Tohoku Math. J.
67 (2015), 39–50

ORDER OF OPERATORS DETERMINED BY OPERATOR MEAN

MASARU NAGISA AND MITSURU UCHIYAMA

(Received July 9, 2013, revised December 20, 2013)

Abstract. Let σ be an operator mean and f a non-constant operator monotone function on $(0,\infty)$ associated with σ. If operators A, B satisfy $0 \le A \le B$, then it holds that $Y\sigma(tA+X) \le Y\sigma(tB+X)$ for any non-negative real number t and any positive, invertible operators X, Y. We show that the condition $Y\sigma(tA+X) \le Y\sigma(tB+X)$ for a sufficiently small $t>0$ implies $A \le B$ if and only if X is a positive scalar multiple of Y or the associated operator monotone function f with σ has the form $f(t) = (at+b)/(ct+d)$, where a, b, c, d are real numbers satisfying $ad - bc > 0$.

1. Introduction. We denote the set of all $n \times n$ matrices over $\mathbb{C}$ by M_n and set

$$H_n = \{A \in M_n;\ A^* = A\} \text{ and } H_n^+ = \{A \in H_n;\ A \ge 0\},$$

where $A \ge 0$ means that A is non-negative, that is, the value of inner product

$$(Ax, x) \ge 0 \quad \text{for all } \ x \in \mathbb{C}^n .$$

Let J be an open interval of the set $\mathbb{R}$ of real numbers. We also denote by $H_n(J)$ the set of $A \in H_n$ with its spectrum $\mathrm{Sp}(A) \subset J$. A real continuous function f defined on J is said to be operator monotone (in short, $f \in \mathbb{P}(J)$) if $A \le B$ implies $f(A) \le f(B)$ for any $n \in \mathbb{N}$ and $A, B \in H_n(J)$. In this paper, we assume that an operator monotone function is not a constant function.

By the definition of $f \in \mathbb{P}(J)$, for $A, B \in H_n$, $A \le B$ implies $f(aI+tA) \le f(aI+tB)$ for a sufficiently small $t > 0$, where $a \in J$. Kubo-Ando [5] defined an operator mean σ which is a binary operation $X\sigma Y$ for $X, Y \in H_n(0,\infty)$, and is given by some $f \in \mathbb{P}(0,\infty)$ with $f(1) = 1$ and $f(t) > 0$ as follows:

$$X\sigma Y = X^{1/2} f(X^{-1/2} Y X^{-1/2}) X^{1/2} .$$

In this paper, we consider the problem whether, for $X, Y \in H_n(0,\infty)$ and an operator mean σ, the following condition for $A, B \in H_n$:

$$(*) \qquad Y\sigma(tA+X) \le Y\sigma(tB+X) \quad \text{for a sufficiently small } t > 0$$

(correctly, there exists a positive number ε satisfying that $Y\sigma(tA+X) \le Y\sigma(tB+X)$ holds whenever $0 \le t \le \varepsilon$) implies $A \le B$ or not. Our results are the following:

2010 *Mathematics Subject Classification.* Primary 47A63; Secondary 15A39.

Key words and phrases. Matrix order, operator mean, operator monotone function, Schur product, Fréchet derivative.

This work was supported in part by Grants-in-Aids for Scientific Research (C)22540220, (C)21540181.

(1) When $X = cY$ for some $c > 0$, the condition $(*)$ implies $A \leq B$ for any operator mean σ ([Corollary 4.2]).
(2) When the associated $f \in \mathbb{P}(0,\infty)$ with σ has the form $\frac{at+b}{ct+d}$, the condition $(*)$ implies $A \leq B$ for any $X, Y \in H_n(0,\infty)$ ([Corollary 4.4]).
(3) When the associated $f \in \mathbb{P}(0,\infty)$ with σ has not the form $\frac{at+b}{ct+d}$ and X is not a scalar multiple of Y, there exists A, B, X and Y satisfying

$$A \not\leq B \text{ and } Y\sigma(tA + X) \leq Y\sigma(tB + X)$$

for a sufficiently small $t > 0$ ([Theorem 4.6]).

Combining these facts, we get the following:

THEOREM 1.1. *Let $X, Y \in H_n(0,\infty)$ and σ an operator mean. Then the condition $(*)$ implies $A \leq B$ if and only if X is a positive scalar multiple of Y or the associated operator monotone function f with σ has the form*

$$f(t) = \frac{at+b}{ct+d}, \quad a, b, c, d \in \mathbb{R},\ ad - bc > 0.$$

2. Schur product. For $A \in M_n$, we define the linear map S_A from M_n to M_n as follows:

$$S_A(B) = A \circ B \qquad (B \in M_n)\,,$$

where $A \circ B$ means the Schur product of A and B, that is,

$$A \circ B = (a_{ij}) \circ (b_{ij}) = (a_{ij}b_{ij})\,.$$

When A belongs to H_n^+, it is well known that S_A is completely positive [2], in particular, $S_A(H_n^+) \subset H_n^+$.

We call A of *strict rank* 1 if there exist $\alpha_i, \beta_i \in \mathbb{C} \setminus \{0\}$ $(i = 1, 2, \ldots, n)$ such that

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 & \alpha_1\beta_2 & \cdots & \alpha_1\beta_n \\ \alpha_2\beta_1 & \alpha_2\beta_2 & \cdots & \alpha_2\beta_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n\beta_1 & \alpha_n\beta_2 & \cdots & \alpha_n\beta_n \end{pmatrix}.$$

When $A \in H_n^+$ is of strict rank 1, A is represented as follows:

$$A = \begin{pmatrix} \bar{\gamma}_1 \\ \bar{\gamma}_2 \\ \vdots \\ \bar{\gamma}_n \end{pmatrix} \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{pmatrix},$$

where $\gamma_1, \gamma_2, \ldots, \gamma_n \in \mathbb{C} \setminus \{0\}$. The following statement plays an important role in this paper.

PROPOSITION 2.1. *For $A = (a_{ij}) \in H_n^+$, the following are equivalent:*
(1) *For $B \in H_n$, $S_A(B) \geq 0 \Rightarrow B \geq 0$.*
(2) *A is of strict rank 1.*
(3) $S_A(H_n^+) = H_n^+$.

(4) $a_{kk} > 0$ *and* $a_{kk}a_{ll} = a_{kl}a_{lk} (= |a_{kl}|^2)$ *for each* k, l.

PROOF. (1) ⇒ (4) We assume $a_{kk} = 0$. Set $B = (b_{ij})$ as follows:

$$b_{ij} = \begin{cases} -1 & \text{if } (i, j) = (k, k) \\ 0 & \text{otherwise} \end{cases} .$$

Since $B \not\geq 0$ and $S_A(B) = A \circ B = 0 \geq 0$, this contradicts the assumption. So $a_{kk} > 0$ for all k.

By the positivity of A, we have

$$\begin{pmatrix} a_{kk} & a_{kl} \\ a_{lk} & a_{ll} \end{pmatrix} \geq 0 ,$$

in particular, $a_{kk}a_{ll} - a_{kl}a_{lk} \geq 0$. We assume that $a_{kk}a_{ll} - a_{kl}a_{lk} > 0$. Then we set $B = (b_{ij})$ as follows:

$$b_{ij} = \begin{cases} \dfrac{|a_{kl}|}{a_{kk}} & \text{if } (i, j) = (k, k) \\ \dfrac{|a_{kl}|}{a_{ll}} & \text{if } (i, j) = (l, l) \\ 1 & \text{if } (i, j) = (k, l) \text{ or } (l, k) \\ 0 & \text{otherwise} \end{cases} .$$

Since $|a_{kl}|^2 = a_{kl}a_{lk} < a_{kk}a_{ll}$, we have $B \not\geq 0$. But we have

$$(A \circ B)_{ij} = \begin{cases} |a_{kl}| & \text{if } (i, j) = (k, k) \text{ or } (l, l) \\ a_{kl} & \text{if } (i, j) = (k, l) \\ a_{lk} & \text{if } (i, j) = (l, k) \\ 0 & \text{otherwise} \end{cases} ,$$

and $A \circ B \geq 0$. This contradicts the assumption. So we have $a_{kk}a_{ll} - a_{kl}a_{lk} = 0$ for each k, l.

(4) ⇒ (2) We set $r_k = \sqrt{a_{kk}} > 0$ $(k = 1, 2, \ldots, n)$. By the assumption, for each k and l, we can choose $\theta(k, l) \in \mathbb{R}$ such that

$$a_{kl} = r_k r_l e^{i\theta(k,l)} .$$

Then we also have $e^{i\theta(k,l)} = e^{-i\theta(l,k)}$ and $e^{i\theta(k,k)} = 1$. If we show the relation

$$e^{i\theta(k,l)} e^{i\theta(l,m)} = e^{i\theta(k,m)}$$

for any k, l and m, then we can see that A is of strict rank 1 as follows:

$$\begin{pmatrix} r_1 \\ r_2 e^{-i\theta(1,2)} \\ \vdots \\ r_n e^{-i\theta(1,n)} \end{pmatrix} \begin{pmatrix} r_1 & r_2 e^{i\theta(1,2)} & \cdots & r_n e^{i\theta(1,n)} \end{pmatrix}$$

$$= \begin{pmatrix} r_1 \\ r_2 e^{i\theta(2,1)} \\ \vdots \\ r_n e^{i\theta(n,1)} \end{pmatrix} \begin{pmatrix} r_1 & r_2 e^{i\theta(1,2)} & \cdots & r_n e^{i\theta(1,n)} \end{pmatrix}$$

$$= \begin{pmatrix} r_1^2 & r_1 r_2 e^{i\theta(1,2)} & \cdots & r_1 r_n e^{i\theta(1,n)} \\ r_2 r_1 e^{i\theta(2,1)} & r_2^2 e^{i\theta(2,1)} e^{i\theta(1,2)} & \cdots & r_2 r_n e^{i\theta(2,1)} e^{i\theta(1,n)} \\ \vdots & \vdots & \ddots & \vdots \\ r_n r_1 e^{i\theta(n,1)} & r_n r_2 e^{i\theta(n,1)} e^{i\theta(1,2)} & \cdots & r_n^2 e^{i\theta(n,1)} e^{i\theta(1,n)} \end{pmatrix}$$

$$= \begin{pmatrix} r_1^2 & r_1 r_2 e^{i\theta(1,2)} & \cdots & r_1 r_n e^{i\theta(1,n)} \\ r_2 r_1 e^{i\theta(2,1)} & r_2^2 & \cdots & r_2 r_n e^{i\theta(2,n)} \\ \vdots & \vdots & \ddots & \vdots \\ r_n r_1 e^{i\theta(n,1)} & r_n r_2 e^{i\theta(n,2)} & \cdots & r_n^2 \end{pmatrix} = A\,.$$

It suffices to show the relation $e^{i\theta(k,l)} e^{i\theta(l,m)} = e^{i\theta(k,m)}$ in the case of each two of k, l, m are different. By the positivity of A, we have

$$\begin{pmatrix} a_{kk} & a_{kl} & a_{km} \\ a_{lk} & a_{ll} & a_{lm} \\ a_{mk} & a_{ml} & a_{mm} \end{pmatrix} \geq 0\,.$$

Since

$$\begin{pmatrix} a_{kk} & a_{kl} & a_{km} \\ a_{lk} & a_{ll} & a_{lm} \\ a_{mk} & a_{ml} & a_{mm} \end{pmatrix} = \begin{pmatrix} r_k^2 & r_k r_l e^{i\theta(k,l)} & r_k r_m e^{i\theta(k,m)} \\ r_l r_k e^{i\theta(l,k)} & r_l^2 & r_l r_m e^{i\theta(l,m)} \\ r_m r_k e^{i\theta(m,k)} & r_m r_l e^{i\theta(m,l)} & r_m^2 \end{pmatrix}$$

$$= \begin{pmatrix} r_k e^{i\theta(k,l)} & & \\ & r_l & \\ & & r_m e^{i\theta(m,l)} \end{pmatrix} \begin{pmatrix} 1 & 1 & \alpha \\ 1 & 1 & 1 \\ \bar{\alpha} & 1 & 1 \end{pmatrix} \begin{pmatrix} r_k e^{i\theta(l,k)} & & \\ & r_l & \\ & & r_m e^{i\theta(l,m)} \end{pmatrix}$$

and

$$\alpha = e^{-i\theta(k,l)} e^{-i\theta(l,m)} e^{i\theta(k,m)}\,,$$

we have

$$\begin{pmatrix} 1 & 1 & \alpha \\ 1 & 1 & 1 \\ \bar{\alpha} & 1 & 1 \end{pmatrix} \geq 0\,.$$

Then $|\alpha| = 1$ and

$$0 \leq \det \begin{pmatrix} 1 & 1 & \alpha \\ 1 & 1 & 1 \\ \bar{\alpha} & 1 & 1 \end{pmatrix} = \alpha + \bar{\alpha} - 2$$

imply $\alpha = 1$. So the desired relation is established.

(2) $\Rightarrow$ (3) When A is of strict rank 1, we construct $A' \in H_n^+$ as follows:

$$A = \begin{pmatrix} \bar{\gamma_1} \\ \bar{\gamma_2} \\ \vdots \\ \bar{\gamma_n} \end{pmatrix} \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{pmatrix}, \quad A' = \begin{pmatrix} 1/\bar{\gamma_1} \\ 1/\bar{\gamma_2} \\ \vdots \\ 1/\bar{\gamma_n} \end{pmatrix} \begin{pmatrix} 1/\gamma_1 & 1/\gamma_2 & \cdots & 1/\gamma_n \end{pmatrix}.$$

Since $S_A(H_n^+) \subset H_n^+$, $S_{A'}(H_n^+) \subset H_n^+$ and $S_A S_{A'} = \mathrm{id} = S_{A'} S_A$, we have that $S_A(H_n^+) = H_n^+$.

(3) $\Rightarrow$ (1) When some (i, j)-th component of A is equal to 0, the (i, j)-th component of $S_A(B)$ is also 0. So we have all components of A are not equal to 0. This means that S_A is injective. If $S_A(B) \geq 0$, then we can choose $C \geq 0$ with $S_A(B) = S_A(C)$ by the assumption (3). The injectivity of S_A implies $B = C \geq 0$. □

REMARK 2.2. When $A = A^* \in H_n$, $S_A(H_n) \subset H_n$. Moreover, if $A \in H_n$ is of strict rank 1, then $S_A(H_n) = H_n$ (i.e., S_A becomes a bijection of H_n). But there is $A \in H_n$ which is not of strict rank 1 and satisfies $S_A(H_n) = H_n$, for example, $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) \in H_2$. In the case of $A \geq 0$, $S_A(H_n^+) \subset H_n^+$ and the fact $S_A(H_n^+) = H_n^+$ is equivalent that A is of strict rank 1.

3. Fréchet derivatives of operator monotone functions. Let f be an operator monotone function on the open interval J and $A \in H_n(J) = \{A \in H_n; \mathrm{Sp}(A) \subset J\}$. We can choose a positive real number ε such that $\mathrm{Sp}(A + H) \subset J$ whenever $H \in H_n$ and $\|H\| < \varepsilon$. So f is a map from an open neighborhood of A to the real vector space H_n and there exists a linear map T from H_n to H_n satisfying

$$\lim_{\|H\| \to 0} \frac{\|f(A + H) - f(A) - T(H)\|}{\|H\|} = 0.$$

We call this T the (Fréchet) derivative of f at A and denote it by $Df(A)$. For every $B \in H_n$, we denote by $Df(A)(B)$ the directional derivative of f at A in the direction B, i.e.,

$$Df(A)(B) = \left.\frac{d}{dt}\right|_{t=0} f(A + tB).$$

For a diagonal matrix $\Lambda \in H_n(J)$, we can compute as follows (see [1], [2], [4]):

$$Df(\Lambda)(B) = f^{[1]}(\Lambda) \circ B,$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad f^{[1]}(\Lambda) = \begin{pmatrix} f^{[1]}(\lambda_1, \lambda_1) & \cdots & f^{[1]}(\lambda_1, \lambda_n) \\ \vdots & \ddots & \vdots \\ f^{[1]}(\lambda_n, \lambda_1) & \cdots & f^{[1]}(\lambda_n, \lambda_n) \end{pmatrix}$$

and

$$f^{[1]}(\lambda, \mu) = \begin{cases} \dfrac{f(\lambda) - f(\mu)}{\lambda - \mu} & \text{if } \lambda \neq \mu \\ f'(\lambda) & \text{if } \lambda = \mu \end{cases}.$$

The Loewner matrix $f^{[1]}(\Lambda)$ is non-negative since f is operator monotone on J. For a general $A \in H_n(J)$, we choose a unitary $U \in M_n$ and a diagonal matrix $\Lambda \in H_n^+$ satisfying $A = U\Lambda U^*$. Then the directional derivatives of f can be represented as follows:

$$Df(A)(B) = U(f^{[1]}(\Lambda) \circ (U^*BU))U^* .$$

If $C \in H_n(J)$, $A, B \in H_n$ and $C + A, C + B \in H_n(J)$, then we have

$$A \le B \Rightarrow f(C + A) \le f(C + B) .$$

We also have

$$f(C + tA) \le f(C + tB) \quad \text{for all } t \in [0, 1] .$$

Since

$$\frac{f(C + tA) - f(C)}{t} \le \frac{f(CI + tB) - f(C)}{t} \quad (0 < t < 1) ,$$

it holds $Df(C)(A) \le Df(C)(B)$ by tending t to 0. So we have implications as follows:

$$\begin{aligned} A \le B &\Rightarrow f(C + tA) \le f(C + tB) \quad (0 \le t \le 1) \\ &\Rightarrow Df(C)(A) \le Df(C)(B) . \end{aligned}$$

We will consider the problem when the reverse implication holds.

THEOREM 3.1. *Let f be an operator monotone function on J and $C \in H_n(J)$. Then the following are equivalent:*

(1) *For $A, B \in H_n$, $Df(C)(A) \le Df(C)(B)$ implies $A \le B$.*
(2) *$f^{[1]}(\Lambda)$ is of strict rank 1, where Λ is a diagonal matrix represented by U^*CU for some unitary U.*
(3) *$Df(C)(H_n^+) = H_n^+$.*

PROOF. First we assume that C is diagonal (i.e., $C = \Lambda$). Then we have

$$Df(C)(A) = Df(\Lambda)(A) = S_L(A) ,$$

where $L = f^{[1]}(\Lambda)$. In this case, we have (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) by Proposition 2.1.

When C has the form $U\Lambda U^*$ (Λ is diagonal and U is unitary), we remark that

$$Df(C)(A) = U(f^{[1]}(\Lambda) \circ (U^*AU))U^* = U(S_L(U^*AU))U^* .$$

Since U is unitary, we have the implication (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3). □

When C is a scalar operator, we can get the following equivalent conditions and the equivalence of (1) and (2) has been proved in [6]: Theorem 2.2.

THEOREM 3.2. *Let f be operator monotone on the open interval J, $A, B \in H_n$ and $c \in J$. Then the following are equivalent:*

(1) *$A \le B$.*
(2) *There exists a sequence $\{t_n\}_{n=1}^{\infty}$ satisfying that*

$$t_n > 0,\ \lim_{n\to\infty} t_n = 0,\ f(cI + t_nA) \le f(cI + t_nB) .$$

(3) *$Df(cI)(A) \le Df(cI)(B)$.*

PROOF. We have already shown (1) $\Rightarrow$ (2) $\Rightarrow$ (3), considering cI as C.

(3) $\Rightarrow$ (1) Since $f \in \mathbb{P}(J)$, $f'(c) > 0$ for $c \in J$. Then we have

$$f^{[1]}(cI) = \begin{pmatrix} f'(c) & \cdots & f'(c) \\ \vdots & \ddots & \vdots \\ f'(c) & \cdots & f'(c) \end{pmatrix} = \begin{pmatrix} \sqrt{f'(c)} \\ \vdots \\ \sqrt{f'(c)} \end{pmatrix} \left(\sqrt{f'(c)} \quad \cdots \quad \sqrt{f'(c)} \right),$$

that is, $f^{[1]}(cI)$ is of strict rank 1. By Theorem 3.1, (3) implies (1). □

4. Operator means and main results. In this section, we only consider the interval J as $(0, \infty)$. Let f be an operator monotone function on J with $f(1) = 1$ and $f(t) > 0$ and σ_f the operator mean associted with f. When $X, Y \in H_n(J)$ and $A, B \in H_n^+$ with $A \leq B$, we have

$$Y\sigma_f(tA + X) \leq Y\sigma_f(tB + X) \quad \text{for all } t \geq 0\,.$$

Our problem is to decide when the condition

$$(*) \qquad Y\sigma_f(tA + X) \leq Y\sigma_f(tB + X) \quad \text{for a sufficiently small } t > 0$$

implies $A \leq B$.

PROPOSITION 4.1. *Let $X, Y \in H_n(J)$ and $A, B \in H_n^+$. If $Df(Y^{-1/2}XY^{-1/2})(H_n^+) = H_n^+$, then the condition $(*)$ implies $A \leq B$.*

PROOF. By the definition of the mean σ_f, we have the following:

$$\begin{aligned}
&Y\sigma_f(tA + X) \leq Y\sigma_f(tB + X) \\
&\Rightarrow f(Y^{-1/2}(tA + X)Y^{-1/2}) \leq f(Y^{-1/2}(tB + X)Y^{-1/2}) \\
&\Rightarrow \lim_{t\to 0+} \frac{f(Y^{-1/2}(tA + X)Y^{-1/2}) - f(Y^{-1/2}XY^{-1/2})}{t} \\
&\qquad\qquad \leq \lim_{t\to 0+} \frac{f(Y^{-1/2}(tB + X)Y^{-1/2}) - f(Y^{-1/2}XY^{-1/2})}{t} \\
&\Rightarrow Df(Y^{-1/2}XY^{-1/2})(Y^{-1/2}AY^{-1/2}) \leq Df(Y^{-1/2}XY^{-1/2})(Y^{-1/2}BY^{-1/2})\,.
\end{aligned}$$

Since $Df(Y^{-1/2}XY^{-1/2})(H_n^+) = H_n^+$, we have $Y^{-1/2}AY^{-1/2} \leq Y^{-1/2}BY^{-1/2}$ by Theorem 3.1. This implies $A \leq B$. □

In the proof of Theorem 3.2, we have already shown that $Df(cI)(H^+) = H^+$ for any $c > 0$. Noticing the fact $Df(cI) = Df(cX^{-1/2}XX^{-1/2})$, we have the following:

COROLLARY 4.2. *Let $X \in H_n(J)$ and $A, B \in H_n^+$. For any operator mean σ_f and $c > 0$, the condition $(*)$ implies $A \leq B$.*

It is well known that the function

$$f(t) = \frac{at + b}{ct + d}$$

is operator monotone on $(-\infty, -d/c)$ and on $(-d/c, \infty)$ if $ad - bc > 0$. So f is operator monotone on $J = (0, \infty)$ whenever $ad - bc > 0$ and $cd \geq 0$.

LEMMA 4.3. *Let* $f(t) = \frac{at+b}{ct+d}$ *with* $ad - bc > 0$ *and* $cd \geq 0$. *Then we have* $Df(X)(H_n^+) = H_n^+$ *for any* $X \in H_n(J)$.

PROOF. Since f is operator monotone on J, we may show that $f^{[1]}(\Lambda)$ is of strict rank 1 for any diagonal matrix $\Lambda \in H_n(J)$. By the argument of Proposition 2.1 (1) $\Rightarrow$ (2), it is sufficient to show that, for any $\lambda, \mu \in J$ ($\lambda \neq \mu$),

$$\begin{pmatrix} f'(\lambda) & f^{[1]}(\lambda, \mu) \\ f^{[1]}(\mu, \lambda) & f'(\mu) \end{pmatrix}$$

is of strict rank 1, that is,

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

By the calculation

$$f'(\lambda) = \frac{ad - bc}{(c\lambda + d)^2}\,, \quad f'(\mu) = \frac{ad - bc}{(c\mu + d)^2}$$

and

$$\begin{aligned} f^{[1]}(\lambda, \mu) &= \frac{f(\lambda) - f(\mu)}{\lambda - \mu} \\ &= \frac{ad - bc}{(c\lambda + d)(c\mu + d)} = (f'(\lambda)f'(\mu))^{1/2}\,, \end{aligned}$$

we get $f'(\lambda)f'(\mu) = f^{[1]}(\lambda, \mu)^2$. So we have done. □

Combining Proposition 4.1 and Lemma 4.3, we can show the following:

COROLLARY 4.4. *Let* $f(t) = \frac{at+b}{ct+d}$ *with* $ad - bc > 0$, $cd \geq 0$ *and* $f(1) = 1$, $X, Y \in H_n(J)$ *and* $A, B \in H_n^+$. *Then the condition* $(*)$ *implies* $A \leq B$.

In the above statements, we have already used the fact $Df(X)(H_n^+) = H_n^+$ as a key tool to get the relation $A \leq B$. But, in many cases, we can not expect that $Df(X)(H_n^+) = H_n^+$ holds.

PROPOSITION 4.5. *Let* f *be operator monotone on* J *and* $X \in H_n(J)$ *not a scalar matrix. If* $Df(X)(H_n^+) = H_n^+$, *then the function* f *is of the form* $\frac{at+b}{ct+d}$.

PROOF. Since f is operator monotone on J, f can be represented as follows:

$$f(\zeta) = \alpha + \beta\zeta + \int_{-\infty}^{0} g(s, \zeta)d\nu(s)\,, \quad (\mathrm{Im}\,\zeta > 0)$$

where $\alpha \in \mathbb{R}$, $\beta \geq 0$, $g(s, \zeta) = \frac{1+s\zeta}{s-\zeta}$ and ν is a positive finite measure on $J^c = (-\infty, 0]$ ([1], [3], [4]). When $\mathrm{Re}\,\zeta > 0$, the function $g(\cdot, \zeta) \in L^\infty(J^c, \nu)$ and

$$|g(s, \zeta)| \leq |\zeta| + \left|\frac{\zeta^2 + 1}{s - \zeta}\right| \leq |\zeta| + \frac{|\zeta|^2 + 1}{|\zeta|}\,.$$

Using the dominated convergence theorem, for $0 < \lambda < \mu$, we can get the following:

$$f'(\lambda) = \beta + \int_{-\infty}^{0} g_t(s, \lambda)d\nu(s) = \beta + \int_{-\infty}^{0} \frac{s^2 + 1}{(s - \lambda)^2}d\nu(s)\,,$$

$$f^{[1]}(\lambda,\mu)=\frac{f(\lambda)-f(\mu)}{\lambda-\mu}=\beta+\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)(s-\mu)}d\nu(s)\,.$$

When $X\in H_n(J)$ is not a scalar operator and

$$Df(X)(H_n^+)=H_n^+\,,$$

we show that f has the form $\frac{at+b}{ct+d}$. We may assume that $\lambda,\mu\in \mathrm{Sp}(X)$, $\lambda\neq\mu$ and

$$\begin{pmatrix} f'(\lambda) & f^{[1]}(\lambda,\mu)\\ f^{[1]}(\mu,\lambda) & f'(\mu)\end{pmatrix}$$

is of strict rank 1. The rank of this matrix is 1 implies that f is rational and of degree 1 by Theorem III ([3], page 38).

For convenience of the reader, we prove this statement in [3], here. By above calculation

$$\begin{aligned} f'(\lambda)f'(\mu)-(f^{[1]}(\lambda,\mu))^2 &=\Big(\beta+\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)^2}d\nu(s)\Big)\Big(\beta+\int_{-\infty}^{0}\frac{s^2+1}{(s-\mu)^2}d\nu(s)\Big)\\ &\quad-\Big(\beta+\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)(s-\mu)}d\nu(s)\Big)^2\\ =\beta\int_{-\infty}^{0}\Big(\frac{\sqrt{s^2+1}}{s-\lambda}-\frac{\sqrt{s^2+1}}{s-\mu}\Big)^2 d\nu(s)&+\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)^2}d\nu(s)\Big)\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\mu)^2}d\nu(s)\Big)\\ &\quad-\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)(s-\mu)}d\nu(s)\Big)^2\\ =0\,. \end{aligned}$$

Applying the Cauchy-Schwarz inequality

$$\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)^2}d\nu(s)\Big)\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\mu)^2}d\nu(s)\Big)\geq\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)(s-\mu)}d\nu(s)\Big)^2$$

to above identity, we have

$$\beta\int_{-\infty}^{0}\Big(\frac{\sqrt{s^2+1}}{s-\lambda}-\frac{\sqrt{s^2+1}}{s-\mu}\Big)^2 d\nu(s)=0$$

and

$$\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)^2}d\nu(s)\Big)\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\mu)^2}d\nu(s)\Big)=\Big(\int_{-\infty}^{0}\frac{s^2+1}{(s-\lambda)(s-\mu)}d\nu(s)\Big)^2.$$

Since the equality holds for the inequality, the functions

$$s\to\frac{\sqrt{s^2+1}}{s-\lambda}\,,\quad s\to\frac{\sqrt{s^2+1}}{s-\mu}$$

are linearly dependent in $L^2(J^c,\nu)$ (i.e., the measure ν is concentrated on one point λ_0) and $\beta=0$. So we have

$$f(t)=\alpha+\nu(\{\lambda_0\})\frac{1+\lambda_0 t}{\lambda_0-t},$$

that is, f has the desired form. □

THEOREM 4.6. *Let f be operator monotone on J with $f(1) = 1$ and not of the form $\frac{at+b}{ct+d}$. If $X, Y \in H_n(J)$ and X is not a scalar multiple of Y, then there exist $A, B \geq 0$ such that $A \not\leq B$ and*

$$Y\sigma_f(tA + X) \leq Y\sigma_f(tB + X)$$

for a sufficiently small $t > 0$.

PROOF. Since X is not a scalar multiple of Y, we may choose a unitary U and a diagonal matrix Λ such that

$$Y^{-1/2}XY^{-1/2} = U\Lambda U^*\,,$$

and Λ has the form

$$\begin{pmatrix} \lambda & & \\ & \mu & \\ & & \ddots \end{pmatrix}, \quad \lambda, \mu > 0 \quad \text{and} \quad \lambda \neq \mu\,.$$

By the assumption for f and Proposition 2.1 and Proposition 4.5, $f^{[1]}(\begin{pmatrix} \lambda & \\ & \mu \end{pmatrix})$ is not of (strict) rank 1, i.e., $f'(\lambda)f'(\mu) > f^{[1]}(\lambda, \mu)^2$. So we choose $H \in H_n$ such that

$$H = \begin{pmatrix} h_{11} & h_{12} & 0 & \cdots & 0 \\ h_{21} & h_{22} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \not\geq 0\,, \quad Df(\Lambda)(H) \geq 0\,.$$

To restrict the argument to the part of M_2, we set

$$H' = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \not\geq 0\,, \quad \Lambda' = \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} > 0\,.$$

Since $f'(\lambda)$, $f'(\mu)$, $f^{[1]}(\lambda, \mu) > 0$ and

$$Df(\Lambda')(H') = \begin{pmatrix} f'(\lambda)h_{11} & f^{[1]}(\lambda, \mu)h_{12} \\ f^{[1]}(\mu, \lambda)h_{21} & f'(\mu)h_{22} \end{pmatrix} \geq 0\,,$$

we have $h_{11}, h_{22} > 0$ and may assume that

$$h_{11}h_{22} < |h_{12}|^2 < \frac{f'(\lambda)f'(\mu)}{f^{[1]}(\lambda, \mu)^2}h_{11}h_{22}\,,$$

in particular, $Df(\Lambda')(H') > 0$. Put $A', B' \geq 0$ as follows:

$$A' = \begin{pmatrix} 0 & 0 \\ 0 & |h_{12}|^2/h_{11} - h_{22} \end{pmatrix}, \quad B' = H' + A' = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & |h_{12}|^2/h_{11} \end{pmatrix}.$$

Since

$$\begin{aligned}0 < Df(\Lambda')(H') &= Df(\Lambda')(B') - Df(\Lambda')(A') \\ &= \lim_{t\to 0}\left(\frac{f(tB'+\Lambda')-f(\Lambda')}{t} - \frac{f(tA'+\Lambda')-f(\Lambda')}{t}\right) \\ &= \lim_{t\to 0}\frac{f(tB'+\Lambda')-f(tA'+\Lambda')}{t},\end{aligned}$$

we have

$$f(tB'+\Lambda') - f(tA'+\Lambda') \geq 0$$

for a sufficiently small $t > 0$.

Put

$$\tilde{A} = \begin{pmatrix} A' & 0 \\ 0 & 0 \end{pmatrix} \in M_n\,, \quad \tilde{B} = \begin{pmatrix} B' & 0 \\ 0 & 0 \end{pmatrix} \in M_n$$

and

$$A = Y^{1/2}U\tilde{A}U^*Y^{1/2}\,, \quad B = Y^{1/2}U\tilde{B}U^*Y^{1/2}\,.$$

Then $A \not\leq B$ because $A' \not\leq B'$. Since

$$\begin{aligned}&U^*Y^{-1/2}\left(Y\sigma_f(tB+X) - Y\sigma_f(tA+X)\right)Y^{-1/2}U \\ &=U^*f(Y^{-1/2}(tB+X)Y^{-1/2})U - U^*f(Y^{-1/2}(tA+X)Y^{-1/2})U \\ &=U^*f(tU\tilde{B}U^* + Y^{-1/2}XY^{-1/2})U - U^*f(tU\tilde{A}U^* + Y^{-1/2}XY^{-1/2})U \\ &=U^*f(U(t\tilde{B}+\Lambda)U^*)U - U^*f(U(t\tilde{A}+\Lambda)U^*)U \\ &=f(t\tilde{B}+\Lambda) - f(t\tilde{A}+\Lambda) \\ &=\begin{pmatrix} f(tB'+\Lambda') - f(tA'+\Lambda') & 0 \\ 0 & 0 \end{pmatrix} \geq 0\,,\end{aligned}$$

we have

$$Y\sigma_f(tA+X) \leq Y\sigma_f(tB+X)$$

for a sufficiently small $t > 0$. □

Acknowledgement. The authors would like to thank the referee for his useful advice. Following his advice, they added the equivalent assertion (4) in Proposition 2.1.

References

[1] R. Bhatia, Matrix analysis, Grad. Texts in Math. 169, Springer-Verlag, New York, 1997.

[2] R. Bhatia, Positive definite matrices, Princeton Ser. Appl. Math., Princeton University Press, Princeton, NJ, 2007.

[3] W. F. Donoghue, Jr., Monotone matrix functions and analytic continuation, Springer-Verlag, New York-Heidelberg, 1974.

[4] F. Hiai, Matrix analysis: matrix monotone functions, matrix means, and majorization, Interdecip. Inform. Sci. 16 (2010), 139–248.

[5] F. Kubo and T. Ando, Means of positive linear operators, Math. Ann. 246(1980), 205–224.

[6] M. UCHIYAMA, A converse of Loewner-Heinz inequality, geometric mean and spectral order, Proc. Edinb. Math. Soc. (2) 57 (2014), 565–571.

DEPARTMENT OF MATHEMATICS AND INFORMATICS
GRADUATE SCHOOL OF SCIENCE
CHIBA UNIVERSITY
CHIBA, 263–8522
JAPAN

E-mail address: nagisa@math.s.chiba-u.ac.jp

RITSUMEIKAN UNIVERSITY
RESEARCH OFFICE (BKC)
NOJI HIGASHI 1 CHOME, 1–1 KUSATSU,
SHIGA 525–8577
JAPAN

E-mail address: uchiyama@fc.ritsumei.ac.jp