Linear and Multilinear Algebra,
60(2012), 255–265.

Characterization of mean transformations

Ádám Besenyei[1] **and Dénes Petz**[2]

Department of Applied Analysis
ELTE, H-1117 Budapest, Pázmány P. sétány 1/c, Hungary

and

Alfréd Rényi Institute of Mathematics
H-1364 Budapest, POB 127, Hungary

Abstract

The mean transformations $M(A,B)$ are linear mappings and they are analogues of the matrix means of $A,B \ge 0$. They are defined by operator monotone functions. In this paper several properties are described and a part of them characterize the concept.

2000 Mathematics Subject Classification. Primary 26E60; Secondary 15A45.

Keywords: Matrix monotone function, matrix means, transformer inequality, trace-preserving mappings.

1 Introduction

Let $m(x,y)$ be a mean of positive numbers and $\mathbf{M}_n$ be the algebra of $n \times n$ complex matrices. The mean of numbers can be extended to matrices if $f(x) = m(1,x)$ is a matrix monotone function. If $0 < A, B \in \mathbf{M}_n$, then

$$0 < m(A,B) = A^{1/2} f(A^{-1/2}BA^{-1/2})A^{1/2} \in \mathbf{M}_n$$

is defined in the paper of Kubo and Ando [7]. The following conditions give an axiomatic approach:

[1]E-mail: badam@cs.elte.hu
[2]E-mail: petz@math.bme.hu

(1) $m(A,A) = A$ for every A,

(2) $m(A,B) = m(B,A)$ for every A and B (symmetry condition),

(3) if $A \le A'$ and $B \le B'$, then $m(A,B) \le m(A',B')$ (joint monotonicity),

(4) m is continuous,

(5) $C\, m(A,B)\, C^* \le m\big(CAC^*, CBC^*\big)$ (transformer inequality).

Condition (1) is equivalent to $f(1) = 1$ and the symmetry condition (2) requires that $xf(x^{-1}) = f(x)$. Matrix monotone functions f with these two properties will be called standard matrix monotone.

Another concept was introduced by Hiai and Kosaki in [5], this is called here mean transformation. $M(A,B)$ is a linear mapping $\mathbf{M}_n \to \mathbf{M}_n$. If $A = \sum_i \lambda_i |x_i\rangle\langle x_i|$ and $B = \sum_j \mu_j |y_j\rangle\langle y_j|$, then

$$M(A,B)|x_i\rangle\langle y_j| = m(\lambda_i, \mu_j)|x_i\rangle\langle y_j|.$$

The operator $M(A,B)$ has positive eigenvalues and orthogonal eigenvectors, so it is a positive operator.

If $m(x,y) = f(x/y)y$ with a function $f : \mathbb{R}^+ \to \mathbb{R}^+$, then we have the equivalent formulation

$$M_f(A,B) = f(\mathbb{L}_A \mathbb{R}_B^{-1})\mathbb{R}_B, \tag{1}$$

where $\mathbb{L}_A H = AH$ and $\mathbb{R}_B H = HB$. Typically f is an operator monotone function and $f(1) = 1$. An important example is $f(x) = (x-1)/\log x$, then

$$M_f(A,B)X = \int_0^1 A^t X B^{1-t}\, dt.$$

For the matrix mean $m(A,A) = A$ and there is nothing to study, but for the mean transformation $M(A,A)$ has non-trivial properties. In the study of abstract quantum Fisher information, the inner product

$$\gamma_D(X,X) = \langle X, M(D,D)^{-1}X\rangle$$

is used, where D is a so-called density matrix ($D > 0$ and $\mathrm{Tr}\, D = 1$), $X = X^*$ and $\langle \cdot, \cdot \rangle$ denotes the Hilbert-Schmidt inner product [9, 10]. (The notation γ_D was motivated by the Riemannian geometry.) A kind of monotonicity under trace-preserving mapping was essential there and this kind of monotonicity will be crucial also in this paper in the characterization of the mean transformation. We note that in the paper [8] the normalization $\mathrm{Tr}\, D = 1$ was skipped and the inverse of the mean transformation $M(D,D)$ was extended.

The main subject of the present paper is the general mean transformation $M(A,B)$. The properties of $M(A,B)$ are rather similar to those of the matrix means. The transformer inequality (5) of the matrix mean has an analogue, a kind of monotonicity property for trace-preserving completely positive mappings. This is essential in the characterization as much as a block matrix formula.

2 Properties of mean transformations

The mean transformation is already defined in the introduction, see (1). First its transformer inequality is discussed.

Lemma 1 *Let $M(A,B) : \mathbf{M}_n \to \mathbf{M}_n$ be a positive linear mapping and $\beta\colon \mathbf{M}_n \to \mathbf{M}_k$ be a linear mapping. Then the inequality*

$$\beta M(A,B)\beta^* \le M(\beta(A),\beta(B)). \tag{2}$$

is equivalent to

$$\beta^* M(\beta(A),\beta(B))^{-1}\beta \le M(A,B)^{-1}. \tag{3}$$

Proof: Clearly, (2) is equivalent to

$$\|M(A,B)^{1/2}\beta^* M(\beta(A),\beta(B))^{-1/2}\| \le 1$$

which holds if and only if

$$\|M(\beta(A),\beta(B))^{-1/2}\beta M(A,B)^{1/2}\| \le 1$$

and this gives the desired inequality (3). □

Theorem 1 *Let $f:\mathbb{R}^+ \to \mathbb{R}^+$ be an operator monotone function and $M(\,\cdot\,,\,\cdot\,)$ be the corresponding mean transformation. If $\beta : \mathbf{M}_n \to \mathbf{M}_k$ is a 2-positive trace-preserving mapping and the matrices $A, B \in \mathbf{M}_n$, $\beta(A), \beta(B) \in \mathbf{M}_k$ are positive definite, then*

$$\beta M(A,B)\beta^* \le M(\beta(A),\beta(B)). \tag{4}$$

Proof: Based on the Löwner theorem [2], we may consider $f(x) = x/(\lambda + x)$ $(\lambda > 0)$. Then

$$M(A,B) = \frac{\mathbb{L}_A}{\lambda I + \mathbb{L}_A \mathbb{R}_B^{-1}}, \qquad M(A,B)^{-1} = (\lambda I + \mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{L}_A^{-1}.$$

By Lemma 1 the statement (4) has the equivalent form (3) which means

$$\langle \beta(X), (\lambda I + \mathbb{L}_{\beta(A)}\mathbb{R}_{\beta(B)}^{-1})\mathbb{L}_{\beta(A)}^{-1}\beta(X)\rangle \le \langle X, (\lambda I + \mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{L}_A^{-1}X\rangle$$

or

$$\lambda \mathrm{Tr}\,\beta(X^*)\beta(A)^{-1}\beta(X) + \mathrm{Tr}\,\beta(X)\beta(B)^{-1}\beta(X^*) \le \lambda \mathrm{Tr}\,X^*A^{-1}X + \mathrm{Tr}\,XB^{-1}X^*.$$

This inequality is true due to the matrix inequality

$$\beta(X^*)\beta(Y)^{-1}\beta(X) \le \beta(X^*Y^{-1}X) \qquad (Y > 0),$$

see [2]. ☐

The property (3) of Lemma 1 has applications in the Fisher information setting. The mean transformation has a monotonicity property similarly to the matrix means.

Theorem 2 *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be an operator monotone function and $M(\,\cdot\,,\,\cdot\,)$ be the corresponding mean transformation. Assume that $0 < A, B \in \mathbf{M}_n$ and $A \le A', B \le B'$. Then $M(A,B) \le M(A',B')$.*

Proof: Based on the Löwner theorem, we may consider $f(x) = x/(\lambda + x)$ ($\lambda > 0$). Then the statement is

$$\mathbb{L}_A(\lambda I + \mathbb{L}_A\mathbb{R}_B^{-1})^{-1} \le \mathbb{L}_{A'}(\lambda I + \mathbb{L}_{A'}\mathbb{R}_{B'}^{-1})^{-1}$$

which is equivalent to the relation

$$\lambda\mathbb{L}_{A'}^{-1} + \mathbb{R}_{B'}^{-1} = (\lambda I + \mathbb{L}_{A'}\mathbb{R}_{B'}^{-1})\mathbb{L}_{A'}^{-1} \le (\lambda I + \mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{L}_A^{-1} = \lambda\mathbb{L}_A^{-1} + \mathbb{R}_B^{-1}.$$

This is true, since $\mathbb{L}_{A'}^{-1} \le \mathbb{L}_A^{-1}$ and $\mathbb{L}_{B'}^{-1} \le \mathbb{L}_B^{-1}$ due to the assumption. ☐

Theorem 3 *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous function with $f(1) = 1$ and M_f be the corresponding mean transformation. It has the following properties:*

(1) $M_f(\lambda A, \lambda B) = \lambda M_f(A,B)$ for every number $\lambda > 0$;

(2) if $xf(x^{-1}) = f(x)$ then $(M_f(A,B)X)^* = M_f(B,A)X^*$;

(3) $M_f(A,A)I = A$;

(4) $\operatorname{Tr} M_f(A,A)Y = \operatorname{Tr} AY$;

(5) $(A,B) \mapsto \langle X, M_f(A,B)Y\rangle$ is continuous;

(6) $(X,Y) \mapsto \langle X, M_f(A,B)Y\rangle$ is an inner product on $\mathbf{M}_n$ for every $n \in \mathbb{N}^+$;

(7) if

$$C := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} > 0,$$

then

$$M_f(C,C)\begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} M_f(A,A)X & M_f(A,B)Y \\ M_f(B,A)Z & M_f(B,B)W \end{bmatrix}. \tag{5}$$

Proof: The properties (1)–(6) are straightforward consequences of the definition of $M_f(A,B)$. Property (7) is easily checked for $f(x)=x^k$ and thus for all polynomials so by passing to the limit for every f. $\square$

Property (7) is very essential, it tells that it is sufficient to know the mean transformation for two identical matrices.

The joint concavity of operator means [7] implies that for every $A,B,A',B'>0$ we have

$$M\Big(\frac{A+A'}{2},\frac{B+B'}{2}\Big)\geq\frac{M(A,B)+M(A',B')}{2}. \tag{6}$$

This can be deduced also from the transformer inequality.

3 Axiomatic characterization

The next theorem is an axiomatic characterization of the mean transformation. We shall use the notation $H_n^+ := \{A\in \mathbf{M}_n : A>0\}$.

Theorem 4 *Assume that the linear operators $N(A,B):\mathbf{M}_n\to\mathbf{M}_n$ are defined for every $A,B\in H_n^+$, $n\in\mathbb{N}^+$ and have the following properties:*

(i) *$(X,Y)\mapsto\langle X,N(A,B)Y\rangle$ is an inner product on $\mathbf{M}_n$ for every $n\in\mathbb{N}^+$;*

(ii) *$(A,B)\mapsto\langle X,N(A,B)Y\rangle$ is continuous for every $A,B\in H_n^+$ and $n\in\mathbb{N}^+$;*

(iii) *for every trace-preserving completely positive mapping $\beta:\mathbf{M}_n\to\mathbf{M}_k$,*

$$\beta\, N(A,B)\,\beta^*\leq N\big(\beta(A),\beta(B)\big)$$

holds;

(iv) *if*

$$C:=\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix}\in H_{2n}^+$$

then

$$N(C,C)\begin{bmatrix} X & Y\\ Z & W\end{bmatrix}=\begin{bmatrix} N(A,A)X & N(A,B)Y\\ N(B,A)Z & N(B,B)W\end{bmatrix}.$$

Then $N(A,B)$ is a mean transformation: $N(A,B)=M_f(\mathbb{L}_A,\mathbb{R}_B)$ with an operator monotone function $f:\mathbb{R}^+\to\mathbb{R}^+$.

Lemma 2 *If $U,V\in\mathbf{M}_n$ are arbitrary unitary matrices then for every $A,B\in H_n^+$ and $X\in\mathbf{M}_n$ we have*

$$\langle X,N(A,B)X\rangle=\langle UXV^*,N(UAU^*,VBV^*)UXV^*\rangle.$$

Proof: For a unitary matrix $U \in \mathbf{M}_n$ define $\beta(A) = U^*AU$. Then $\beta\colon \mathbf{M}_n \to \mathbf{M}_n$ is trace-preserving completely positive, further, $\beta^*(A) = \beta^{-1}(A) = UAU^*$. Thus by double application of (iii) we obtain

$$\begin{aligned}\langle X, N(A,A)X\rangle &= \langle X, N(\beta\beta^{-1}(A), \beta\beta^{-1}(A))X\rangle \\ &\geq \langle X, \beta N(\beta^{-1}(A), \beta^{-1}(A))\beta^*(X)\rangle \\ &= \langle \beta^*(X), N(\beta^{-1}(A), \beta^{-1}(A))\beta^*(X)\rangle \\ &\geq \langle \beta^*(X), \beta^{-1}N(A,A)(\beta^{-1})^*\beta^*(X)\rangle \\ &= \langle X, N(A,A)X\rangle,\end{aligned}$$

hence

$$\langle X, N(A,A)X\rangle = \langle UXU^*, N(UAU^*, UAU^*)UXU^*\rangle.$$

Now for the matrices

$$C = \begin{bmatrix} A & 0 \\ 0 & B\end{bmatrix} \in H_{2n}^+, \qquad Y = \begin{bmatrix} 0 & X \\ 0 & 0\end{bmatrix} \in \mathbf{M}_{2n} \qquad \text{and} \qquad W = \begin{bmatrix} U & 0 \\ 0 & V\end{bmatrix} \in \mathbf{M}_{2n}$$

it follows by (iv) that

$$\begin{aligned}\langle X, N(A,B)X\rangle &= \langle Y, N(C,C)Y\rangle \\ &= \langle WYW^*, N(WCW^*, WCW^*)WYW^*\rangle \\ &= \langle UXV^*, N(UAU^*, VBV^*)UXV^*\rangle\end{aligned}$$

and we have the statement. □

Lemma 3 *Suppose that $N(A,B)$ is defined by the axioms (i)–(iv). Then there exists a unique continuous function $d\colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$d(r\lambda, r\mu) = rd(\lambda,\mu) \qquad (r, \lambda, \mu > 0)$$

and for every $A = \text{Diag}\,(\lambda_1, \ldots, \lambda_n) \in H_n^+$, $B = \text{Diag}\,(\mu_1, \ldots, \mu_n) \in H_n^+$,

$$\langle X, N(A,B)X\rangle = \sum_{j,k=1}^n d(\lambda_j, \mu_k)|X_{jk}|^2.$$

Proof: The uniqueness of such a function d is clear, we focus on the existence.

Denote by $E(jk)^{(n)}$ and I_n the $n\times n$ matrix units and the $n \times n$ unit matrix, respectively. We assume $A = \text{Diag}\,(\lambda_1, \ldots, \lambda_n) \in H_n^+$, $B = \text{Diag}\,(\mu_1, \ldots, \mu_n) \in H_n^+$.

We first show that

$$\langle E(jk)^{(n)}, N(A,A)E(\ell m)^{(n)}\rangle = 0 \quad \text{if} \quad (j,k) \neq (\ell, m). \tag{7}$$

Indeed, if $j \neq k, \ell, m$ we let $U_j = \mathrm{Diag}\,(1, \ldots, 1, \mathrm{i}, 1, \ldots, 1)$ where the imaginary unit is the jth entry and $j \neq k, \ell, m$. Then by Lemma 2 one has

$$\begin{aligned}\langle E(jk)^{(n)}, N(A,A)E(\ell m)^{(n)}\rangle &= \langle U_j E(jk)^{(n)} U_j^*, N(U_j A U_j^*, U_j A U_j^*) U_j E(lm)^{(n)} U_j^*\rangle \\ &= \langle \mathrm{i} E(jk)^{(n)}, N(A,A)E(\ell m)^{(n)}\rangle \\ &= -\mathrm{i}\langle E(jk)^{(n)}, N(A,A)E(\ell m)^{(n)}\rangle\end{aligned}$$

hence $\langle E(jk)^{(n)}, N(A,A)E(\ell m)^{(n)}\rangle = 0$. If one of the indexes j, k, ℓ, m is different from the others then (7) follows analogously. Finally, applying condition (iv) we obtain that

$$\langle E(jk)^{(n)}, N(A,B)E(\ell m)^{(n)}\rangle = \langle E(j, k+n)^{(2n)}, N(C,C)E(\ell, m+n)^{(2n)}\rangle = 0$$

if $(j,k) \neq (\ell, m)$, because $C = \mathrm{Diag}\,(\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n) \in H_{2n}^+$ and one of the indexes j, $(k+n)$, ℓ, $(m+n)$ are different from the others.

Now we claim that $\langle E(jk)^{(n)}, N(A,B)E(jk)^{(n)}\rangle$ is determined by λ_j, and μ_k. More specifically,

$$\|E(jk)^{(n)}\|_{A,B}^2 = \|E(12)^{(2)}\|_{\mathrm{Diag}\,(\lambda_j,\mu_k)}^2, \tag{8}$$

where for brevity we introduced the notations

$$\|X\|_{A,B}^2 = \langle X, N(A,B)X\rangle \qquad \text{and} \qquad \|X\|_A^2 = \|X\|_{A,A}^2.$$

(The above notations are correct due to condition (i)). Indeed, if $U_{j,k+n} \in \mathbf{M}_{2n}$ denotes the unitary matrix which interchanges the first and the jth, further, the second and the $(k+n)$th coordinates then by condition (iv) and Lemma 2 it follows that

$$\begin{aligned}\|E(jk)^{(n)}\|_{A,B}^2 &= \|E(j,k+n)^{(2n)}\|_C^2 = \|U_{j,k+n} E(j,k+n)^{(2n)} U_{j,k+n}^*\|_{U_{j,k+n} C U_{j,k+n}^*}^2 \\ &= \|E(12)^{(2n)}\|_{\mathrm{Diag}\,(\lambda_j,\mu_k,\lambda_3,\ldots,\mu_n)}^2\end{aligned}$$

thus it suffices to prove

$$\|E(12)^{(2n)}\|_{\mathrm{Diag}\,(\eta_1,\eta_2,\ldots,\eta_{2n})}^2 = \|E(12)^{(2)}\|_{\mathrm{Diag}\,(\eta_1,\eta_2)}^2. \tag{9}$$

Condition (iv) with $X = E(12)^{(n)}$ and $Y = Z = W = 0$ yields

$$\|E(12)^{(2n)}\|_{\mathrm{Diag}\,(\eta_1,\eta_2,\ldots,\eta_{2n})}^2 = \|E(12)^{(n)}\|_{\mathrm{Diag}\,(\eta_1,\eta_2,\ldots,\eta_n)}^2. \tag{10}$$

Further, the mappings ($n \geq 4$) $\beta_n \colon \mathbf{M}_n \to \mathbf{M}_{n-1}$,

$$\beta_n(E(jk)^{(n)}) := \begin{cases} E(jk)^{(n-1)}, & \text{if } 1 \leq j,k \leq n-1, \\ E(n-1,n-1)^{(n-1)}, & \text{if } j = k = n, \\ 0, & \text{otherwise,}\end{cases}$$

and $\tilde{\beta}_n \colon \mathbf{M}_{n-1} \to \mathbf{M}_n$,

$$\tilde{\beta}_n(E(jk)^{(n-1)}) := \begin{cases} E(jk)^{(n)}, & \text{if } 1 \leq j,k \leq n-2, \\ \dfrac{\eta_{n-1}E(n-1,n-1)^{(n)} + \eta_n E(nn)^{(n)}}{\eta_{n-1}+\eta_n}, & \text{if } j = k = n-1, \\ 0, & \text{otherwise}\end{cases}$$

are trace-preserving completely positive hence by (iii)

$$\begin{aligned}
\|E(12)^{(n)}\|^2_{\mathrm{Diag}\,(\eta_1,\dots,\eta_n)} &= \|E(12)^{(n)}\|^2_{\tilde{\beta}_n\beta_n\mathrm{Diag}\,(\eta_1,\dots,\eta_n)} \\
&\geq \|\tilde{\beta}^*_n E(12)^{(n)}\|^2_{\beta_n\mathrm{Diag}\,(\eta_1,\dots,\eta_n)} \\
&\geq \|\beta^*_n\tilde{\beta}^*_n E(12)^{(n)}\|^2_{\mathrm{Diag}\,(\eta_1,\dots,\eta_n)} \\
&= \|E(12)^{(n)}\|^2_{\mathrm{Diag}\,(\eta_1,\dots,\eta_n)}.
\end{aligned}$$

Thus equality holds which implies

$$\|E(12)^{(n)}\|^2_{\mathrm{Diag}\,(\eta_1,\dots,\eta_{n-1},\eta_n)} = \|E(12)^{(n-1)}\|^2_{\mathrm{Diag}\,(\eta_1,\dots,\eta_{n-2},\eta_{n-1}+\eta_n)}. \tag{11}$$

Now repeated application of (10) and (11) yields (9) and therefore also (8) follows.

For $\lambda, \mu > 0$ let

$$d(\lambda,\mu) := \|E(12)^{(2)}\|^2_{\mathrm{Diag}\,(\lambda,\mu)}.$$

Condition (i) yields that $d > 0$, moreover (ii) implies the continuity of d. We further claim that d is homogeneous of order one, that is,

$$d(r\lambda, r\mu) = rd(\lambda,\mu) \qquad (\lambda,\mu,r>0).$$

First let $r = k \in \mathbb{N}^+$. Then the mappings $\alpha_k\colon \mathbf{M}_2 \to \mathbf{M}_{2k}$, $\tilde{\alpha}_k\colon \mathbf{M}_{2k} \to \mathbf{M}_2$

$$\alpha_k(X) = \frac{1}{k} I_k \otimes X \qquad \text{and} \qquad \tilde{\alpha}_k\left(\begin{bmatrix} X_{11} & X_{22} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kk} \end{bmatrix}\right) = X_{11} + X_{22} + \ldots + X_{kk}$$

are trace-preserving completely positive, further, $\tilde{\alpha}^*_k = k\alpha_k$ so applying condition (iii) twice it follows

$$\begin{aligned}
\|E(12)^{(2)}\|^2_{\mathrm{Diag}\,(\lambda,\mu)} &= \|E(12)^{(2)}\|^2_{\tilde{\alpha}_k\alpha_k\mathrm{Diag}\,(\lambda,\mu)} \\
&\geq \|\tilde{\alpha}^*_k E(12)^{(2)}\|^2_{\alpha_k\mathrm{Diag}\,(\lambda,\mu)} \\
&\geq \|\alpha^*_k\tilde{\alpha}^*_k E(12)^{(2)}\|^2_{\mathrm{Diag}\,(\lambda,\mu)} \\
&= \|E(12)^{(2)}\|^2_{\mathrm{Diag}\,(\lambda,\mu)}
\end{aligned}$$

hence equality holds which means

$$\|E(12)^{(2)}\|^2_{\mathrm{Diag}\,(\lambda,\mu)} = \|I_k \otimes E(12)^{(2)}\|^2_{\frac{1}{k}I_k\otimes\mathrm{Diag}\,(\lambda,\mu)}.$$

Thus by applying (7) and (8) we obtain

$$\begin{aligned} d(\lambda,\mu) &= \|I_k \otimes E(12)^{(2)}\|^2_{\frac{1}{k}I_k\otimes \mathrm{Diag}\,(\lambda,\mu)} \\ &= \sum_{j=1}^{k} \|E(jj)^{(k)} \otimes E(12)^{(2)}\|^2_{\frac{1}{k}I_k\otimes \mathrm{Diag}\,(\lambda,\mu)} \\ &= k\|E(11)^{(k)} \otimes E(12)^{(2)}\|^2_{\frac{1}{k}I_k\otimes \mathrm{Diag}\,(\lambda,\mu)} \\ &= kd\left(\frac{\lambda}{k},\frac{\mu}{k}\right). \end{aligned}$$

If $r=\ell/k$ where $\ell,k\in\mathbb{N}^+$ then

$$d(r\lambda,r\mu)=d\left(\frac{\ell}{k}\lambda,\frac{\ell}{k}\mu\right)=\frac{1}{k}d(\ell\lambda,\ell\mu)=\frac{\ell}{k}d(\lambda,\mu).$$

The continuity of d yields the homogeneity for every $r>0$.

We finish the proof by using (7) and (8) and obtain

$$\|X\|^2_{A,B}=\sum_{j,k=1}^{n} d(\lambda_j,\mu_k)|X_{jk}|^2.$$

□

Lemma 4 *Suppose that $N(A,B)=f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B\colon \mathbf{M}_n\to\mathbf{M}_n$ where $f:\mathbb{R}^+\to\mathbb{R}^+$ is a continuous function. If N has the property (iii) then f is an operator monotone function.*

Proof: Let $0\le\lambda\le1$, $A_1,A_2,B_1,B_2,\in H_n^+$, $X\in\mathbf{M}_n$ and put $A=\lambda A_1+(1-\lambda)A_2\in H_n^+$, $B=\lambda B_1+(1-\lambda)B_2\in H_n^+$. We prove that

$$\langle X,(\lambda f(\mathbb{L}_{A_1}\mathbb{R}_{B_1}^{-1})\mathbb{R}_{B_1}+(1-\lambda)f(\mathbb{L}_{A_2}\mathbb{R}_{B_2}^{-1})\mathbb{R}_{B_2})X\rangle\le\langle X,f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_BX\rangle, \tag{12}$$

i.e., $(A,B)\mapsto N(A,B)$ is jointly concave. By choosing $B_1=B_2=I_n$ in (12) it follows

$$\langle X,(\lambda f(\mathbb{L}_{A_1})+(1-\lambda)f(\mathbb{L}_{A_2}))X\rangle\le\langle X,f(\mathbb{L}_{\lambda A_1+(1-\lambda)A_2})X\rangle$$

or equivalently

$$\langle X,(\lambda f(A_1)+(1-\lambda)f(A_2))X\rangle\le\langle X,f(\lambda A_1+(1-\lambda)A_2)X\rangle$$

meaning that f is operator concave which is equivalent to the operator monotonicity [4].

In order to show (12) define $\beta_{2n}\colon\mathbf{M}_{2n}\to\mathbf{M}_{2n}$ as

$$\beta_{2n}\left(\begin{bmatrix}X_{11}&X_{12}\\X_{21}&X_{22}\end{bmatrix}\right)=\frac{1}{2}\begin{bmatrix}X_{11}+X_{22}&0\\0&X_{11}+X_{22}\end{bmatrix}.$$

It is easily seen that β_{2n} is trace-preserving completely positive and Hermitian with respect to the Hilbert-Schmidt inner product. Denote

$$\tilde{A} = \begin{bmatrix} \lambda A_1 & 0 \\ 0 & (1-\lambda)A_2 \end{bmatrix} \in H_{2n}^+, \quad \tilde{B} = \begin{bmatrix} \lambda B_1 & 0 \\ 0 & (1-\lambda)B_2 \end{bmatrix} \in H_{2n}^+, \quad \tilde{X} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \in \mathbf{M}_{2n}.$$

Then condition (iii) implies

$$\langle \tilde{X}, \beta_{2n} f(\mathbb{L}_{\tilde{A}}\mathbb{R}_{\tilde{B}}^{-1})\mathbb{R}_{\tilde{B}}\beta_{2n}^*(\tilde{X})\rangle \leq \langle \tilde{X}, f(\mathbb{L}_{\beta_{2n}(\tilde{A})}\mathbb{R}_{\beta_{2n}(\tilde{B})}^{-1})\mathbb{R}_{\beta_{2n}(\tilde{B})}\tilde{X}\rangle. \tag{13}$$

Since $\beta_{2n}^*(\tilde{X}) = \beta_{2n}(\tilde{X}) = \tilde{X}$ by simple calculation we obtain

$$\begin{aligned}
&\langle \tilde{X}, \beta_{2n} f(\mathbb{L}_{\tilde{A}}\mathbb{R}_{\tilde{B}}^{-1})\mathbb{R}_{\tilde{B}}\beta_{2n}^*(\tilde{X})\rangle \\
&= \langle \tilde{X}, f(\mathbb{L}_{\tilde{A}}\mathbb{R}_{\tilde{B}}^{-1})\mathbb{R}_{\tilde{B}}\tilde{X}\rangle \\
&= \mathrm{Tr}\begin{bmatrix} X^* f(\mathbb{L}_{\lambda A_1}\mathbb{R}_{\lambda B_1}^{-1})\mathbb{R}_{\lambda B_1}X & 0 \\ 0 & X^* f(\mathbb{L}_{(1-\lambda)A_2}\mathbb{R}_{(1-\lambda)B_2}^{-1})\mathbb{R}_{(1-\lambda)B_2}X \end{bmatrix} \\
&= \langle X, (\lambda f(\mathbb{L}_{A_1}\mathbb{R}_{B_1}^{-1})\mathbb{R}_{B_1} + (1-\lambda) f(\mathbb{L}_{A_2}\mathbb{R}_{B_2}^{-1})\mathbb{R}_{B_2})X\rangle.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\langle \tilde{X}, f(\mathbb{L}_{\beta_{2n}(\tilde{A})}\mathbb{R}_{\beta_{2n}(\tilde{B})}^{-1})\mathbb{R}_{\beta_{2n}(\tilde{B})}\tilde{X}\rangle \\
&= \frac{1}{2}\mathrm{Tr}\begin{bmatrix} X^* f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B X & 0 \\ 0 & X^* f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B X \end{bmatrix} \\
&= \langle X, f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B X\rangle
\end{aligned}$$

so that from (13) follows (12). □

Notice that Lemma 4 is the converse of Theorem 1. Now we are ready to prove Theorem 4.

Proof of Theorem 4: Let $d : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ be the unique function according to Lemma 3 and put $f(x) := d(x, 1)$. Then applying the homogeneity of d we obtain that for every $A, B \in H_n^+$ (which might be assumed to be diagonal due to Lemma 2)

$$\begin{aligned}
\langle X, N(A,B)X\rangle &= \sum_{j,k=1}^{n} d(\lambda_j, \mu_k)|X_{jk}|^2 \\
&= \sum_{j,k=1}^{n} |X_{jk}|^2 \left\langle E(jk)^{(n)}, f\left(\frac{\lambda_j}{\mu_k}\right)\mu_k E(jk)^{(n)}\right\rangle \\
&= \langle X, f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B X\rangle
\end{aligned}$$

hence $N(A,B) = f(\mathbb{L}_A\mathbb{R}_B^{-1})\mathbb{R}_B$ and therefore $f : \mathbb{R}^+ \to \mathbb{R}^+$ is operator monotone by Lemma 4. □

As a consequence of Theorem 4 we obtain the characterization of the inverse of a mean transformation.

Theorem 5 *Assume that the linear operators $L(A,B) : \mathbf{M}_n \to \mathbf{M}_n$ are defined for every $A,B \in H_n^+$, $n \in \mathbb{N}^+$ and have the following properties:*

(i') $(X,Y) \mapsto \langle X, L(A,B)Y\rangle$ *is an inner product on $\mathbf{M}_n$ for every $n \in \mathbb{N}^+$;*

(ii') $(A,B) \mapsto \langle X, L(A,B)Y\rangle$ *is continuous for every $A,B \in H_n^+$ and $n \in \mathbb{N}^+$;*

(iii') for every trace-preserving completely positive mapping $\beta : \mathbf{M}_n \to \mathbf{M}_k$,

$$\beta^* L(A,B)\,\beta \le L\big(\beta(A), \beta(B)\big)$$

holds;

(iv') if

$$C := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in H_{2n}^+$$

then

$$L(C,C)\begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} L(A,A)X & L(A,B)Y \\ L(B,A)Z & L(B,B)W \end{bmatrix}.$$

Then $L(A,B)$ is the inverse of a mean transformation: $L(A,B) = M_f(\mathbb{L}_A, \mathbb{R}_B)^{-1}$ with an operator monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$.

Proof: Condition (i) implies that $N(A,B) := L^{-1} : \mathbf{M}_n \to \mathbf{M}_n$ is well-defined for every $A,B \in H_n^+$, $n \in \mathbb{N}^+$. Clearly, conditions (i'), (ii') and (iv') for L are equivalent to conditions (i), (ii), (iv) for $N(A,B)$ in Theorem 4. In addition, (iii) and (iii') are equivalent due to Lemma 1. Therefore, by Theorem 4 it follows that $N(A,B) = M_f(A,B)$ with an operator monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$ and hence $L(A,B) = M_f(A,B)^{-1}$. □

4 Complete positivity

In this section we analyze the complete positivity of M_f for standard matrix monotone functions f (i.e. $xf(x^{-1}) = f(x)$ and $f(1) = 1$). If we require the positivity of $M_f(A,B)X$ for $X \ge 0$, then from the formula

$$(M_f(A,B)X)^* = M_f(B,A)X^*$$

we need $A = B$. If $A = \sum_i \lambda_i |x_i\rangle\langle x_i|$ and $X = \sum_{i,j} X_{ij} |x_i\rangle\langle x_j|$ with an orthonormal basis $\{|x_i\rangle : i\}$, then

$$\Big(M_f(A,A)X\Big)_{ij} = X_{ij} m_f(\lambda_i, \lambda_j)$$

where $m_f(x,y) = f(x/y)y$. The choice $X_{ij} = 1$ shows that the positivity of the matrix

$$K_{ij}^f = m_f(\lambda_i, \lambda_j) \tag{14}$$

is necessary. Given the positive numbers $\{\lambda_i : 1 \le i \le n\}$, the matrix (14) is called an $n \times n$ mean matrix. From the previous argument the positivity of $M(A, A) : \mathbf{M}_n \to \mathbf{M}_n$ implies the positivity of the $n \times n$ mean matrices of the (numerical) mean m_f. It is easy to see [1] that if the mean matrices of any size are positive, then $M_f(A, A) : \mathbf{M}_n \to \mathbf{M}_n$ is a completely positive mapping for every $A > 0$. There are many examples in the paper [3] and the paper [1] studied the complete positivity of $M(A, A)^{-1}$.

Example 1 The *power mean* or *binomial mean*

$$m(x, y) = \left(\frac{x^p + y^p}{2}\right)^{\frac{1}{p}} \qquad (-\infty < p < +\infty)$$

is induced by

$$f_p(x) = \left(\frac{x^p + 1}{2}\right)^{\frac{1}{p}}$$

(notice that f_{-1} is that harmonic mean, f_0 is the geometric mean as limit and f_1 is the arithmetic mean.) It is shown in [1] that f_p is matrix monotone if and only if $-1 \le p \le 1$. Furthermore, for $-1 \le p \le 0$ the mean matrices K^{f_p} are positive, see [3], so $M_{f_p}(A, A)$ is completely positive for every $A > 0$. For $0 \le p \le 1$ the mappings $M_{f_p}(A, A)^{-1}$ are completely positive. □

Example 2 The function

$$f(x) := \frac{1}{2}\left(\frac{x+1}{2} + \frac{2x}{x+1}\right)$$

is matrix monotone, but $M_f(A, A)^{-1}$ is not completely positive, see [1]. □

References

[1] Á. Besenyei and D. Petz, Completely positive mappings and mean matrices, Lin. Alg. Appl. **435**(2011), 984–997.

[2] R. Bhatia, *Matrix Analysis*, Springer-Verlag, 1997.

[3] R. Bhatia and H. Kosaki, Mean matrices and infinite divisibility, Linear Algebra Appl. **424**(2007), 36–54.

[4] F. Hansen and G.K. Pedersen, Jensen's inequality for operators and Löwner's theorem, Math. Ann. **258**(1982), 229–241.

[5] F. Hiai and H. Kosaki, *Means of Hilbert space operators*, Lecture Notes in Mathematics, 1820. Springer-Verlag, Berlin, 2003.

[6] F. Hiai and D. Petz, From quasi-entropy, to be published.

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

[8] W. Kumagai, A characterization of extended monotone metrics, Linear Algebra Appl. **434**(2011), 224–231.

[9] D. Petz, Monotone metrics on matrix spaces, Linear Algebra Appl. **244**(1996), 81–96.

[10] D. Petz, Covariance and Fisher information in quantum mechanics, J. Phys. A: Math. Gen. **35**(2003), 79–91.