Journal of Statistical Physics, Vol. 124, No. 1, July 2006 (© 2006)
DOI: 10.1007/s10955-006-9155-2

Extensions of Lieb's Concavity Theorem

Frank Hansen[1]

Received January 18, 2006; accepted June 8, 2006
Published Online: July 21, 2006

The operator function $(A, B) \to Tr f(A, B)(K^*)K$, defined in pairs of bounded self-adjoint operators in the domain of a function f of two real variables, is convex for every Hilbert Schmidt operator K, if and only if f is operator convex. We obtain, as a special case, a new proof of Lieb's concavity theorem for the function $(A, B) \to Tr\, A^p K^* B^q K$, where p and q are non-negative numbers with sum $p + q \le 1$. In addition, we prove concavity of the operator function

$$(A, B) \to Tr\left[\frac{A}{A+\mu_1}K^*\frac{B}{B+\mu_2}K\right]$$

in its natural domain $D_2(\mu_1, \mu_2)$, cf. Definition 3.

KEY WORDS: Lieb's concavity theorem, operator convex function, generalized Hessian.

1. INTRODUCTION

Let $f : D \to \mathbf{R}$ be a function of two variables defined in a set $D \subseteq \mathbf{R}^2$, and let $M_{n\times m}$ denote the set of complex $n \times m$ matrices (with the abbreviation M_n for $M_{n\times n}$). We say that two Hermitian matrices $(A, B) \in M_n \times M_m$ are in the domain of f, if the product $\sigma(A) \times \sigma(B)$ of the spectra is included in D. We shall consider two different but related notions of matrix functions associated with f.

1.1. The Functional Calculus

Following Korányi[16] we introduce the functional calculus

$$f(A, B) = \sum_{i=1}^{p}\sum_{j=1}^{q} f(\lambda_i, \mu_j) P_i \otimes Q_j \tag{1}$$

[1] Department of Economics, University of Copenhagen, Studiestraede 6, DK-1455 Copenhagen K, Denmark

for functions f of two variables, where

$$A = \sum_{i=1}^{p} \lambda_i P_i \quad \text{and} \quad B = \sum_{j=1}^{q} \mu_j Q_j \tag{2}$$

are the spectral decompositions of A and B. If f can be written as a product $f(t, s) = g(t)h(s)$ of two functions each depending only on one variable then $f(A, B) = g(A) \otimes h(B)$. We say that f is matrix convex of order (n, m), if D is convex and

$$f(\lambda A + (1-\lambda)B, \lambda C + (1-\lambda)D) \leq \lambda f(A, C) + (1-\lambda) f(B, D)$$

for all pairs of Hermitian matrices $(A, C), (B, D) \in M_n \times M_m$ in the domain of f and $\lambda \in [0, 1]$. Note that $(\lambda A + (1-\lambda)B, \lambda C + (1-\lambda)D)$ automatically is in the domain of f.

This type of functional calculus may for continuous functions be extended to bounded, linear and self-adjoint operators on a Hilbert space by replacing sums with integrals, hence

$$f(A, B) = \int f(\lambda, \mu)\, dE_A(\lambda) \otimes dE_B(\mu), \tag{3}$$

where $E_A \otimes E_B$ is the product measure constructed from the two spectral measures E_A and E_B. It is well-defined on products of Borel sets in $\mathbf{R}$ since $E_A \otimes 1$ and $1 \otimes E_B$ commute, and it may be extended to Borel sets in $\mathbf{R}^2$. The support of the measure is contained in $\sigma(A) \times \sigma(B)$.

The function f is said to be operator convex, if D is convex and the operator function $(A, B) \to f(A, B)$ is convex in pairs of operators in the domain of f. It is not difficult to establish that f is operator convex, if an only if it is matrix convex of all orders. The proof follows a suggestion by Löwner (for operator monotone functions) as reported by Bendat and Sherman[3] (Lemma 2.2) and can easily be adapted to the present situation. Note finally that this type of functional calculus may be generalized to functions of k variables, together with the notion of operator convexity or matrix convexity of a fixed order $(n_1, \ldots, n_k)$.

1.2. The Variant Functional Calculus

We may also define an endomorphism $K \to f(A, B)(K)$ of $M_{n\times m}$ by setting

$$f(A, B)(K) = \sum_{i=1}^{p} \sum_{j=1}^{q} f(\lambda_i, \mu_j) P_i K Q_j \tag{4}$$

for each $K \in M_{n\times m}$. If f can be written as a product $f(t, s) = g(t)h(s)$ of two functions each depending only on one variable then $f(A, B)(K) = g(A)Kh(B)$. This type of functional calculus is difficult to extend to bounded linear operators on a Hilbert space H, since there is no obvious way of constructing a measure on H from the two spectral measures E_A and E_B. These questions "were extensively investigated by Birman and Solomyak[5,6] within the very general scope of their theory of double operator integrals," and it is only possible to extend the type of functional calculus in (4) to bounded linear operators for a special class of functions, cf. also Ref. 15. The variant functional calculus is in the literature sometimes expressed in terms of "super operators" acting on $M_{n\times m}$ by setting

$$f(A,B)(K) = f(L_A, R_B)K,$$

where L_A and R_B are commuting left and right multiplication operators (by A and B).

1.3. Convexity Statements

The two types of functional calculus are connected by the following construction. Let H_1 and H_2 be Hilbert spaces of finite dimensions n_1 and n_2 equipped with fixed orthonormal bases $(e_1^1, \ldots, e_{n_1}^1)$ and $(e_1^2, \ldots, e_{n_2}^2)$. Let furthermore

$$\{e_{ij}\}_{i=1,\ldots,n_1;\, j=1,\ldots,n_2}$$

be the system of matrix units in $B(H_2, H_1)$ such that

$$e_{ij}e_m^2 = \delta_{jm}e_i^1 \qquad j, m = 1, \ldots, n_2;\ i = 1, \ldots, n_1.$$

Let $\bar{H}_2$ denote the Hilbert space conjugate[2] to H_2 and consider the linear bijection $\Phi\colon H_1 \otimes \bar{H}_2 \to B(H_2, H_1)$ such that

$$\Phi\big(e_i^1 \otimes e_j^2\big) = e_{ij} \qquad i = 1, \ldots, n_1;\ j = 1, \ldots, n_2.$$

It is not difficult to establish that Φ is unitary and that

$$\Phi(f(A,B)\varphi) = f(A,B)(\Phi(\varphi)), \tag{5}$$

hence

$$(f(A,B)\varphi \mid \varphi)_{H_1\otimes\bar{H}_2} = Tr\,(f(A,B)(\Phi(\varphi))\Phi(\varphi)^*) \tag{6}$$

for self-adjoint operators (A, B) in the domain of f such that A is acting on H_1 and B is acting on H_2, and every vector $\varphi \in H_1 \otimes \bar{H}_2$. We consequently obtain:

Theorem 1.1 *Let $f : D \to \mathbf{R}$ be a function defined in a convex set $D \subseteq \mathbf{R}^2$. The matrix function*

$$(A,B) \to Trf(A,B)(K^*)K,$$

[2] This means that H_2 and $\bar{H}_2$ are identical as complex vector spaces, but the inner products are conjugate to each other.

defined in pairs of Hermitian matrices $(A, B) \in M_n \times M_m$ *in the domain of* f, *is convex for all matrices* $K \in M_{m\times n}$ *if and only if* f *is matrix convex of order* (n, m).

Lieb's concavity theorem states that the mapping

$$(A, B) \to Tr\, A^p K^* B^q K,$$

defined in pairs of positive definite operators, is concave for arbitrary Hilbert Schmidt operators K and non-negative exponents p and q with $p + q \leq 1$. Let us therefore, for these exponents, consider the function $f(t, s) = t^p s^q$ defined in the first quadrant. Since

$$Tr f(A, B)(K^*)K = Tr\, A^p K^* B^q K$$

we realize by Theorem 1.1 that Lieb's concavity theorem is a reflection of the operator concavity of the function f. But Theorem 1.1 also sets the scope for the largest possible extension of Lieb's theorem, not only for operators but for each class of matrices. These distinctions are significant because of the richness of the class of matrix convex functions. In a forthcoming paper[14] we show that there to any interval I different from the real line and to each natural number n exist a function in I which is matrix convex of order n, but not matrix convex of order $n + 1$.

2. SOME OPERATOR CONCAVE FUNCTIONS

In this section we study some well-known operator concave functions with the aim to give truly elementary or otherwise illuminating proofs. The basic tool is the geometric mean # for positive operators A and B introduced by Pusz and Woronowicz.[2,17,23] It is increasing, concave and given by

$$A \# B = A^{1/2}(A^{-1/2} B A^{-1/2})^{1/2} A^{1/2},$$

if A is invertible. Note that $A \# B = (AB)^{1/2}$ if A and B commute. The geometric mean $A \# B$ may be characterized as the maximum of all self-adjoint C such that the block matrix

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

is positive semi-definite. Adapting the reasoning in Ref. 2 (Corollary 2.2) we obtain:

Proposition 2.1. *Let* f *and* g *be non-negative operator concave functions of* k *variables defined in some convex domain* D *in* $\mathbf{R}^k$. *The function*

$$F(t_1, \ldots, t_k) = f(t_1, \ldots, t_k)^{1/2} g(t_1, \ldots, t_k)^{1/2}$$

is then also operator concave in the domain D.

Proof: We consider k-tuples $(A_1, \ldots, A_k)$ and $(B_1, \ldots, B_k)$ of self-adjoint operators in the domain D and note that

$$F(A_1, \ldots, A_k) = f(A_1, \ldots, A_k) \# g(A_1, \ldots, A_k).$$

The statement now follows from the calculation

$$\begin{aligned}
&F\left(\frac{A_1 + B_1}{2}, \ldots, \frac{A_k + B_k}{2}\right) \\
&= f\left(\frac{A_1 + B_1}{2}, \ldots, \frac{A_k + B_k}{2}\right) \# g\left(\frac{A_1 + B_1}{2}, \ldots, \frac{A_k + B_k}{2}\right) \\
&\geq \frac{f(A_1, \ldots, A_k) + f(B_1, \ldots, B_k)}{2} \# \frac{g(A_1, \ldots, A_k) + g(B_1, \ldots, B_k)}{2} \\
&\geq \frac{f(A_1, \ldots, A_k) \# g(A_1, \cdots, A_k)}{2} + \frac{f(B_1, \ldots, B_k) \# g(B_1, \cdots, B_k)}{2} \\
&= \frac{F(A_1, \ldots, A_k) + F(B_1, \cdots, B_k)}{2},
\end{aligned}$$

where we used the concavity of f and g and monotonicity of the geometric mean in the first inequality, and the concavity of the geometric mean in the second. □

Note that the above proposition may be formulated also for classes of matrix concave functions of a fixed order $(n_1, \ldots, n_k)$.

Corollary 2.2 *The functions $(t_1, \ldots, t_k) \to t_1^{p_1} \cdots t_k^{p_k}$ are operator concave in $\mathbf{R}_+^k$ for non-negative exponents $p_1, \ldots, p_k$ with sum $p_1 + \cdots + p_k \leq 1$.*

Proof: Consider the simplex $S = \{(p_1, \ldots, p_k) \mid p_i \geq 0, p_1 + \cdots + p_k \leq 1\}$ and the set of exponents

$$E = \left\{(p_1, \ldots, p_k) \in S \mid t_1^{p_1} \cdots t_k^{p_k} \text{ is operator concave in } \mathbf{R}_+^k\right\}.$$

The vertices $(0, 0, \ldots, 0)$ and $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$ of the convex polytope S are in E, hence $S = \text{conv}(E)$. Since E is closed and mid-point convex by Proposition 2.1, we therefore obtain $E = S$. □

This gives for $k = 1$ the operator concavity in the positive half-axis of the function $t \to t^p$ for $0 \leq p \leq 1$. For $k = 2$ we obtain concavity in the first quadrant of the function $(t, s) \to t^p s^q$ for non-negative exponents with sum $p + q \leq 1$. This is essentially Lieb's concavity theorem, cf. also Ando[(2)] (Corollary 6.2) who gave a different proof. The method of considering convex sets of exponents to prove concavity of the map $A \to A^p \otimes A^q$ appeared in the unpublished notes (1, Theorem IV.3) by Ando. The same technique also appeared in a study of operator monotone functions,[22] and very recently in a study of Morozova-Chentsov functions.[12] (Remark 2.4).

3. NEW OPERATOR CONCAVE FUNCTIONS

Let us henceforth consider the functions

$$f(t_1, \ldots, t_k) = \frac{t_1}{t_1 + \mu_1} \cdots \frac{t_k}{t_k + \mu_k} \qquad t_1, \ldots, t_k > 0, \tag{7}$$

where $\mu_1, \ldots, \mu_k > 0$.

Definition 3.1. We define the domain $D_k(\mu_1, \ldots, \mu_k) \subset \mathbf{R}_+^k$ (abbreviated D_k when there is no confusion) as the set of k-tuples $(t_1, \ldots, t_k) \in \mathbf{R}_+^k$ such that the matrix

$$A_k(t_1, \ldots, t_k) = \begin{pmatrix} \dfrac{2t_1}{\mu_1} & -1 & \cdots & -1 \\ -1 & \dfrac{2t_2}{\mu_2} & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & \dfrac{2t_k}{\mu_k} \end{pmatrix} \tag{8}$$

is positive semi-definite.

It readily follows from the above definition that D_k is a closed convex set, and that $(ct_1, \ldots, ct_k) \in D_k$ for $(t_1, \ldots, t_k) \in D_k$ and $c \geq 1$.

Proposition 3.2 *The function f defined in (7) is concave in the convex domain D_k. Furthermore, any open convex set in $\mathbf{R}_+^k$ in which f is concave is already contained in D_k.*

Proof: The Hessian matrix $H_f(t_1, \ldots, t_k)$ of f is given by

$$f(t_1, \ldots, t_k) \begin{pmatrix} \dfrac{-2\mu_1}{t_1(t_1+\mu_1)^2} & \dfrac{\mu_1\mu_2}{t_1 t_2(t_1+\mu_1)(t_2+\mu_2)} & \cdots \\ \dfrac{\mu_2\mu_1}{t_2 t_1(t_2+\mu_2)(t_1+\mu_1)} & \dfrac{-2\mu_2}{t_2(t_2+\mu_2)^2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

If we introduce the manifestly positive semi-definite matrix

$$P(t_1, \dots, t_k) = f(t_1, \dots, t_k) \left(\frac{\mu_i \mu_j}{t_i t_j (t_i + \mu_i)(t_j + \mu_j)} \right)_{i,j=1}^{k}$$

then the Hessian can be written as the Hadamard product

$$H_f(t_1, \dots, t_k) = -A_k(t_1, \dots, t_k) \circ P(t_1, \dots, t_k),$$

and since a Hadamard product is a principal submatrix of the tensor product, it follows that $H_f(t_1, \dots, t_k)$ is negative semi-definite in the domain D_k. It hence follows that f is concave in D_k. Even though $P(t_1, \dots, t_k)$ is a rank one operator it has a Hadamard inverse

$$P^{\circ -1}(t_1, \dots, t_k) = \frac{1}{f(t_1, \dots, t_k)} \left(\frac{t_i t_j (t_i + \mu_i)(t_j + \mu_j)}{\mu_i \mu_j} \right)_{i,j=1}^{k},$$

which is manifestly positive semi-definite in every point $(t_1, \dots, t_k) \in \mathbf{R}_+^k$, thus

$$A_k(t_1, \dots, t_k) = -H_f(t_1, \dots, t_k) \circ P^{\circ -1}(t_1, \dots, t_k).$$

If the Hessian were negative semi-definite in a point $(t_1, \dots, t_k) \in \mathbf{R}_+^k$ outside of D_k it would then follow that also $A_k(t_1, \dots, t_k)$ is positive semi-definite, and this contradicts the definition of D_k. Therefore f is not concave in any open convex set outside of D_k. □

We have shown that the function f defined in (7) is concave in the domain D_k and nowhere concave outside of this domain. We will prove that f is in fact also operator concave in D_k, but first we need some preliminaries.

3.1. Generalized Hessian Matrices

Matrix or operator convexity of a function of one or several variables may be inferred by calculating the so called generalized Hessian matrices.[(9)] The theory is based on the structure theorem[3] for the second Fréchet differential of the corresponding matrix function.

Let $f: D \to \mathbf{R}$ be a continuous function defined in an open set $D \subseteq \mathbf{R}^k$. We say that a k-tuple of bounded self-adjoint operators $(x_1, \dots, x_k)$ acting on Hilbert spaces $H_1, \dots, H_k$ is contained in the domain of f, if the product of the spectra $\sigma(x_1) \times \cdots \times \sigma(x_k)$ is contained in D. We may then proceed as in (3) to define the bounded self-adjoint operator $f(x_1, \dots, x_k)$ acting on the tensor product $H_1 \otimes \cdots \otimes H_k$.

[3] In the reference we only considered functions defined in a product of open intervals, but the structure theorem is valid for functions defined in arbitrary open sets in $\mathbf{R}^k$.

A data set Λ for f of order $(n_1, \ldots, n_k)$ is a set of points in the domain D written on the form

$$\Lambda = \{(\lambda_{m_1}(1), \ldots, \lambda_{m_k}(k)) \in D \mid m_i = 1, \ldots, n_i \text{ for } i = 1, \ldots, k\}. \tag{9}$$

It may naturally be constructed from the eigenvalues of a k-tuple of Hermitian matrices $(x_1, \ldots, x_k)$ of order $(n_1, \ldots, n_k)$ in the domain of f.

Suppose now that $f\colon D \to \mathbf{R}$ has continuous partial derivatives up to the second order. To a data set Λ for f of order $(n_1, \ldots, n_k)$ as in (9) and a k-tuple of natural numbers $(m_1, \ldots, m_k)$ such that $m_i \le n_i$ for $i = 1, \ldots, k$, the generalized Hessian matrix $H(m_1, \ldots, m_k)$ is defined[9] (Definition 3.1) as the block matrix

$$H(m_1, \ldots, m_k) = \begin{pmatrix} H_{11}(m_1, \ldots, m_k) & \cdots & H_{1k}(m_1, \ldots, m_k) \\ \vdots & \ddots & \vdots \\ H_{k1}(m_1, \ldots, m_k) & \cdots & H_{kk}(m_1, \ldots, m_k) \end{pmatrix},$$

where for $u \neq s$ the $n_u \times n_s$ matrix

$$H_{us}(m_1, \ldots, m_k) = \\ ([\lambda_{m_1}(1)| \cdots |\lambda_{m_s}(s), \lambda_j(s)| \cdots |\lambda_p(u), \lambda_{m_u}(u)| \cdots |\lambda_{m_k}(k)]_f)_{p,j}$$

while the $n_s \times n_s$ matrix

$$H_{ss}(m_1, \ldots, m_k) = (2[\lambda_{m_1}(1)| \cdots |\lambda_{m_s}(s), \lambda_p(s), \lambda_j(s)| \cdots |\lambda_{m_k}(k)]_f)_{p,j}$$

for $s = 1, \ldots, k$. The entries are second order partial divided differences of f (the notation does not imply any particular order of the entries). Note that each generalized Hessian matrix is a quadratic and real symmetric matrix of order $n_1 + \cdots + n_k$.

Theorem 3.3. (The second Fréchet differential) *Let $f\colon D \to \mathbf{R}$ be a real $p > 2 + k/2$ times continuously differentiable function defined in an open set $D \subseteq \mathbf{R}^k$. Then the operator function*

$$(x_1, \ldots, x_n) \to f(x_1, \ldots, x_k),$$

defined in k-tuples $(x_1, \ldots, x_k)$ of bounded self-adjoint operators in the domain of f, is twice Fréchet differentiable. If this function is restricted to k-tuples of Hermitian matrices $(x_1, \ldots, x_k)$ of order $(n_1, \ldots, n_k)$ in the domain of f, then the expectation value of the second Fréchet differential can be written on the form

$$(d^2 f(x)(h, h)\varphi \mid \varphi) \\ = \sum_{m_1=1}^{n_1} \cdots \sum_{m_k=1}^{n_k} (H(m_1, \ldots, m_k)\Phi^h(m_1, \ldots, m_k) \mid \Phi^h(m_1, \ldots, m_k)),$$

where $H(m_1,\dots,m_k)$ is a generalized Hessian matrix associated with f and the data set Λ constructed from the eigenvalues of the matrices $(x_1,\dots,x_k)$. The vectors $\Phi^h(m_1,\dots,m_k)$ are given by

$$\Phi^h(m_1,\dots,m_k)=\begin{pmatrix}\Phi^h_1(m_1,\dots,m_k)\\ \vdots \\ \Phi^h_k(m_1,\dots,m_k)\end{pmatrix},$$

the k-tuple of Hermitian matrices $h=(h^1,\dots,h^k)$ is arbitrary but of order $(n_1,\dots,n_k)$ and the vectors

$$\Phi^h_s(m_1,\dots,m_k)_{j_s}=h^s_{m_s j_s}\varphi(m_1,\dots,m_{s-1},j_s,m_{s+1},\dots,m_k)$$

for $j_s=1,\dots,n_s$ and $s=1,\dots,k$, and the tensor

$$\varphi=\sum_{m_1=1}^{n_1}\cdots\sum_{m_k=1}^{n_k}\varphi(m_1,\dots,m_k)\,e^1_{m_1}\otimes\cdots\otimes e^k_{m_k}$$

is expressed in terms of orthonormal bases of eigenvectors $(e^i_{m_i})_{m_i=1,\dots,n_i}$ of each Hermitian matrix x_i in the k-tuple $(x_1,\dots,x_k)$.

The form of the second Fréchet differential implies[(8)] (Exercises 3.1.8 and 3.6.4) the following result:

Corollary 3.4 *A real $p>2+k/2$ times continuously differentiable function $f\colon D\to\mathbf{R}$ defined in an open convex set $D\subseteq\mathbf{R}^k$ is matrix convex of order $(n_1,\dots,n_k)$, if to each data set Λ for f of order $(n_1,\dots,n_k)$ all of the generalized Hessian matrices $H(m_1,\dots,m_k)$ are positive semi-definite.*

Theorem 3.5 *Let $\mu_1,\dots,\mu_k>0$ be positive real constants. The function*

$$f(t_1,\dots,t_k)=\frac{t_1}{t_1+\mu_1}\cdots\frac{t_k}{t_k+\mu_k}$$

is operator concave in the domain $D_k(\mu_1,\dots,\mu_k)$.

Proof: It is sufficient to prove that f is matrix concave of arbitrary order $(n_1,\dots,n_k)$. For this purpose we consider an arbitrary data set Λ for f of order $(n_1,\dots,n_k)$ written as in (9). The multiplicative form of the function makes it simple to calculate the generalized Hessian matrices. We introduce the vectors

$$a(i)=\left(\frac{\mu_i}{\lambda_1(i)+\mu_i},\dots,\frac{\mu_i}{\lambda_{n_i}(i)+\mu_i}\right)\in\mathbf{R}^{n_i}$$

for $i = 1, \ldots, k$ and calculate for $u \neq s$ the entries

$$
\begin{aligned}
&[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s), \lambda_{j_s}(s)|\cdots|\lambda_{p_u}(u), \lambda_{m_u}(u)|\cdots|\lambda_{m_k}(k)]_f \\
&\quad = \frac{\lambda_{m_1}(1)}{\lambda_{m_1}(1) + \mu_1} \cdots \frac{\mu_s}{(\lambda_{m_s}(s) + \mu_s)(\lambda_{j_s}(s) + \mu_s)} \cdots \\
&\qquad \cdots \frac{\mu_u}{(\lambda_{p_u}(u) + \mu_u)(\lambda_{m_u}(u) + \mu_u)} \cdots \frac{\lambda_{m_k}(k)}{\lambda_{m_k}(k) + \mu_k} \\
&\quad = \frac{f(\lambda_{m_1}(1), \ldots, \lambda_{m_k}(k))}{\lambda_{m_s}(s)\lambda_{m_u}(u)} a(u)_{p_u} a(s)_{j_s}
\end{aligned}
$$

hence the block

$$
H_{us}(m_1, \ldots, m_k) = \frac{f(\lambda_{m_1}(1), \ldots, \lambda_{m_k}(k))}{\lambda_{m_s}(s)\lambda_{m_u}(u)} a(u)^t a(s).
$$

Similarly, we calculate the entries in the diagonal blocks

$$
\begin{aligned}
&2[\lambda_{m_1}(1)|\cdots|\lambda_{m_s}(s), \lambda_{p_s}(s), \lambda_{j_s}(s)|\cdots|\lambda_{m_k}(k)]_f \\
&= \frac{2\lambda_{m_1}(1)}{\lambda_{m_1}(1) + \mu_1} \cdots \frac{-\mu_s}{(\lambda_{m_s}(s) + \mu_s)(\lambda_{p_s}(s) + \mu_s)(\lambda_{j_s}(s) + \mu_s)} \cdots \frac{\lambda_{m_k}(k)}{\lambda_{m_k}(k) + \mu_k} \\
&= -2\frac{f(\lambda_{m_1}(1), \ldots, \lambda_{m_k}(k))}{\mu_s \lambda_{m_s}(s)} a(s)_{p_s} a(s)_{j_s}
\end{aligned}
$$

hence the block

$$
H_{ss}(m_1, \ldots, m_k) = -2\frac{f(\lambda_{m_1}(1), \ldots, \lambda_{m_k}(k))}{\mu_s \lambda_{m_s}(s)} a(s)^t a(s).
$$

In conclusion, the generalized Hessian matrices $H(m_1, \ldots, m_k)$ associated with the function (7) and the data set (9) can be written on the form

$$
f(\lambda_{m_1}, \ldots, \lambda_{m_k})
\begin{pmatrix}
\frac{-2a(1)^t a(1)}{\mu_1 \lambda_{m_1}(1)} & \frac{a(1)^t a(2)}{\lambda_{m_1}(1)\lambda_{m_2}(2)} & \cdots & \frac{a(1)^t a(k)}{\lambda_{m_1}(1)\lambda_{m_k}(k)} \\
\frac{a(2)^t a(1)}{\lambda_{m_2}(2)\lambda_{m_1}(1)} & \frac{-2a(2)^t a(2)}{\mu_2 \lambda_{m_2}(2)} & \cdots & \frac{a(2)^t a(k)}{\lambda_{m_2}(2)\lambda_{m_k}(k)} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{a(k)^t a(1)}{\lambda_{m_k}(k)\lambda_{m_1}(1)} & \frac{a(k)^t a(2)}{\lambda_{m_k}(k)\lambda_{m_2}(2)} & \cdots & \frac{-2a(k)^t a(k)}{\mu_k \lambda_{m_k}(k)}
\end{pmatrix}
$$

where $a(i)^t$ denotes the transpose of $a(i)$. It can be written as the Hadamard product of the manifestly positive semi-definite block matrix

$$f(\lambda_{m_1},\ldots,\lambda_{m_k})\begin{pmatrix} \dfrac{a(1)^t a(1)}{\lambda_{m_1}(1)^2} & \dfrac{a(1)^t a(2)}{\lambda_{m_1}(1)\lambda_{m_2}(2)} & \cdots & \dfrac{a(1)^t a(k)}{\lambda_{m_1}(1)\lambda_{m_k}(k)} \\ \dfrac{a(2)^t a(1)}{\lambda_{m_2}(2)\lambda_{m_1}(1)} & \dfrac{a(2)^t a(2)}{\lambda_{m_2}(2)^2} & \cdots & \dfrac{a(2)^t a(k)}{\lambda_{m_2}(2)\lambda_{m_k}(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{a(k)^t a(1)}{\lambda_{m_k}(k)\lambda_{m_1}(1)} & \dfrac{a(k)^t a(2)}{\lambda_{m_k}(k)\lambda_{m_2}(2)} & \cdots & \dfrac{a(k)^t a(k)}{\lambda_{m_k}(k)^2} \end{pmatrix}$$

and the matrix $-A_k(\lambda_{m_1}(1),\ldots,\lambda_{m_k}(k))$ defined in (8).

All of the generalized Hessian matrices associated with f and Λ are thus negative semi-definite, hence it follows from Corollary 3.4 that f is matrix concave of order $(n_1,\ldots,n_k)$, and since this order is arbitrary, we conclude that f is operator concave. $\square$

Since the above function f is operator concave in the largest domain in which it is concave, we realize that the associated generalized Hessian matrices of a certain order $(n_1,\ldots,n_k)$ are negative semi-definite, if and only if f is matrix concave of the same order. This is in line with the conjecture (known to be true for functions of one variable) that positive semi-definiteness of the generalized Hessian matrices are not only sufficient but also necessary conditions for matrix convexity.

Corollary 3.6 *Let μ_1 and μ_2 be positive real numbers, and let K be a Hilbert Schmidt operator. The operator function*

$$(A,B)\to Tr\left[\frac{A}{A+\mu_1}K^*\frac{B}{B+\mu_2}K\right],$$

defined in pairs (A,B) of positive definite operators, is concave in the convex domain

$$D_2(\mu_1,\mu_2)=\{(t_1,t_2)\in\mathbf{R}_+^2 \mid t_1t_2\geq\mu_1\mu_2/4\}.$$

Note that the operator function in the corollary, for non-vanishing K, is not concave in any open convex set outside of $D_2(\mu_1,\mu_2)$, not even its restriction to pairs of positive real numbers.

APPENDIX

Theorem 4.1 *The function*

$$f(t_1, \ldots, t_k) = \frac{1}{t_1 \cdots t_k}$$

is operator convex in $\mathbf{R}_+^k$.

Proof: Let Λ be a data set for f of order $(n_1, \ldots, n_k)$ as in (9) and set

$$a(i) = \left(\frac{1}{\lambda_1(i)}, \ldots, \frac{1}{\lambda_{n_i}(i)} \right) \in \mathbf{R}_+^{n_i} \qquad i = 1, \ldots, k.$$

It is easy to calculate the generalized Hessian $H(m_1, \ldots, m_k)$ as

$$f(\lambda_{m_1}, \ldots, \lambda_{m_k}) \begin{pmatrix} 2a(1)^t a(1) & a(1)^t a(2) & \cdots & a(1)^t a(k) \\ a(2)^t a(1) & 2a(2)^t a(2) & \cdots & a(2)^t a(k) \\ \vdots & \vdots & \ddots & \vdots \\ a(k)^t a(1) & a(k)^t a(2) & \cdots & 2a(k)^t a(k) \end{pmatrix}$$

for any k-tuple $(m_1, \ldots, m_k) \leq (n_1, \ldots, n_k)$. Since this matrix is manifestly positive semi-definite the assertion follows from Corollary 3.4. $\square$

The above Theorem is due to Ando[2] (Theorem 5) who gave a very different proof. For $k = 2$ the result may be derived from[18] (Corollary 8.1) by using the identification Φ introduced in the introduction. The result is fitting since $-f$ is operator monotone as a function of k variables, cf. (Ref. 11, Page 17).

Corollary 4.2 *The function*

$$f(t_1, \ldots, t_k) = \frac{1}{t_1^{p_1} \cdots t_k^{p_k}}$$

is for arbitrary exponents $p_1, \ldots, p_k \in [0, 1]$ *operator convex in* $\mathbf{R}_+^k$.

Lieb proved (18, Corollary 3.1) convexity of the mapping

$$(A, B, K) \to \int_0^\infty Tr \left[\frac{1}{A + u} K^* \frac{1}{B + u} K \right] du$$

in $B(H)_+ \times B(H)_+ \times B(H)_{\mathrm{HS}}$, cf. also Ref. 21, 24. It is a triviality that the constituent mappings

$$(A, B, K) \to Tr \left[\frac{1}{A + u} K^* \frac{1}{B + u} K \right] \qquad u > 0$$

are not (jointly) convex in $B(H)_+ \times B(H)_+ \times B(H)_{\mathrm{HS}}$. But they are, as noted above, (jointly) convex in the first two variables.

Proposition 4.3 *The mapping* $(A, \xi) \to (A^{-1}\xi \mid \xi)$ *is (jointly) convex for positive invertible operators* A *on a Hilbert space* H, *and vectors* $\xi \in H$.

Proof: Ando noted[4] (2) (Page 208) that the harmonic mean $2(A^{-1} + B^{-1})^{-1}$ of two positive invertible operators A and B on a Hilbert space H can be characterized as the maximum of all Hermitian operators C for which

$$\begin{pmatrix} C & C \\ C & C \end{pmatrix} \le 2 \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{10}$$

Replacing A and B with their inverses and inserting the Harmonic mean $2(A + B)^{-1}$ of A^{-1} and B^{-1} for C, we obtain the inequality

$$\begin{pmatrix} (A+B)^{-1} & (A+B)^{-1} \\ (A+B)^{-1} & (A+B)^{-1} \end{pmatrix} \le \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix} \tag{11}$$

which evaluated in block vectors (ξ, η) for $\xi, \eta \in H$ may be written as

$$\left(\left(\frac{A+B}{2}\right)^{-1}\left(\frac{\xi+\eta}{2}\right) \mid \left(\frac{\xi+\eta}{2}\right)\right) \le \frac{1}{2}\Big((A^{-1}\xi \mid \xi) + (B^{-1}\eta \mid \eta)\Big).$$

But this inequality is the desired result. □

The mapping $A \to A \otimes B$ is linear for a fixed B, thus the mapping

$$(A, \xi) \to ((A^{-1} \otimes B^{-1})\xi \mid \xi)$$

is (jointly) convex for positive invertible operators A and B on a Hilbert space H and vectors $\xi \in H \otimes H$. By using the unitary map $\Phi\colon H \otimes \bar{H} \to B(H)$ introduced in the introduction, we obtain:

Proposition 4.4 *The mapping*

$$(A, B, K) \to Tr\left[\frac{1}{A+u} K^* \frac{1}{B+v} K\right] \qquad u, v > 0$$

defined in $B(H)_+ \times B(H)_+ \times B(H)_{HS}$ *is (jointly) convex in any two of the three variables.*

[4] Since Ando offered no proof, we sketch (10) in the case where C is chosen as the harmonic mean. Use the identity $2(A^{-1} + B^{-1}) = 2A^{1/2}(1 + A^{1/2}B^{-1}A^{1/2})^{-1}A^{1/2}$ and multiply the inequality from the left and from the right with a diagonal block matrix with $A^{-1/2}$ in the diagonal. This transformation reduces (10) to an inequality between commuting operators.

The joint convexity in say (A, K) may also be derived directly from the Lieb-Ruskai convexity theorem (19, Remark after Theorem 1) stating that the mapping $(A, K) \to K^* A^{-1} K$ is convex, where A is positive definite and invertible, and K is arbitrary.

Remark 4.5 Lieb pointed out that Proposition 4.3 may be obtained also as a direct consequence of the Lieb-Ruskai theorem in the following way: Let B_ξ for an arbitrary vector ξ be defined as the operator $B_\xi u = (u \mid v)\xi$ where v is a fixed unit vector. The mapping $\xi \to B_\xi$ is linear, so the composed mapping $(A, \xi) \to B_\xi^* A^{-1} B_\xi$ is jointly convex. The desired result now follows by taking the expectation value in the vector v.

Remark 4.6 One may ask for which functions f defined in $\mathbf{R}_+$ the mapping

$$(A, \xi) \to (f(A)\xi \mid \xi)$$

is (jointly) convex. Obviously f has to be operator convex, and it follows immediately from Proposition 4.3 that any function of the form

$$f(t) = \beta + \int_0^\infty \frac{1}{t+s}\, d\mu(s) \qquad \beta \in \mathbf{R}, \tag{12}$$

where μ is a positive measure with support in $[0, \infty)$ such that the integrals $\int (s^2+1)^{-1}\, d\mu(s)$ and $\int s(s^2+1)^{-1}\, d\mu(s)$ both are finite, has the property. The functions of the form (12) coincide with the class of operator monotone decreasing functions defined in the positive half-axis and bounded from below[13] (Page 9). But not all operator convex functions have the property. If we set $f(t) = t^2$ and choose the projections

$$A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad A_2 = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

together with the vectors $\xi_1 = (1, 0)$ and $\xi_2 = (0, -1)$, then the difference

$$\frac{\left(A_1^2\xi_1 \mid \xi_1\right) + \left(A_2^2\xi_2 \mid \xi_2\right)}{2} - \left(\left(\frac{A_1 + A_2}{2}\right)^2 \left(\frac{\xi_1 + \xi_2}{2}\right) \,\Big|\, \frac{\xi_1 + \xi_2}{2}\right) = -\frac{1}{16}$$

is negative, and this remains so if we perturb A_1 and A_2 slightly such that they become strictly positive.

REFERENCES

1. T. Ando, *Topics on Operator Inequalities* (Sapporo, 1978. Unpublished notes).
2. T. Ando, Concavity of certain maps of positive definite matrices and applications to Hadamard products. *Linear Algebra Appl.* **26**:203–241 (1979).

3. J. Bendat and S. Sherman, Monotone and convex operator functions. *Trans. Amer. Math. Soc.* **79**:58–71 (1955).
4. R. Bhatia. *Matrix analysis* (New York, Springer, 1997).
5. M. Sh. Birman and M. Z. Solomyak, Notes on the function of spectral shift. *Zap. Nauchn. Semin. LOMI* **27**:33–46 (1972). (Russian).
6. M. Sh. Birman and M. Z. Solomyak, Double Stieltjes operator integrals, III. *Problems of Math. Physics, no. 6*, pp. 27–53, (1973). (Russian).
7. H. Epstein, Remarks on two theorems of E. Lieb. *Comm. Math. Phys.* **31**:317–325 (1973).
8. T. M. Flett, *Differential Analysis* (Cambridge, Cambridge University Press, 1980).
9. F. Hansen, Operator convex functions of several variables. *Publ. RIMS, Kyoto Univ.* **33**:443–463 (1997).
10. F. Hansen, Operator inequalities associated with Jensen's inequality. In: T. M. Rassias (edr.), *Survey on Classical Inequalities* (Kluwer Academic Publishers, 2000) pp. 67–98.
11. F. Hansen, Operator monotone functions of several variables. *Math. Ineq. Appl.* **6**:1–17 (2003).
12. F. Hansen, Characterization of symmetric monotone metrics on the the state space of quantum systems. *arXiv:math-ph/0601056 v3*, pp. 1–12, 2006. To appear in Quantum Information and Computation.
13. F. Hansen, Trace functions as Laplace transforms. *J. Math. Phys.* **47**:(043504) 1–11 (2006).
14. F. Hansen and J. Tomiyama, Differential analysis of matrix convex functions. *arXiv:math. OA/0601290 v1*, pp. 1–17, (2006). To appear in Linear Algebra Appl.
15. F. Hiai and H. Kosaki, *Means of Hilbert space operators*. Lecture Notes in Mathematics. (Berlin, Springer, 2003).
16. A. Korányi, On some classes of analytic functions of several variables. *Trans Amer. Math. Soc.* **101**:520–554 (1961).
17. F. Kubo and T. Ando, Means of positive linear operators. *Math. Ann.* **246**:205–224 (1980).
18. E. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture. *Advances in Math.* **11**:267–288 (1973).
19. E. H. Lieb and M. B. Ruskai, Some operator inequalities of the Schwarz type. *Adv. in Math.* **12**:269–273 (1974).
20. N. E. Nörlund, *Vorlesungen Über Differenzenrechnung* (Berlin, Springer Verlag, 1924).
21. M. Ohya and D. Petz, *Quantum Entropy and its Use* (Heidelberg, Springer Verlag, 1993).
22. G. K. Pedersen, Some operator monotone functions. *Proc. Amer. Math. Soc.* **36**:309–310 (1972).
23. W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map. *Rep. Math. Phys.* **8**:159–170 (1975).
24. M. B. Ruskai. Lieb's simple proof of concavity of $(A, B) \to Tr\ A^p K^\dagger B^{1-p} K$ and remarks on related inequalities. *arXiv:quant-ph/0404126 v3*, pp. 1–14, (2005).