proposition 3.1, the proof is completed. □

In the proof of Theorem 3.3, a criterion of non matrix convexity of order 2 is obtained:

Corollary 3.4. *Let f be a function as in Theorem 3.3. Then f is not matrix convex of order 2.*

The same argument works for the following theorem whose proof is left to the reader:

Theorem 3.5. *Let φ be a faithful positive linear functional on $M_n(\mathbb{C})$ and let f be a continuously differentiable increasing function on $[0,\infty)$ with*

$$f^{[1]}(0,0) = 0, \quad f^{[1]}(1,0) > 0,$$

where $f^{[1]}$ is the first divided difference of f. Then

$$\varphi(f(X)) \leqq \varphi(f(Y)) \quad \textit{whenever} \quad O \leqq X \leqq Y$$

if and only if φ is a positive scalar multiple of the trace.

$f(t) = t^p$ $(p > 1)$ on $[0,\infty)$ is such an example.

We remark that a proof can be obtained by the formula of the first divided difference in [1, p. 12] for

$$\frac{d}{dt}\bigg|_{t=0} f(A+tB)$$

as in the proof of Theorem 3.3.

REFERENCES

[1] T. ANDO, *Topics on Operator Inequalities,* Hokkaido University, Sapporo (1978).

[2] R. BHATIA, *Matrix Analysis,* Springer-Verlag (1996).

[3] C. DAVIS, Notions generalizing convexity for functions defined on spaces of matrices, in *Convexity: Proceedings of Symposia in Pure Mathematics*, American Mathematical Society: 187–201 (1963).

[4] W.F. DONOGHUE, Jr., *Monotone Matrix Functions and Analytic Continuation,* Springer-Verlag, (1974).

[5] F. HANSEN AND G.K. PEDERSEN, Jensen's trace inequality in several variables, *Int. J. Math.*, **14** (2003), 667–681.

[6] F. HANSEN, Trace functions as Laplace transforms, preprint (2005).

[7] C. HILLAR AND C.R. JOHNSON, On the positivity of the coefficients of a certain polynomial defined by two positive definite matrices, *J. Stat. Phys.*, **118** (2005), 781–789.

[8] C.R. JOHNSON AND C. HILLAR, Eigenvalues of words in two positive definite letters, *SIAM J. Matrix Anal. Appl.*, **23** (2002), 916–928.

[9] E.H. LIEB AND R. SEIRINGER, Equivalent forms of the Bessis-Moussa-Villani conjecture, *J. Stat. Phys.*, **115** (2004), 185–190.

[10] M. OHYA AND D. PETZ, *Quantum Entropy and Its Use,* Texts and Monographs in Physics, Springer-Verlag (1993).