ANNALES
POLONICI MATHEMATICI
LXII.1 (1995)

Convex and monotone operator functions

by Jaspal Singh Aujla (Jalandhar) and
H. L. Vasudeva (Chandigarh)

Abstract. The purpose of this note is to provide characterizations of operator convexity and give an alternative proof of a two-dimensional analogue of a theorem of Löwner concerning operator monotonicity.

1. Introduction. For $m \in \mathbb{N}$, let M_m be the algebra of all hermitian $m \times m$ complex matrices. Let I be an interval of $\mathbb{R}$. We denote by $M_m(I)$ the set of all members of M_m whose spectrum is contained in I. Let f be a real function of a real variable x in I. Let $A = \sum_{i=1}^{m'} \lambda_i P_i$ $(m' \le m)$ be the spectral resolution of an $A \in M_m(I)$. By $f(A)$ we understand the matrix $f(A) = \sum_{i=1}^{m'} f(\lambda_i) P_i$.

A real function f on an interval I is *operator monotone* if for each $m \in \mathbb{N}$ and for every pair $A, B \in M_m(I)$ with $A \le B$, we have $f(A) \le f(B)$. Likewise we say that f is *operator convex* if for each $m \in \mathbb{N}$, $f(tA + (1-t)B) \le tf(A) + (1-t)f(B)$ for all $A, B \in M_m(I)$ and every $t \in [0, 1]$.

For known results on operator monotone functions and operator convex functions we refer to Ando [1] and Donoghue [3] rather than to original sources.

Let $m, n \in \mathbb{N}$ and I, J be intervals of $\mathbb{R}$. Let f be a real-valued function of two real variables x in I and y in J. Let $A \in M_m(I)$ and $B \in M_n(J)$ have spectral resolutions $A = \sum_{i=1}^{m'} \lambda_i P_i$ $(m' \le m)$ and $B = \sum_{j=1}^{n'} \mu_j Q_j$ $(n' \le n)$. Then $f(A, B)$, as in Korányi [5], is the matrix

$$f(A, B) = \sum_{i=1}^{m'} \sum_{j=1}^{n'} f(\lambda_i, \mu_j) P_i \otimes Q_j.$$

f is called *operator monotone* on $I \times J$ if for each $m, n \in \mathbb{N}$ and for every $A, A_1 \in M_m(I)$ and $B, B_1 \in M_n(J)$ with $A \le A_1$ and $B \le B_1$ we have the

1991 *Mathematics Subject Classification*: 15A45, 26B25.
Key words and phrases: operator monotone function, operator convex function.

inequality

$$f(A_1, B_1) - f(A_1, B) - f(A, B_1) + f(A, B) \geq 0.$$

Likewise f is called *operator convex* on $I \times J$ if for each $m, n \in \mathbb{N}$,

$$f(t(A, B) + (1-t)(A_1, B_1)) \leq tf(A, B) + (1-t)f(A_1, B_1)$$

for all $A, A_1 \in M_m(I)$, $B, B_1 \in M_n(J)$ and for every $t \in [0, 1]$.

It is the purpose of this note to give characterizations of convex operator functions analogous to those of real-valued convex functions of one or more real variables [8, pp. 98–103]. In the final section, we also provide an alternative proof of a theorem of Korányi [5, Th. 4].

2. One variable case. Consider an open interval I in $\mathbb{R}$ and a continuously differentiable function f on I. Fix $n \in \mathbb{N}$ and take $A \in M_n(I)$. If $\{e_{ij} : 1 \leq i, j \leq n\}$ is a system of matrix units for M_n such that $A = \sum \lambda_i e_{ii}$, we shall denote by $f^{[1]}(A)$ the element in M_n with

$$f^{[1]}(A)|_{i,j} = \begin{cases} (\lambda_i - \lambda_j)^{-1}(f(\lambda_i) - f(\lambda_j)) & \text{if } \lambda_i \neq \lambda_j, \\ f'(\lambda_i) & \text{if } \lambda_i = \lambda_j. \end{cases}$$

Recall that the spectral resolution of a matrix in M_n yields a system of matrix units and that a system of matrix units yields the Hadamard product operation on matrices. The symbol $\circ_X$ shall denote the Hadamard product of matrices in a basis that diagonalizes X.

2.1. Lemma [4, Lemma 3.1 or 2, III]. *With f and A as above, we have*

$$\lim_{\varepsilon \to 0} \varepsilon^{-1}\{f(A + \varepsilon H) - f(A)\} = f^{[1]}(A) \circ_A H$$

for every $H \in M_n$.

For a proof of the lemma, the reader is referred to [4, Lemma 3.1] or [2, III].

2.2. Theorem. *For a function $f \in C^1(I)$, the following statements are equivalent*:

(i) *f is operator convex on I,*
(ii) *$f(A) - f(B) - f^{[1]}(B) \circ_B (A - B) \geq 0$ for all $A, B \in M_n(I)$ and*
(iii) *$f^{[1]}(A) \circ_A (A - B) - f^{[1]}(B) \circ_B (A - B) \geq 0$ for all $A, B \in M_n(I)$.*

Proof. (i)⇒(ii). Fix n and take $B \in M_n(I)$. Choose a system of matrix units for M_n that diagonalizes B. For $A \in M_n(I)$ and $t \in [0, 1]$, we have

$$f(B + t(A - B)) \leq (1-t)f(B) + tf(A).$$

This implies

$$\begin{aligned}0 &= \lim_{t\to 0} t^{-1}\{f(B+t(A-B)) - f(B) - tf^{[1]}(B)\circ_B (A-B)\}\\ &\le f(A) - f(B) - f^{[1]}(B)\circ_B (A-B),\end{aligned}$$

by Lemma 2.1 and the foregoing inequality.

(ii)⇒(iii). Fix $n \in \mathbb{N}$. Let $A, B \in M_n(I)$. Then

$$f(A) - f(B) \ge f^{[1]}(B)\circ_B (A-B)$$

and

$$f(B) - f(A) \ge f^{[1]}(A)\circ_A (B-A).$$

On adding the last two inequalities, one gets

$$f^{[1]}(A)\circ_A (A-B) - f^{[1]}(B)\circ_B (A-B) \ge 0.$$

(iii)⇒(i). Let $\varphi : [0,1] \to M_n$ be defined by

$$\varphi(t) = f(tA + (1-t)B), \qquad A, B \in M_n(I).$$

For $0 \le t_1 < t_2 \le 1$, let $U_i = t_iA + (1-t_i)B$, $i = 1, 2$. Then $U_2 - U_1 = (t_2 - t_1)(A-B)$. In view of the given condition, we have

$$f^{[1]}(U_2)\circ_{U_2} (U_2 - U_1) - f^{[1]}(U_1)\circ_{U_1} (U_2 - U_1) \ge 0.$$

Observe that

$$\varphi'(t) = \lim_{h\to 0} h^{-1}\{\varphi(t+h) - \varphi(t)\} = f^{[1]}(tA + (1-t)B)\circ_X (A-B),$$

where $X = tA + (1-t)B$, by Lemma 2.1. Now,

$$\varphi'(t_1) = f^{[1]}(U_1)\circ_{U_1} (U_2 - U_1) \le f^{[1]}(U_2)\circ_{U_2} (U_2 - U_1) = \varphi'(t_2),$$

i.e. φ' is increasing. Consequently, φ is convex. Therefore,

$$\begin{aligned}f(tA + (1-t)B) &= \varphi(t) = \varphi(t\cdot 1 + (1-t)\cdot 0)\\ &\le t\varphi(1) + (1-t)\varphi(0) = tf(A) + (1-t)f(B). \ \blacksquare\end{aligned}$$

For our next result, we need the following lemma.

2.3. Lemma (cf. [2, III]). *Let $f \in C^2(I)$. Fix n and take $A \in M_n(I)$. If $\{e_{ij} : 1 \le i, j \le n\}$ is a system of matrix units for M_n such that $A = \sum \lambda_i e_{ii}$, then*

$$\begin{aligned}&e_{ii} \lim_{\varepsilon\to 0} \varepsilon^{-2}\{f(A + \varepsilon H) - f(A) - \varepsilon f^{[1]}(A)\circ_A H\}e_{jj}\\ &\qquad = \frac{1}{\lambda_i - \lambda_j}\sum_k \left\{\frac{f(\lambda_i) - f(\lambda_k)}{\lambda_i - \lambda_k} - \frac{f(\lambda_j) - f(\lambda_k)}{\lambda_j - \lambda_k}\right\} h_{ik}h_{kj}e_{ij}\end{aligned}$$

for every $H \in M_n$ and for $\lambda_i \ne \lambda_j \ne \lambda_k$. In case $\lambda_i = \lambda_j$ or $\lambda_j = \lambda_k$ or $\lambda_i = \lambda_k$, the difference quotient on the right hand side is to be replaced by the appropriate derivative.

Proof. If $f(x)=x^p$ then the second order term in ε of $f(A+\varepsilon H)$ is $\sum A^m HA^r HA^s$, the summation being taken over all $m,r,s\ge 0$ such that $m+r+s=p-2$. Consequently,

$$\begin{aligned} e_{ii}\sum_{m,r,s} A^m HA^r HA^s e_{jj} &= \sum_{m,r,s} \lambda_i^m \lambda_j^s e_{ii} HA^r He_{jj} \\ &= \sum_{m,r,s,k} \lambda_i^m \lambda_j^s \lambda_k^r h_{ik}h_{kj}e_{ij} \\ &= (\lambda_i-\lambda_j)^{-1}\sum_k \left\{\frac{\lambda_i^p-\lambda_k^p}{\lambda_i-\lambda_k}-\frac{\lambda_j^p-\lambda_k^p}{\lambda_j-\lambda_k}\right\} h_{ik}h_{kj}e_{ij} \end{aligned}$$

as desired. Since the linear span of such functions is dense in $C^2(I)$ in the topology of uniform convergence on compact sets, the result follows. ■

2.4. THEOREM ([1], Lemma 3.1). *If $f\in C^2(-1,1)$ and $f(0)=0$, then f is operator convex iff $f^{[2]}(A;\mu)\ge 0$, where $A=\sum_i \lambda_i e_{ii}$ and*

$$f^{[2]}(A;\mu)|_{i,j}=(\lambda_i-\lambda_j)^{-1}\left\{\frac{f(\lambda_i)-f(\mu)}{\lambda_i-\mu}-\frac{f(\lambda_j)-f(\mu)}{\lambda_j-\mu}\right\},$$

and $\mu\in(-1,1)$ is arbitrary. The right hand of the above equality is to be interpreted appropriately in case $\lambda_i=\lambda_j$ or $\lambda_i=\mu$ or $\lambda_j=\mu$.

Proof. Fix n and take $A\in M_n(I)$ ($I=(-1,1)$) and choose a system of matrix units for M_n such that A is diagonal. Let $\mu\in I$ be arbitrary. Let $A'=\mathrm{diag}(\lambda_1,\dots,\lambda_n,\lambda_{n+1})$, where $\lambda_{n+1}=\mu$. Suppose f is operator convex. Then

$$f(A'+\delta\varepsilon H)=f((1-\delta)A'+\delta(A'+\varepsilon H))\le(1-\delta)f(A')+\delta f(A'+\varepsilon H).$$

Dividing by δ and letting $\delta\to 0$ gives

$$0\le f(A'+\varepsilon H)-f(A')-\varepsilon f^{[1]}(A')\circ_A H,$$

using Lemma 2.1. Lemma 2.3 then implies that the matrix whose (i,j)th entry is

$$(1)\qquad (\lambda_i-\lambda_j)^{-1}\sum_{k=1}^{n+1}\left\{\frac{f(\lambda_i)-f(\lambda_k)}{\lambda_i-\lambda_k}-\frac{f(\lambda_j)-f(\lambda_k)}{\lambda_j-\lambda_k}\right\}h_{ik}h_{kj}$$

is non-negative. Choose

$$H=\begin{pmatrix} 0 & 0 & \dots & 0 & \overline{\xi}_1 \\ 0 & 0 & \dots & 0 & \overline{\xi}_2 \\ \dots & \dots & \dots & \dots & \overline{\xi}_n \\ \xi_1 & \xi_2 & \dots & \xi_n & 0 \end{pmatrix}.$$

Then (1) becomes

$$(\lambda_i - \lambda_j)^{-1}\left\{\frac{f(\lambda_i) - f(\mu)}{\lambda_i - \mu} - \frac{f(\lambda_j) - f(\mu)}{\lambda_j - \mu}\right\}\bar{\xi}_i\xi_j.$$

Since the vector $(\xi_1, \xi_2, \dots, \xi_n)$ is arbitrary, it follows that $f^{[2]}(A;\mu) \geq 0$.

On the other hand, suppose that $f^{[2]}(A;\mu) \geq 0,\ \mu \in I$. In particular, $f^{[2]}(A;0) \geq 0$, which is $g^{[1]}(A) \geq 0$, where $g(x) = f(x)/x$ for $x \neq 0$ and $g(0) = f'(0)$. This implies, by using Lemma 2.1 of [1], that $g(x)$ is operator monotone. Consequently, $f(x)$ is operator convex by Lemma 3.1 of [1].

3. Functions of several variables. In this section, we provide characterizations of operator convexity for functions of two variables.

Consider a real-valued function of two real variables x and y in $I = (-1, 1)$. Assume that (i) $f(x, 0) = f(0, y) = 0$ for all x, y in I and (ii) the first partial derivatives and the mixed second partial derivative of f exist and are continuous. Fix $m, n \in \mathbb{N}$ and take A in $M_n(I)$ and B in $M_n(J)$. Let $\{e_{ij} : 1 \leq i, j \leq m\}$ and $\{f_{ij} : 1 \leq i, j \leq n\}$ be matrix units for M_m and M_n respectively such that $A = \sum_{i=1}^m \lambda_i e_{ii}$ and $B = \sum_{i=1}^n \mu_i f_{ii}$. Let $H \in M_m$ and $K \in M_n$ be arbitrary. We shall denote by $f^{[1,0]}(A, B)$, $f^{[0,1]}(A, B)$, $f^{[1,1]}(A, B)$, $f^{[2,0]}(A, B)$ and $f^{[0,2]}(A, B)$ the elements in $M_m \otimes M_n$ defined by:

$$f^{[1,0]}(A, B)|_{ij;kl} = (\lambda_i - \lambda_k)^{-1}\{f(\lambda_i, \mu_j) - f(\lambda_k, \mu_j)\}$$

if $\lambda_i \neq \lambda_k$, and equal to $\frac{\partial f}{\partial x}(\lambda_i, \mu_j)$ if $\lambda_i = \lambda_k$;

$$f^{[0,1]}(A, B)|_{ij;kl} = (\mu_j - \mu_l)^{-1}\{f(\lambda_i, \mu_j) - f(\lambda_i, \mu_l)\}$$

if $\mu_j \neq \mu_l$, and equal to $\frac{\partial f}{\partial y}(\lambda_i, \mu_j)$ if $\mu_j = \mu_l$;

$$f^{[1,1]}(A, B)|_{ij;kl} = \frac{f(\lambda_i, \mu_j) - f(\lambda_i, \mu_l) - f(\lambda_k, \mu_j) + f(\lambda_k, \mu_l)}{(\lambda_i - \lambda_k)(\mu_j - \mu_l)}$$

if $\lambda_i \neq \lambda_k$ and $\mu_j \neq \mu_l$, and the divided difference is to be interpreted appropriately when $\lambda_i = \lambda_k$ or $\mu_j = \mu_l$;

$$\begin{aligned} &f^{[2,0]}(A, B) \circ (H^2 \otimes I)|_{ij;kl} \\ &= (\lambda_i - \lambda_k)^{-1}\sum_\alpha \left\{\frac{f(\lambda_i, \mu_j) - f(\lambda_\alpha, \mu_j)}{\lambda_i - \lambda_\alpha} - \frac{f(\lambda_k, \mu_j) - f(\lambda_\alpha, \mu_j)}{\lambda_k - \lambda_\alpha}\right\} h_{i\alpha}h_{\alpha l}\delta_{\alpha l} \end{aligned}$$

if $\lambda_i \neq \lambda_k$, $\lambda_i \neq \lambda_\alpha$, $\lambda_k \neq \lambda_\alpha$, and the divided difference is to be interpreted appropriately when $\lambda_i = \lambda_k$ or $\lambda_i = \lambda_\alpha$ or $\lambda_k = \lambda_\alpha$; finally,

$$\begin{aligned} &f^{[0,2]}(A, B) \circ (I \otimes K^2)|_{ij;kl} \\ &= (\mu_j - \mu_l)^{-1}\sum_\alpha \left\{\frac{f(\lambda_i, \mu_j) - f(\lambda_i, \mu_\alpha)}{\mu_j - \mu_\alpha} - \frac{f(\lambda_i, \mu_l) - f(\lambda_i, \mu_\alpha)}{\mu_l - \mu_\alpha}\right\} k_{i\alpha}k_{\alpha l}\delta_{i\alpha} \end{aligned}$$

if $\mu_j \neq \mu_l$, $\mu_j \neq \mu_\alpha$, $\mu_l \neq \mu_\alpha$, and the divided difference is to be interpreted appropriately when $\mu_j = \mu_l$ or $\mu_j = \mu_\alpha$ or $\mu_l = \mu_\alpha$.

If $f(x,y) = g(x)h(y)$, the operator $f(A,B)$ coincides with $g(A) \otimes h(B)$ and the following formulae hold for the Hadamard product $\circ$ of matrices in a basis obtained from the basis that diagonalizes A and B:

$$\begin{aligned}
f^{[1,0]}(A,B) \circ H \otimes I &= (g^{[1]}(A) \circ_A H) \otimes h(B),\\
f^{[0,1]}(A,B) \circ I \otimes K &= g(A) \otimes (h^{[1]}(B) \circ_B K),\\
f^{[1,1]}(A,B) \circ H \otimes K &= (g^{[1]}(A) \circ_A H) \otimes (h^{[1]}(B) \circ_B K),\\
f^{[2,0]}(A,B) \circ H^2 \otimes I &= (g^{[2]}(A) \circ_A H^2) \otimes h(B),\\
f^{[0,2]}(A,B) \circ I \otimes K^2 &= g(A) \otimes (h^{[2]}(B) \circ_B K^2).
\end{aligned}$$

Since every f with the properties stipulated at the beginning of the section is the uniform limit of a sequence of linear combinations of such functions, the following lemma holds:

3.1. Lemma. *With f and A, B as above, we have*

(i) $\lim_{\varepsilon\to 0} \varepsilon^{-1}\{f(A+\varepsilon H, B) - f(A,B)\} = f^{[1,0]}(A,B) \circ H \otimes I$,

(ii) $\lim_{\varepsilon\to 0} \varepsilon^{-1}\{f(A, B+\varepsilon K) - f(A,B)\} = f^{[0,1]}(A,B) \circ I \otimes K$,

(iii) $\lim_{\varepsilon\to 0} \varepsilon^{-1}\{f(A+\varepsilon H, B+\varepsilon K) - f(A,B)\} = f^{[1,0]}(A,B) \circ H \otimes I + f^{[0,1]}(A,B) \circ I \otimes K$,

(iv) $\lim_{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}} \varepsilon_1^{-1}\varepsilon_2^{-1}\{f(A+\varepsilon_1 H, B+\varepsilon_2 K) - f(A+\varepsilon_1 H, B) - f(A, B+\varepsilon_2 K) + f(A,B)\} = f^{[1,1]}(A,B) \circ H \otimes K$,

(v) $\lim_{\varepsilon\to 0} \varepsilon^{-2}\{f(A+\varepsilon H, B) - f(A,B) - \varepsilon f^{[1,0]}(A,B) \circ H \otimes I\} = f^{[2,0]}(A,B) \circ H^2 \otimes I$,

(vi) $\lim_{\varepsilon\to 0} \varepsilon^{-2}\{f(A, B+\varepsilon K) - f(A,B) - \varepsilon f^{[0,1]}(A,B) \circ I \otimes K\} = f^{[0,2]}(A,B) \circ I \otimes K^2$,

for every $H \in M_m$ and $K \in M_n$.

Remark. $f^{[1,0]}(A,B)$, $f^{[0,1]}(A,B)$, $f^{[1,1]}(A,B)$, $f^{[2,0]}(A,B)$, and $f^{[0,2]}(A,B)$ resemble in many ways the appropriate derivatives of f. However, they depend on the basis considered for $M_m \otimes M_n$.

3.2. Theorem. *Let f be a real-valued function of two real variables x and y in $I = (-1,1)$. Assume that* (i) $f(x,0) = f(0,y) = 0$ *for all x, y in I and* (ii) *the first partial derivatives and the mixed second partial derivative of f exist and are continuous. Then the following statements are equivalent*:

(i) *f is operator convex on $I \times I$*,

(ii) $f(A,B) - f(A_0,B_0) \geq f^{[1,0]}(A_0,B_0) \circ (A - A_0) \otimes I$
$$+ f^{[0,1]}(A_0,B_0) \circ I \otimes (B - B_0)$$

for all $A_0, A \in M_m(I)$, $B_0, B \in M_n(I)$ *and for every* $m,n \in \mathbb{N}$, *and*

(iii) $f^{[1,0]}(A_2,B_2) \circ (A_2 - A_1) \otimes I + f^{[0,1]}(A_2,B_2) \circ I \otimes (B_2 - B_1)$
$$- f^{[1,0]}(A_1,B_1) \circ (A_2 - A_1) \otimes I - f^{[0,1]}(A_1,B_1) \circ I \otimes (B_2 - B_1) \geq 0$$

for all A_1, A_2 *in* $M_m(I)$, *all* B_1, B_2 *in* $M_n(I)$ *and all* $m,n \in \mathbb{N}$.

Proof. (i)$\Rightarrow$(ii). Fix m and n in $\mathbb{N}$ and take $A_0 \in M_m(I)$ and $B_0 \in M_n(I)$. Choose systems of matrix units for M_m and M_n that diagonalize A_0 and B_0. For $A \in M_m(I)$ and $B \in M_n(I)$ and $t \in [0,1]$, we have

$$\begin{aligned} f(A_0 + t(A - A_0), B_0 + t(B - B_0)) &= f((1-t)(A_0,B_0) + t(A,B)) \\ &\leq (1-t)f(A_0,B_0) + tf(A,B). \end{aligned}$$

This implies

$$\lim_{t\to 0} t^{-1}\{f(A_0+t(A-A_0), B_0+t(B-B_0)) - f(A_0,B_0)\} \leq f(A,B) - f(A_0,B_0),$$

i.e.

$$f(A,B) - f(A_0,B_0) \geq f^{[1,0]}(A_0,B_0) \circ (A-A_0) \otimes I + f^{[0,1]}(A_0,B_0) \circ I \otimes (B-B_0)$$

by Lemma 3.1(iii).

(ii)$\Rightarrow$(iii). Fix $m,n \in \mathbb{N}$. Let A_i $(i=1,2)$ be in $M_m(I)$ and B_i $(i=1,2)$ be in $M_n(I)$. Then

$$\begin{aligned} f(A_2,B_2) - f(A_1,B_1) \geq{} & f^{[1,0]}(A_1,B_1) \circ (A_2 - A_1) \otimes I \\ & + f^{[0,1]}(A_1,B_1) \circ I \otimes (B_2 - B_1) \end{aligned}$$

and

$$\begin{aligned} f(A_1,B_1) - f(A_2,B_2) \geq{} & f^{[1,0]}(A_2,B_2) \circ (A_1 - A_2) \otimes I \\ & + f^{[0,1]}(A_2,B_2) \circ I \otimes (B_1 - B_2). \end{aligned}$$

On adding the above inequalities, we get the desired result.

(iii)$\Rightarrow$(i). Let $\varphi : [0,1] \to M_m \otimes M_n$ be defined by

$$\begin{aligned} \varphi(t) &= f(t(A_2,B_2) + (1-t)(A_1,B_1)) \\ &= f(tA_2 + (1-t)A_1, tB_2 + (1-t)B_1). \end{aligned}$$

Let $t_1, t_2 \in [0,1]$ be such that $0 \leq t_1 < t_2 \leq 1$. Set

$$U_i = t_iA_2 + (1-t_i)A_1 \quad \text{and} \quad V_i = t_iB_2 + (1-t_i)B_1, \quad i = 1,2.$$

Then $U_2 - U_1 = (t_2 - t_1)(A_2 - A_1)$ and $V_2 - V_1 = (t_2 - t_1)(B_2 - B_1)$. The given condition then implies

$$\begin{aligned} & f^{[1,0]}(U_1,V_1) \circ (A_2 - A_1) \otimes I + f^{[0,1]}(U_1,V_1) \circ I \otimes (B_2 - B_1) \\ & \quad \leq f^{[1,0]}(U_2,V_2) \circ (A_2 - A_1) \otimes I + f^{[0,1]}(U_2,V_2) \circ I \otimes (B_2 - B_1). \end{aligned}$$

Observe that

$$\varphi'(t_i) = f^{[1,0]}(U_i, V_i) \circ (A_2 - A_1) \otimes I + f^{[0,1]}(U_i, V_i) \circ I \otimes (B_2 - B_1),$$

$i = 1, 2$, by Lemma 3.1(iii). The above inequality then becomes $\varphi'(t_1) \le \varphi'(t_2)$, i.e., φ' is increasing and hence φ is convex. Now,

$$\begin{aligned} f(t(A_2, B_2) + (1-t)(A_1, B_1)) &= \varphi(t) = \varphi(t \cdot 1 + (1-t) \cdot 0) \\ &\le t\varphi(1) + (1-t)\varphi(0) \\ &= tf(A_2, B_2) + (1-t)f(A_1, B_1). \ \blacksquare \end{aligned}$$

3.3. THEOREM. *Let $f \in C^2(I \times I)$, where $I = (-1, 1)$, be such that $f(x, 0) = f(0, y) = 0$ for x, y in I. Then f is operator convex on $I \times I$ iff the matrix*

$$f^{[2,0]}(A, B) \circ H^2 \otimes I + f^{[1,1]}(A, B) \circ H \otimes K + f^{[0,2]}(A, B) \circ I \otimes K^2 \ge 0$$

for A in $M_m(I)$ and B in $M_n(I)$.

Proof. Let $H \in M_m$ and $K \in M_n$ be arbitrary. Set $\varphi(t) = f(A + tH, B + tK)$. Then $\varphi(t)$ is a convex function of t in some neighbourhood of the origin. Since $f \in C^2(I \times I)$ then so also is φ, and $\varphi''(0) \ge 0$. But

$$\frac{\varphi''(0)}{2!} = \lim_{t \to 0} t^{-2}\{\varphi(t) - \varphi(0) - t\varphi'(0)\}.$$

Also

$$\begin{aligned} \varphi'(0) &= \lim_{t \to 0} t^{-1}\{\varphi(t) - \varphi(0)\} \\ &= \lim_{t \to 0} t^{-1}\{f(A + tH, B + tK) - f(A, B)\} \\ &= f^{[1,0]}(A, B) \circ H \otimes I + f^{[0,1]}(A, B) \circ I \otimes K, \end{aligned}$$

by Lemma 3.1(iii). Now,

$$\begin{aligned} t^{-2}\{f(A+tH, B+tK) &- f(A, B) - tf^{[1,0]}(A, B) \circ H \otimes I - tf^{[0,1]}(A, B) \circ I \otimes K\} \\ &= t^{-2}\{f(A + tH, B + tK) - f(A + tH, B) - f(A, B + tK) + f(A, B)\} \\ &\quad + t^{-2}\{f(A + tH, B) - f(A, B) - tf^{[1,0]}(A, B) \circ H \otimes I\} \\ &\quad + t^{-2}\{f(A, B + tK) - f(A, B) - tf^{[0,1]}(A, B) \circ I \otimes K\}. \end{aligned}$$

Consequently,

$$\frac{\varphi''(0)}{2!} = f^{[2,0]}(A, B) \circ H^2 \otimes I + f^{[1,1]}(A, B) \circ H \otimes K + f^{[0,2]}(A, B) \circ I \otimes K^2.$$

Thus, if f is operator convex, it then follows that the matrix

$$f^{[2,0]}(A, B) \circ H^2 \otimes I + f^{[1,1]}(A, B) \circ H \otimes K + f^{[0,2]}(A, B) \circ I \otimes K^2$$

is non-negative.

Conversely, if the condition is satisfied then $\varphi(t)$ is convex and hence f is operator convex since its restriction to any line segment in $M_m(I) \times M_n(I)$ is operator convex. ∎

4. Operator monotonicity of functions of two variables. Lemmas 3.1(iv) and 4.1 provide an alternative proof of a theorem of Korányi [5, Th. 4] and a theorem of Vasudeva [9, Th. 3]—a complete analogue to Löwner's theorem for functions of two variables and its finite-dimensional version respectively.

4.1. LEMMA. *If $s \to A(s)$ [resp. $t \to B(t)$] is a C^1 function from $[0,1]$ to the space of $m \times m$ matrices [resp. $n \times n$ matrices] with spectrum in $I = (-1,1)$ and if $f \in C^1(I \times I)$, then*

$$f(A(1),B(1)) - f(A(1),B(0)) - f(A(0),B(1)) + f(A(0),B(0)) = \int_0^1 \int_0^1 f^{[1,1]}(A(s),B(t)) \circ \left(\frac{dA(s)}{ds} \otimes \frac{dB(t)}{dt}\right) ds\, dt.$$

P r o o f. From Lemma 3.1(iv) we observe that

$$f(A(s'),B(t')) - f(A(s'),B(t)) - f(A(s),B(t')) + f(A(s),B(t)) = f^{[1,1]}(A(s),B(t)) \circ [(A(s') - A(s)) \otimes (B(t') - B(t))] + \theta(s',s,t',t),$$

where $\theta(s',s,t',t) \to 0$ as $|s'-s| \to 0$, $|t'-t| \to 0$. Choose $\varepsilon > 0$. Then there exist integers $m', n' \in \mathbb{N}$ such that $\|\theta(s',s,t',t)\| < \varepsilon$ whenever $|s'-s| < 1/m'$, $|t'-t| < 1/n'$. With $A_k = A(k/m')$, $0 \le k \le m'$, and $B_l = B(l/n')$, $0 \le l \le n'$, we therefore have

$$\begin{aligned} f(A(1),B(1)) &- f(A(1),B(0)) - f(A(0),B(1)) + f(A(0),B(0)) \\ &= \sum_{k,l} [f(A_{k+1},B_{l+1}) - f(A_{k+1},B_l) - f(A_k,B_{l+1}) + f(A_k,B_l)] \\ &= \sum_{k,l} \{f^{[1,1]}(A_k,B_k) \circ [(A_{k+1} - A_k) \otimes (B_{k+1} - B_k)] \\ &\quad + \theta((k+1)/m', k/m', (l+1)/n', l/n')\}. \end{aligned}$$

The first part of the sum converges to the Riemann integral

$$\int_0^1 \int_0^1 f^{[1,1]}(A(s),B(t)) \circ \left(\frac{dA(s)}{ds} \otimes \frac{dB(t)}{dt}\right) ds\, dt$$

as the area of the mesh of the subdivision tends to zero. The second term of the sum is less than the preassigned positive number ε. Hence the result follows. ∎

4.2. THEOREM. *Let f be a real-valued function of two real variables x and y in $I = (-1, 1)$. Assume that* (i) $f(x, 0) = f(0, y) = 0$ *for all x, y in I and* (ii) *the first partial derivatives and the mixed second partial derivative of f exist and are continuous. Then f is a monotone operator function of two variables iff $f^{[1,1]}(A, B) \geq 0$ for every $A \in M_m(I)$ and $B \in M_n(I)$ and for $m, n \in \mathbb{N}$.*

Proof. Fix $m, n \in \mathbb{N}$ and let $A = \sum \lambda_i e_{ii}$ and $B = \sum \mu_i f_{ii}$, where $\{e_{ij} : 1 \leq i, j \leq m\}$ and $\{f_{ij} : 1 \leq i, j \leq n\}$ are matrix units for M_m and M_n respectively. Suppose f is operator monotone. Choose $H = \sum e_{ij}$ and $K = \sum f_{ij}$. Then $m^{-1}H$, $n^{-1}K$ are one-dimensional projections. Consequently,

$$0 \leq \varepsilon_1^{-1}\varepsilon_2^{-1}\{f(A+\varepsilon_1 H, B+\varepsilon_2 K)-f(A+\varepsilon_1 H, B)-f(A, B+\varepsilon_2 K)+f(A, B)\},$$

whence $f^{[1,1]}(A, B) \geq 0$, by Lemma 3.1(iv), since $H \otimes K$ is the unit for the Hadamard product.

Conversely, suppose that $f^{[1,1]}(A, B) \geq 0$. Choose $A' \in M_m(I)$ and $B' \in M_n(I)$ such that $A' \geq A$ and $B' \geq B$. Set $A(s) = (1 - s)A + sA'$ and $B(t) = (1 - t)B + tB'$. Then

$$\frac{dA(s)}{ds} = A' - A \geq 0, \qquad \frac{dB(t)}{dt} = B' - B \geq 0.$$

Consequently, using Lemma 4.1, we have

$$\begin{aligned} f(A', B') &- f(A', B) - f(A, B') + f(A, B) \\ &= \int_0^1 \int_0^1 f^{[1,1]}(A(s), B(t)) \circ (A' - A) \otimes (B' - B)\, ds\, dt \geq 0, \end{aligned}$$

because the Hadamard product of positive matrices is again positive.

Acknowledgements. The authors would like to thank Prof. Ajit Iqbal Singh (née Ajit Kaur Chilana) for useful suggestions.

References

[1] T. Ando, *Topics on Operator Inequalities*, lecture notes (mimeographed), Hokkaido University, Sapporo, 1978.

[2] C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, in: Proc. Sympos. Pure Math. 7, Amer. Math. Soc., 1963, 187–201.

[3] W. F. Donoghue, Jr., *Monotone Matrix Functions and Analytic Continuation*, Springer, Heidelberg, 1974.

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

[5] A. Korányi, *On a class of analytic functions of several variables*, Trans. Amer. Math. Soc. 101 (1961), 521–554.

[6] F. Kraus, *Über konvexe Matrixfunktionen*, Math. Z. 41 (1936), 18–41.

[7] C. Löwner, *Über monotone Matrixfunktionen*, ibid. 38 (1934), 177–216.

[8] A. W. Roberts and D. E. Varberg, *Convex Functions*, Academic Press, New York, 1973.

[9] H. Vasudeva, *On monotone matrix functions of two variables*, Trans. Amer. Math. Soc. 176 (1973), 303–318.

DEPARTMENT OF APPLIED MATHEMATICS
REGIONAL ENGINEERING COLLEGE
JALANDHAR, PUNJAB, INDIA

DEPARTMENT OF MATHEMATICS
PANJAB UNIVERSITY
CHANDIGARH, INDIA

Reçu par la Rédaction le 20.5.1993