Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER
Linear Algebra and its Applications 369 (2003) 217–233
www.elsevier.com/locate/laa

Weak majorization inequalities and convex functions

Jaspal Singh Aujla [a,*], Fernando C. Silva [b]

[a] *Department of Applied Mathematics, National Institute of Technology, Jalandhar 144011, Punjab, India*
[b] *Center for Linear Structures and Combinatorics, University of Lisbon, Av. Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal*

Received 30 September 2002; accepted 10 December 2002

Submitted by R. Bhatia

Abstract

Let f be a convex function defined on an interval I, $0 \leqslant \alpha \leqslant 1$ and A, B $n \times n$ complex Hermitian matrices with spectrum in I. We prove that the eigenvalues of $f(\alpha A + (1-\alpha)B)$ are weakly majorized by the eigenvalues of $\alpha f(A) + (1-\alpha) f(B)$. Further if f is log convex we prove that the eigenvalues of $f(\alpha A + (1-\alpha)B)$ are weakly majorized by the eigenvalues of $f(A)^{\alpha} f(B)^{1-\alpha}$. As applications we obtain generalizations of the famous Golden–Thompson trace inequality, a representation theorem and a harmonic–geometric mean inequality. Some related inequalities are discussed.

AMS classification: 47A30; 47B15; 15A60

Keywords: Convex function; Weak majorization; Unitarily invariant norm

1. Introduction

Throughout $\mathscr{M}_n$ denotes the set of $n \times n$ complex matrices and $\mathscr{H}_n$ denotes the set of all Hermitian matrices in $\mathscr{M}_n$. We denote by $\mathscr{S}_n$, the set of all positive semidefinite matrices in $\mathscr{M}_n$. The set of all positive definite matrices in $\mathscr{M}_n$ is denoted by

* Corresponding author. Current address: Center for Linear Structures and Combinatorics, University of Lisbon, Av. Prof. Gama Pinto 2, 1649-003 Lisboa, Portugal.
E-mail address: aujlajs@yahoo.com (J.S. Aujla).

0024-3795/03/$ - see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0024-3795(02)00720-6

$\mathscr{P}_n$. Let I be an interval in $\mathscr{R}$. We denote by $\mathscr{H}_n(I)$, the set of all Hermitian matrices in $\mathscr{M}_n$ whose spectrum is contained in I.

Let f be a real valued function defined on I. The function f is called convex if

$$f(\alpha s + (1-\alpha)t) \leqslant \alpha f(s) + (1-\alpha) f(t)$$

for all $0 \leqslant \alpha \leqslant 1$ and $s, t \in I$. Likewise f is called concave if $-f$ is convex. Further if f is positive then f is called log convex if

$$f(\alpha s + (1-\alpha)t) \leqslant f(s)^{\alpha} f(t)^{1-\alpha}$$

and is called log concave if

$$f(s)^{\alpha} f(t)^{1-\alpha} \leqslant f(\alpha s + (1-\alpha)t).$$

If $I = (0, \infty)$ and f is positive then f is called multiplicativily convex if

$$f(s^{\alpha} t^{1-\alpha}) \leqslant f(s)^{\alpha} f(t)^{1-\alpha}$$

for all $0 \leqslant \alpha \leqslant 1$ and $s, t \in I$.

The reader is referred to [11] for general properties of convex and log convex functions. If f is multiplicativily convex then the function $t \to f(e^t)$ is log convex on $(-\infty, \infty)$. The functions exp, sinh, cosh are multiplicativily convex. For more examples and properties of multiplicativily convex functions the reader is referred to [10].

A norm $|||\cdot|||$ on $\mathscr{M}_n$ is called unitarily invariant or symmetric if

$$|||UAV||| = |||A|||$$

for all $A \in \mathscr{M}_n$ and for all unitaries $U, V \in \mathscr{M}_n$. The most basic unitarily invariant norms are the Ky Fan norms $\|\cdot\|_{(k)}$, $(k = 1, 2, \ldots, n)$, defined as

$$\|A\|_{(k)} = \sum_{j=1}^{k} s_j(A) \quad (k = 1, 2, \ldots, n)$$

and the Schatten p-norms defined as

$$\|A\|_p = \left(\sum_{j=1}^{n} (s_j(A))^p \right)^{1/p} \qquad 1 \leqslant p < \infty,$$

where $s_1(A) \geqslant s_2(A) \geqslant \cdots \geqslant s_n(A)$ are the singular values of A, that is, the eigenvalues of $|A| = (A^*A)^{1/2}$. It is customary to assume a normalization condition that $|||\mathrm{diag}(1, 0, \ldots, 0)||| = 1$. The spectral norm (or operator norm) is given by $\|A\| = s_1(A)$. An $A \in \mathscr{M}_n$ is called a contraction if $\|A\| \leqslant 1$.

Throughout $|||\cdot|||$ denotes an arbitrary unitarily invariant norm on $\mathscr{M}_n$. For (column) vectors $x, y \in \mathscr{C}^n$ their inner product is denoted by $\langle x, y\rangle = y^*x$. For an $A \in \mathscr{H}_n$, $\lambda_j(A)$, $1 \leqslant j \leqslant n$ denote the eigenvalues of A arranged in the decreasing order. We use the notation $\lambda(A)$ to denote the row vector $(\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A))$. We then define

$$\lambda(A) \circ \lambda(B) = (\lambda_1(A)\lambda_1(B), \lambda_2(A)\lambda_2(B), \ldots, \lambda_n(A)\lambda_n(B)).$$

Let $A \in \mathscr{H}_n(I)$ have spectral decomposition

$$A = U^*\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)U,$$

where U is a unitary and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. Let f be a real valued function defined on I. Then $f(A)$ is defined by

$$f(A) = U^* \text{diag}(f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n))U.$$

For $A, B \in \mathscr{H}_n$ we consider three kinds of ordering:

(i) $B \leqslant A$ (or $A \geqslant B$) $\overset{\text{def}}{\Longleftrightarrow}$ $A - B$ positive semidefinite,

(ii) (eigenvalue inequalities)

$$\lambda(B) \leqslant \lambda(A) \overset{\text{def}}{\Longleftrightarrow} \lambda_j(B) \leqslant \lambda_j(A) \quad (j = 1, 2, \ldots, n),$$

(iii) (weak majorization)

$$\lambda(B) \prec_w \lambda(A) \overset{\text{def}}{\Longleftrightarrow} \sum_{j=1}^{k} \lambda_j(B) \leqslant \sum_{j=1}^{k} \lambda_j(A) \quad (k = 1, 2, \ldots, n).$$

We can see

$$B \leqslant A \Longrightarrow \lambda(B) \leqslant \lambda(A) \Longrightarrow \lambda(B) \prec_w \lambda(A).$$

For f increasing on I, $A, B \in \mathscr{H}_n(I)$, $\lambda(B) \leqslant \lambda(A) \Longrightarrow \lambda(f(B)) \leqslant \lambda(f(A))$.

For f increasing and convex on I, $A, B \in \mathscr{H}_n(I)$, $\lambda(B) \prec_w \lambda(A) \Longrightarrow \lambda(f(B)) \prec_w \lambda(f(A))$.

A function f on I is called operator convex if

$$f(\alpha A + (1-\alpha)B) \leqslant \alpha f(A) + (1-\alpha)f(B)$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$. Thus if the function f is operator convex we have the inequalities at the strongest level (i).

The purpose of this paper is to show that if we replace operator convexity by mere convexity we get weak majorization inequalities of kind (iii). If in addition the function f is also increasing (or decreasing) we get eigenvalue inequalities of kind (ii). Similar inequalities are proved for log convex functions. These include a number of known inequalities.

2. Convex functions

The following lemmas will be used to prove the main results in this section. The reader may refer to [4] for their proofs.

Lemma 2.1 [4, p. 281]. *Let $A \in \mathscr{H}_n(I)$ and let f be a convex function on I. Then for every unit vector $x \in \mathscr{C}^n$,*

$$f(\langle Ax, x\rangle) \leqslant \langle f(A)x, x\rangle.$$

Lemma 2.2 [4, p. 35]. *Let $A \in \mathscr{H}_n$. Then*

$$\sum_{j=1}^{k} \lambda_j(A) = \max \sum_{j=1}^{k} \langle Au_j, u_j\rangle \quad (k = 1, 2, \ldots, n),$$

where the maximum is taken over all choices of orthonormal vectors $u_1, u_2, \ldots, u_k$.

Theorem 2.3. *Let f be a convex function on I. Then*

$$\lambda(f(\alpha A + (1-\alpha)B)) \prec_w \lambda(\alpha f(A) + (1-\alpha)f(B))$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$. If further $0 \in I$ and $f(0) \leqslant 0$ then

$$\lambda(f(X^*AX)) \prec_w \lambda(X^* f(A)X)$$

for all $A \in \mathscr{H}_n(I)$ and contractions $X \in \mathscr{M}_n$.

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of $\alpha A + (1-\alpha)B$ and let $u_1, u_2, \ldots, u_n$ be the corresponding orthonormal eigenvectors arranged such that $f(\lambda_1) \geqslant f(\lambda_2) \geqslant \cdots \geqslant f(\lambda_n)$. Let $k = 1, 2, \ldots, n$. Then

$$\begin{aligned}
\sum_{j=1}^{k} \lambda_j(f(\alpha A + (1-\alpha)B)) &= \sum_{j=1}^{k} f(\langle(\alpha A + (1-\alpha)B)u_j, u_j\rangle) \\
&= \sum_{j=1}^{k} f(\alpha\langle Au_j, u_j\rangle + (1-\alpha)\langle Bu_j, u_j\rangle) \\
&\leqslant \sum_{j=1}^{k} [\alpha f(\langle Au_j, u_j\rangle) + (1-\alpha) f(\langle Bu_j, u_j\rangle)] \\
&\leqslant \sum_{j=1}^{k} [\alpha\langle f(A)u_j, u_j\rangle + (1-\alpha)\langle f(B)u_j, u_j\rangle] \\
&= \sum_{j=1}^{k} \langle(\alpha f(A) + (1-\alpha)f(B))u_j, u_j\rangle \\
&\leqslant \sum_{j=1}^{k} \lambda_j(\alpha f(A) + (1-\alpha)f(B)),
\end{aligned}$$

using convexity of f, Lemmas 2.1 and Lemma 2.2 respectively. This proves the first assertion. To prove the second assertion, let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of X^*AX and let $u_1, u_2, \ldots, u_n$ be the corresponding orthonormal eigenvectors arranged such that $f(\lambda_1) \geqslant f(\lambda_2) \geqslant \cdots \geqslant f(\lambda_n)$. Since $f(0) \leqslant 0$, to prove the desired inequality we can assume that $\|Xu_j\| \neq 0,\ j = 1, 2, \ldots, n$. Then

$$\begin{aligned}\sum_{j=1}^{k}\lambda_j(f(X^*AX)) &= \sum_{j=1}^{k} f(\langle X^*AXu_j, u_j\rangle)\\ &= \sum_{j=1}^{k} f\left(\|Xu_j\|^2\left\langle A\frac{Xu_j}{\|Xu_j\|}, \frac{Xu_j}{\|Xu_j\|}\right\rangle\right.\\ &\qquad \left. + (1-\|Xu_j\|^2)\cdot 0\right)\\ &\leqslant \sum_{j=1}^{k}\left(\|Xu_j\|^2 f\left(\left\langle A\frac{Xu_j}{\|Xu_j\|}, \frac{Xu_j}{\|Xu_j\|}\right\rangle\right)\right.\\ &\qquad \left. + (1-\|Xu_j\|^2) f(0)\right)\\ &\leqslant \sum_{j=1}^{k}\left(\|Xu_j\|^2\left\langle f(A)\frac{Xu_j}{\|Xu_j\|}, \frac{Xu_j}{\|Xu_j\|}\right\rangle\right)\\ &= \sum_{j=1}^{k}\langle X^*f(A)Xu_j, u_j\rangle\\ &\leqslant \sum_{j=1}^{k}\lambda_j(X^*f(A)X),\end{aligned}$$

using convexity of f, the condition $f(0) \leqslant 0$, Lemmas 2.1 and 2.2 respectively. This completes the proof. $\square$

The following result is proved in [3]. This follows from Theorem 2.3 taking $f(t) = t^r$, $r \leqslant 0$ and $I = (0, \infty)$.

Corollary 2.4. *Let $A, B \in \mathscr{P}_n$. Then*

$$\lambda(2^{1-r}(A+B)^r) \prec_w \lambda(A^r + B^r)$$

for all $r \leqslant 0$.

Every nonnegative decreasing function f on $[0, \infty)$ satisfies $f(2t) \leqslant 2f(t)$, $t \in [0, \infty)$. The next corollary gives an inequality similar to the inequalities proved in [2] for operator monotone functions.

Corollary 2.5. *Let f be a convex function on $[0, \infty)$ such that $f(2t) \leqslant 2f(t)$ for all $t \in [0, \infty)$. Then*

$$\lambda(f(A+B)) \prec_w \lambda(f(A) + f(B))$$

for all $A, B \in \mathscr{S}_n$.

Proof. By Theorem 2.3, we have

$$\lambda\left(f\left(\frac{A+B}{2}\right)\right) \prec_w \lambda\left(\frac{f(A)+f(B)}{2}\right).$$

Now on replacing A by $2A$ and B by $2B$ in the above inequality, we get

$$\lambda(f(A+B)) \prec_w \lambda\left(\frac{f(2A)+f(2B)}{2}\right). \tag{1}$$

The condition $f(2t) \leqslant 2f(t)$ implies $f(2A) \leqslant 2f(A)$ and $f(2B) \leqslant 2f(B)$. Therefore

$$\lambda\left(\frac{f(2A)+f(2B)}{2}\right) \prec_w \lambda(f(A)+f(B)). \tag{2}$$

Now (1) and (2) give the desired result. □

The following corollary follows on using the Fan Dominance Theorem [4, p. 93].

Corollary 2.6. *Let f be a nonnegative convex function on I. Then*

$$|||f(\alpha A+(1-\alpha)B)||| \leqslant |||\alpha f(A)+(1-\alpha)f(B)|||$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$. If further $0 \in I$ and $f(0)=0$ then

$$|||f(X^*AX)||| \leqslant |||X^*f(A)X|||$$

for all $A \in \mathscr{H}_n(I)$ and contractions $X \in \mathscr{M}_n$.

Remark 2.7. Corollary 2.6 may not be true if f is not nonnegative. To see this one may take $f(t) = -\log t$.

Remark 2.8. For $A, B \in \mathscr{H}_n$, the inequality (see [4, p. 294])

$$|||(A-B)^{2m+1}||| \leqslant 2^{2m}|||A^{2m+1}-B^{2m+1}|||$$

is equivalent to

$$|||(A+B)^{2m+1}||| \leqslant 2^{2m}|||A^{2m+1}+B^{2m+1}|||, \quad m=1,2,\ldots$$

Choosing the nonnegative convex function $f(t)=|t|^r$, $r \geqslant 1$, on $(-\infty,\infty)$. Corollary 2.6 provides an analogue of the above inequality,

$$|||\,|A+B|^r\,||| \leqslant 2^{r-1}|||\,|A|^r+|B|^r|||, \quad r \geqslant 1.$$

Another particular case of Corollary 2.6 when $f(t)=t^r$, $r \geqslant 1$ is Theorem 1 in [7].

If in addition, in Theorem 2.3 we assume that f is increasing (or decreasing) we have the following stronger result.

Theorem 2.9. *Let f be an increasing (or decreasing) convex function on I. Then*

$$\lambda(f(\alpha A+(1-\alpha)B)) \leqslant \lambda(\alpha f(A)+(1-\alpha)f(B))$$

for all $A, B \in \mathscr{H}_n(I)$ *and* $0 \leqslant \alpha \leqslant 1$. *If, in addition*, $0 \in I$ *and* $f(0) \leqslant 0$, *then*

$$\lambda(f(X^*AX)) \leqslant \lambda(X^* f(A)X)$$

for all $A \in \mathscr{H}_n(I)$ *and contractions* $X \in \mathscr{M}_n$.

Proof. Since f is increasing, for any $H \in \mathscr{H}_n(I)$

$$\lambda_j(f(H)) = f(\lambda_j(H)) \quad (j = 1, 2, \ldots, n).$$

It is known [4, p. 58] that the eigenvalue $\lambda_j(H)$ admits the following max–min characterization:

$$\lambda_j(H) = \max_{\dim \mathscr{M}=j} \min\{\langle Hx, x\rangle;\ \|x\| = 1, x \in \mathscr{M}\} \tag{3}$$

where $\mathscr{M}$ is a subspace of $\mathscr{C}^n$. Then since f is increasing

$$\begin{aligned}\lambda_j(f(H)) = f(\lambda_j(H)) &= f\left(\max_{\dim \mathscr{M}=j} \min\{\langle Hx, x\rangle;\ \|x\| = 1, x \in \mathscr{M}\}\right)\\ &= \max_{\dim \mathscr{M}=j} \min\{f(\langle Hx, x\rangle);\ \|x\| = 1, x \in \mathscr{M}\}.\end{aligned}$$

Applying this to $H = \alpha A + (1-\alpha)B$ we have

$$\begin{aligned}\lambda_j(f(\alpha A + (1-\alpha)B)) = \max_{\dim \mathscr{M}=j} \min\{&f(\langle(\alpha A + (1-\alpha)B)x, x\rangle);\\ &\|x\| = 1, x \in \mathscr{M}\}.\end{aligned}$$

By convexity of f and Lemma 2.1, we get

$$\begin{aligned}f(\langle(\alpha A + (1-\alpha)B)x, x\rangle) &= f(\alpha\langle Ax, x\rangle + (1-\alpha)\langle Bx, x\rangle)\\ &\leqslant \alpha f(\langle Ax, x\rangle) + (1-\alpha) f(\langle Bx, x\rangle)\\ &\leqslant \langle(\alpha f(A) + (1-\alpha) f(B))x, x\rangle, \quad (\|x\| = 1).\end{aligned}$$

Now using formula (3), we have

$$\lambda_j\,(f(\alpha A + (1-\alpha)B)) \leqslant \lambda_j\,(\alpha f(A) + (1-\alpha) f(B))\,.$$

This completes the proof of the first assertion. The proof of the second assertion is similar. □

Remark 2.10. Theorem 2.9 may not be true if f is not increasing (or decreasing). To see this one may take $f(t) = |t|$, $t \in (-\infty, \infty)$,

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Remark 2.11. Ando and Zhan [2] proved that

$$\lambda(A^r + B^r) \prec_w \lambda((A+B)^r), \quad r \geqslant 1, \tag{4}$$

$$\lambda((A+B)^r) \prec_w \lambda(A^r + B^r), \quad 0 \leqslant r \leqslant 1, \tag{5}$$

for $A, B \in \mathscr{S}_n$. Taking the convex function $f(t) = t^r$, $r \geqslant 1$ in Theorem 2.9, we get

$$\lambda((A+B)^r) \leqslant \lambda(2^{r-1}(A^r + B^r)),$$

which in turn gives a sharp upper bound for inequality (4). Now let $0 \leqslant r \leqslant 1$. Applying Theorem 2.9 to the decreasing convex function $g(t) = -t^r$, we get

$$\lambda(2^{r-1}(A^r + B^r)) \leqslant \lambda((A+B)^r).$$

This provides a sharp lower bound for inequality (5). Taking the decreasing convex function $f(t) = t^r$, $r \leqslant 0$ in Theorem 2.9, we get

$$\lambda(2^{1-r}(A+B)^r) \leqslant \lambda(A^r + B^r),$$

which gives a stronger result than Corollary 2.4.

Theorem 2.12. *Let f be a nonnegative continuous function on $[0, \infty)$. Then f is increasing and convex with $f(0) = 0$ if and only if*

$$2^{1/p-1}\|f(A) + f(B)\|_p \leqslant \|f(A+B)\|_p$$

for all $A, B \in \mathscr{S}_n$ and $p \geqslant 1$.

Proof. First suppose f is increasing and convex with $f(0) = 0$. Then by [4, Theorem IV.2.13] and the Fan Dominance Theorem we see that

$$\frac{1}{2}\left|\left|\left|\begin{pmatrix} f(A)+f(B) & O \\ O & f(A)+f(B) \end{pmatrix}\right|\right|\right| \leqslant \left|\left|\left|\begin{pmatrix} f(A+B) & O \\ O & O \end{pmatrix}\right|\right|\right|.$$

This implies

$$2^{1/p-1}\left[\sum_{j=1}^{n}\left(\lambda_j(f(A)+f(B))\right)^p\right]^{1/p} \leqslant \left[\sum_{j=1}^{n}\left(\lambda_j(f(A+B))\right)^p\right]^{1/p}.$$

Thus

$$2^{1/p-1}\|f(A) + f(B)\|_p \leqslant \|f(A+B)\|_p.$$

The converse follows (using the given inequality for $p = 1$) as in [3]. This completes the proof. $\square$

Remark 2.13. Since a nonnegative decreasing function f satisfies $f(2t) \leqslant 2f(t)$ by Corollary 2.5, we have

$$\lambda(f(A+B)) \prec_w \lambda(f(A) + f(B)),$$

if f is also convex. Thus one might conjecture that for any nonnegative increasing convex function f on $[0, \infty)$ with $f(0) = 0$,

$$\lambda(f(A) + f(B)) \prec_w \lambda(f(A+B)). \tag{6}$$

The inequality (4) supports this when $f(t) = t^r$, $r \geqslant 1$. Our next theorem for which we need the following lemmas further strengthens this belief.

Lemma 2.14 [4, p. 54]. *Let $A, B, C \in \mathscr{S}_n$. Then*

$$\lambda(A) \prec_w \lambda(B)$$

implies

$$\lambda(A) \circ \lambda(C) \prec_w \lambda(B) \circ \lambda(C).$$

Lemma 2.15. *Let $A, B, C \in \mathscr{S}_n$ be such that*

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix} \geqslant O.$$

Then

$$\lambda(C^2) \prec_w \lambda(A) \circ \lambda(B).$$

Proof. The positive semidefiniteness of the given matrix implies that there exists a contraction $K \in \mathscr{M}_n$ (see [1, p. 13]) such that

$$C = A^{1/2} K B^{1/2}.$$

Then using a standard argument with antisymmetric tensor products as in [4, p. 94], we get the desired inequality. □

Theorem 2.16. *Let f, g be nonnegative* (*continuous*) *functions on* $[0, \infty)$ *which satisfy inequality* (6). *Then the functions $f + g$, $f \circ g$ and fg satisfy inequality* (6).

Proof. By Theorem 2.12 f, g are increasing and convex. Let $k = 1, 2, \ldots, n$. Then

$$\begin{aligned}
\sum_{j=1}^{k} \lambda_j((f+g)(A) + (f+g)(B)) &\leqslant \sum_{j=1}^{k} \lambda_j(f(A) + f(B)) \\
&\quad + \sum_{j=1}^{k} \lambda_j(g(A) + g(B)) \\
&\leqslant \sum_{j=1}^{k} \lambda_j(f(A+B)) + \sum_{j=1}^{k} \lambda_j(g(A+B)) \\
&= \sum_{j=1}^{k} \left[\lambda_j \left(f(A+B) + g(A+B)\right)\right],
\end{aligned}$$

using the fact that f and g are increasing. This proves the result for $f + g$. To prove the result for $f \circ g$, note that

$$\lambda(f \circ g(A) + f \circ g(B)) \prec_w \lambda(f(g(A) + g(B))).$$

Since f is increasing and convex, the inequality

$$\lambda(g(A) + g(B)) \prec_w \lambda(g(A + B))$$

implies

$$\lambda(f(g(A) + g(B))) \prec_w \lambda(f \circ g(A + B))$$

and hence

$$\lambda(f \circ g(A) + f \circ g(B)) \prec_w \lambda(f \circ g(A + B)).$$

This completes a proof of the second assertion. Next note that the inequality

$$\lambda(f(A) + f(B)) \prec_w \lambda(f(A + B))$$

and Lemma 2.14 imply

$$\begin{aligned} \lambda(f(A) + f(B)) \circ \lambda(g(A + B)) &\prec_w \lambda(f(A + B)) \circ \lambda(g(A + B)) \\ &= \lambda(f(A + B)g(A + B)). \end{aligned} \tag{7}$$

Again we have

$$\lambda(f(A) + f(B)) \circ \lambda(g(A) + g(B)) \prec_w \lambda(f(A) + f(B)) \circ \lambda(g(A + B)). \tag{8}$$

Therefore inequalities (7) and (8) give

$$\lambda(f(A) + f(B)) \circ \lambda(g(A) + g(B)) \prec_w \lambda(f(A + B)g(A + B)). \tag{9}$$

Now observe that the matrices

$$\begin{pmatrix} f(A) & (f(A)g(A))^{1/2} \\ (f(A)g(A))^{1/2} & g(A) \end{pmatrix}$$

and

$$\begin{pmatrix} f(B) & (f(B)g(B))^{1/2} \\ (f(B)g(B))^{1/2} & g(B) \end{pmatrix}$$

are positive semidefinite. Hence the matrix

$$\begin{pmatrix} f(A) + f(B) & (f(A)g(A))^{1/2} + (f(B)g(B))^{1/2} \\ (f(A)g(A))^{1/2} + (f(B)g(B))^{1/2} & g(A) + g(B) \end{pmatrix}$$

is positive semidefinite. Therefore by Lemma 2.15

$$\begin{aligned} &\lambda\left(\left[(f(A)g(A))^{1/2} + (f(B)g(B))^{1/2}\right]^2\right) \\ &\quad \prec_w \lambda(f(A) + f(B)) \circ \lambda(g(A) + g(B)). \end{aligned} \tag{10}$$

Thus from (9) and (10), we get

$$\lambda\left(\left[(f(A)g(A))^{1/2}+(f(B)g(B))^{1/2}\right]^2\right)\prec_w \lambda(f(A+B)g(A+B)).$$

Then inequality (4) for $r=2$ and the above inequality imply

$$\lambda(f(A)g(A)+f(B)g(B))\prec_w \lambda(f(A+B)g(A+B)).$$

This completes the proof. □

Corollary 2.17. *Let $p(t)$ be a polynomial (or a power series) in $t\in[0,\infty)$ with nonnegative coefficients and $p(0)=0$. Then*

$$\lambda(p(A)+p(B))\prec_w \lambda(p(A+B))$$

for all $A, B\in\mathscr{S}_n$.

Proof. Applying Theorem 2.16 with suitable f, g repeatedly, the inequality

$$\lambda(A+B)\prec_w \lambda(A+B)$$

implies

$$\lambda(A^m+B^m)\prec_w \lambda((A+B)^m),\quad m=1,2,\ldots$$

Hence the result follows on applying Theorem 2.16 again repeatedly. □

Remark 2.18. The weak majorization

$$\lambda(A^m+B^m)\prec_w \lambda((A+B)^m),\quad m=1,2,\ldots,$$

has been proved in [5]. This inequality is a special case of inequality (4). From Corollary 2.17 we also see that

$$\lambda(e^A+e^B-2I)\prec_w \lambda(e^{A+B}-I).$$

This has been proved in [2]. We further remark that if the inequality (6) holds for the nonnegative functions $h_\lambda(t)$, $\lambda, t\geqslant 0$, then it holds for all functions $f(t)$ given by

$$f(t)=\int_0^\infty h_\lambda(t)\,\mathrm{d}\mu(\lambda),$$

where μ is a positive measure on $[0,\infty)$. The same is true for a function f on $[0,\infty)$ with $f(0)=0$ if it is the limit of polynomials $p_k(t)$, $k=1,2,\ldots$ with nonnegative coefficients and $p_k(0)=0$.

Thus we make the following conjecture.

Conjecture 2.19. *If f is a nonnegative increasing convex function on $[0,\infty)$ with $f(0)=0$, then*

$$\lambda(f(A)+f(B))\prec_w \lambda(f(A+B))$$

for all $A, B\in\mathscr{S}_n$.

3. Log convex functions

We begin this section with some lemmas. For a proof of the following two lemmas the reader is referred to [1].

Lemma 3.1 [1, p. 56]. *Let $A, B \in \mathscr{P}_n$ and $0 < r < 1$. Then*

$$\lambda\left(\frac{1}{r}\log(A^{r/2}B^rA^{r/2})\right) \prec_w \lambda(\log(A^{1/2}BA^{1/2})).$$

The following lemma is known as Trotter's formula.

Lemma 3.2 [1, p. 57]. *Let $A, B \in \mathscr{P}_n$. Then*

$$\lim_{r\to 0+}\left[\frac{1}{r}\log(A^{r/2}B^rA^{r/2})\right] = \log A + \log B.$$

The next lemma follows from Lemmas 3.1 and 3.2.

Lemma 3.3. *Let $A, B \in \mathscr{P}_n$. Then*

$$\lambda(\log A + \log B) \prec_w \lambda(\log(A^{1/2}BA^{1/2})).$$

Theorem 3.4. *Let f be a log convex function on I. Then*

$$\lambda(f(\alpha A + (1-\alpha)B)) \prec_w \lambda(f(A)^{\alpha}f(B)^{1-\alpha})$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$.

Proof. The function $\log f(t)$ is a convex function on I. Therefore by Theorem 2.3 and Lemma 3.3, we get

$$\begin{aligned}\lambda(\log f(\alpha A + (1-\alpha)B)) &\prec_w \lambda(\alpha \log f(A) + (1-\alpha)\log f(B))\\ &= \lambda(\log f(A)^{\alpha} + \log f(B)^{1-\alpha})\\ &\prec_w \lambda(\log[f(A)^{\alpha/2}f(B)^{1-\alpha}f(A)^{\alpha/2}]).\end{aligned}$$

Since the function $t \to e^t$ is increasing and convex, we get

$$\begin{aligned}\lambda(f(\alpha A + (1-\alpha)B)) &\prec_w \lambda(f(A)^{\alpha/2}f(B)^{1-\alpha}f(A)^{\alpha/2})\\ &= \lambda(f(A)^{\alpha}f(B)^{1-\alpha}).\end{aligned}$$

This completes the proof. □

Since for any $X \in \mathscr{M}_n$, we have $|\lambda(X)| \prec_w \lambda(|X|)$ (see [4, p. 42]) by the Fan Dominance Theorem we get a proof of the following corollary.

Corollary 3.5. *Let f be a log convex function on I. Then*

$$|||f(\alpha A + (1-\alpha)B)||| \leqslant |||f(A)^{\alpha}f(B)^{1-\alpha}|||$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$.

Corollary 3.6. *Let $a > 1$ and $A, B \in \mathscr{H}_n$. Then*

$$\lambda(a^{A+B}) \prec_w \lambda(a^A a^B).$$

Proof. Let $p = \max\{\|A\|, \|B\|\}$. Then $-pI \leqslant A, B \leqslant pI$. The function $f(t) = a^t$ is log convex on $[-p, p]$. Therefore by Theorem 3.4, we get

$$\lambda(a^{\alpha A+(1-\alpha)B}) \prec_w \lambda(a^{\alpha A} a^{(1-\alpha)B})$$

for $0 \leqslant \alpha \leqslant 1$. Now by taking $\alpha = 1/2$ and then replacing A by $2A$ and B by $2B$ in the above inequality, we get the desired result. □

Remark 3.7. As a special case of Corollary 3.6 when $a = e$ we obtain the famous Golden–Thompson inequality:

$$\operatorname{tr}(e^{A+B}) \leqslant \operatorname{tr}(e^A e^B)$$

for $A, B \in \mathscr{H}_n$. Here for $X \in \mathscr{M}_n$, $\operatorname{tr}(X)$ denotes the trace of X. The following corollary may be considered as another generalization of the Golden–Thompson inequality.

Corollary 3.8 [8, p. 513–514]. *Let f be a multiplicatively convex function on* $(0, \infty)$. *Then*

$$\lambda(f(e^{\alpha A+(1-\alpha)B})) \prec_w \lambda(f(e^A)^\alpha f(e^B)^{1-\alpha})$$

for all $0 \leqslant \alpha \leqslant 1$ *and* $A, B \in \mathscr{H}_n$.

As another application of Theorem 3.4, we obtain a generalized harmonic–geometric mean (*Young*'*s*) inequality.

Corollary 3.9. *Let $A, B \in \mathscr{P}_n$ and $0 \leqslant \alpha \leqslant 1$. Then*

$$\lambda([\alpha A^{-1} + (1-\alpha)B^{-1}]^{-r}) \prec_w \lambda(A^{\alpha r} B^{(1-\alpha)r})$$

for all $r \geqslant 0$.

Proof. Let $p = \max\{\|A\|, \|A^{-1}\|, \|B\|, \|B^{-1}\|\}$. Then $-pI \leqslant A, A^{-1}, B, B^{-1} \leqslant pI$ and the function $t \to t^{-r}$ is log convex on $(0, p]$. Therefore by Theorem 3.4

$$\lambda([\alpha A + (1-\alpha)B]^{-r}) \prec_w \lambda(A^{-\alpha r} B^{-(1-\alpha)r}).$$

Now on replacing A by A^{-1} and B by B^{-1} in the above inequality, we get

$$\lambda([\alpha A^{-1} + (1-\alpha)B^{-1}]^{-r}) \prec_w \lambda(A^{\alpha r} B^{(1-\alpha)r}).$$

This completes the proof. □

Remark 3.10. For an increasing log convex function f

$$\lambda(f(\alpha A + (1-\alpha)B)) \leqslant \lambda(f(A)^\alpha f(B)^{1-\alpha})$$

does not hold. In fact, let $A, B \in \mathscr{H}_n$ and $f(t) = e^t$. It is known (see [4, p. 260]) that

$$\prod_{j=1}^{k} \lambda_j[\exp(\alpha A + (1-\alpha)B)] \leqslant \prod_{j=1}^{k} \lambda_j[\exp(\alpha A)\exp((1-\alpha)B)] \quad (k = 1, 2, \ldots, n).$$

But

$$\begin{aligned}\prod_{j=1}^{n} \lambda_j[\exp(\alpha A + (1-\alpha)B)] &= \det[\exp(\alpha A + (1-\alpha)B)] \\ &= \det[\exp(\alpha A)\exp((1-\alpha)B)] \\ &= \prod_{j=1}^{n} \lambda_j[\exp(\alpha A)\exp((1-\alpha)B)].\end{aligned}$$

Thus it follows that we can find $A, B \in \mathscr{H}_n$ and an i, $1 \leqslant i \leqslant n$ such that

$$\lambda_i(\exp(\alpha A + (1-\alpha)B)) \geqslant \lambda_i(\exp(\alpha A)\exp((1-\alpha)B)).$$

Remark 3.11. Note that Theorem 2.3 with $\prec_w$ replaced by $\prec^w$ (see [4, p. 30]) and Theorem 2.9 with inequalities in the reverse order hold when “convex function” is replaced by appropriate “concave function”. Then for a log concave function f on I one might conjecture that

$$\lambda(f(\alpha A + (1-\alpha)B)) \prec^w \lambda(f(A)^\alpha f(B)^{1-\alpha})$$

for all $A, B \in \mathscr{H}_n(I)$ and $0 \leqslant \alpha \leqslant 1$. However this fails. To see this one may take $f(t) = t^6$, $I = (0, \infty)$, $\alpha = 1/2$,

$$A = \begin{pmatrix} 4 & -5 \\ -5 & 7 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 9 & -1 \\ -1 & 1 \end{pmatrix}.$$

Lemma 3.12 [4, p. 267]. *Let $A, B \in \mathscr{P}_n$ and $0 \leqslant \alpha \leqslant 1$. Then*

$$|||A^\alpha B^{1-\alpha}||| \leqslant |||A|||^\alpha |||B|||^{1-\alpha}.$$

Theorem 3.13. *Let $A_j \in \mathscr{P}_n$, $j = 1, 2, \ldots, m$ and let f be a positive increasing function on $(0, \infty)$. Let*

$$g(t) = \left|\left|\left| f\left(\sum_{j=1}^{m} A_j^t\right) \right|\right|\right|.$$

If f is convex on $(0, \infty)$ then the function g is convex on $(-\infty, \infty)$. If f is log convex on $(0, \infty)$ then the function g is log convex on $(-\infty, \infty)$.

Proof. Let $s, t \in (-\infty, \infty)$. Note that the matrices

$$\begin{pmatrix} A_j^s & A_j^{(s+t)/2} \\ A_j^{(s+t)/2} & A_j^t \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A_j^t & A_j^{(s+t)/2} \\ A_j^{(s+t)/2} & A_j^s \end{pmatrix}$$

are positive semidefinite and hence so is the matrix

$$\begin{pmatrix} A_j^s + A_j^t & 2A_j^{(s+t)/2} \\ 2A_j^{(s+t)/2} & A_j^s + A_j^t \end{pmatrix}.$$

This implies that the map $t \to A_j^t$ is mid-point convex and hence by continuity it is convex. Since the sum of convex maps is convex it follows that the map $t \to \sum_{j=1}^m A_j^t$ is convex. Let $0 \leqslant \alpha \leqslant 1$. Therefore, we have

$$O \leqslant \sum_{j=1}^m A_j^{\alpha s+(1-\alpha)t} \leqslant \alpha \sum_{j=1}^m A_j^s + (1-\alpha) \sum_{j=1}^m A_j^t.$$

Since a unitarily invariant norm is monotone in the sence that $O \leqslant X \leqslant Y$ implies $|||X||| \leqslant |||Y|||$, the above inequality gives

$$\left|\left|\left| \sum_{j=1}^m A_j^{\alpha s+(1-\alpha)t} \right|\right|\right| \leqslant \left|\left|\left| \alpha \sum_{j=1}^m A_j^s + (1-\alpha) \sum_{j=1}^m A_j^t \right|\right|\right|. \tag{11}$$

Now if f is increasing and convex then inequality (11) together with the Fan Dominance Theorem and Corollary 2.6 imply

$$\left|\left|\left| f\left(\sum_{j=1}^m A_j^{\alpha s+(1-\alpha)t} \right) \right|\right|\right| \leqslant \left|\left|\left| \alpha f\left(\sum_{j=1}^m A_j^s \right) + (1-\alpha) f\left(\sum_{j=1}^m A_j^t \right) \right|\right|\right|.$$

Then use of the triangle inequality for norms gives the convexity of g. If f is increasing and log convex then inequality (11) together with the Fan Dominance Theorem, Corollary 3.5 and Lemma 3.12, imply

$$\begin{aligned} \left|\left|\left| f\left(\sum_{j=1}^m A_j^{\alpha s+(1-\alpha)t} \right) \right|\right|\right| &\leqslant \left|\left|\left| f\left(\alpha \sum_{j=1}^m A_j^s + (1-\alpha) \sum_{j=1}^m A_j^t \right) \right|\right|\right| \\ &\leqslant \left|\left|\left| \left(f\left(\sum_{j=1}^m A_j^s \right) \right)^{\alpha} \left(f\left(\sum_{j=1}^m A_j^t \right) \right)^{1-\alpha} \right|\right|\right| \\ &\leqslant \left|\left|\left| f\left(\sum_{j=1}^m A_j^s \right) \right|\right|\right|^{\alpha} \left|\left|\left| f\left(\sum_{j=1}^m A_j^t \right) \right|\right|\right|^{1-\alpha}. \end{aligned}$$

This proves the log convexity of g. $\square$

Remark 3.14. The special case of Theorem 3.13 when $f(t) = t^r$, $r \geqslant 1$ is Theorem 4 in [6].

Let $A = (a_{ij})$ and $B = (b_{ij})$ be in $\mathscr{M}_n$. Then their Hadamard product denoted by $A \circ B$ is the $n \times n$ matrix whose (i, j) entry is $(a_{ij} b_{ij})$. We have the following theorem whose proof is exactly similar to the proof of Theorem 3.13 and is therefore not included.

Theorem 3.15. *Let $A_j, B_j \in \mathscr{P}_n$, $j = 1, 2, \ldots, m$ and let f be a positive increasing function on $(0, \infty)$. Let*

$$g(t) = \left|\left|\left| f\left(\sum_{j=1}^{m}(A_j^t \circ B_j^t)\right)\right|\right|\right|.$$

If f is convex on $(0, \infty)$ then the function g is convex on $(-\infty, \infty)$. If f is log convex on $(0, \infty)$ then the function g is log convex on $(-\infty, \infty)$.

Next we prove a representation theorem.

Theorem 3.16. *Let $p, q > 1$ be such that $(1/p) + (1/q) = 1$ and $A \in \mathscr{P}_n$. Then*

$$\max_{X \in \Sigma} |||AX||| = |||A^p|||^{1/p}$$

where $\Sigma = \{X \in \mathscr{P}_n : |||X^q||| = 1\}$.

Proof. By Lemma 3.12, we have

$$|||A^{1/p} X^{1/q}||| \leqslant |||A|||^{1/p} |||X|||^{1/q}.$$

Now replace A by A^p and X by X^q to get

$$|||AX||| \leqslant |||A^p|||^{1/p}$$

using that $|||X^q||| = 1$ if $X \in \Sigma$. Equality occurs in the above inequality if we take $X^q = A^p/|||A^p|||$. This completes the proof. □

For the monotonicity and limit properties of the function $f(p) = |||A^p|||^{1/p}$ see [6, Corollary 9].

The following corollary is the well known Minkowski's inequality (see [4, p. 88]) for unitarily invariant norms.

Corollary 3.17. *Let $A, B \in \mathscr{P}_n$ and $p > 1$. Then*

$$|||(A + B)^p|||^{1/p} \leqslant |||A^p|||^{1/p} + |||B^p|||^{1/p}.$$

Proof. Let $q = p/(p - 1)$. Then $(1/p) + (1/q) = 1$. Therefore by Theorem 3.16, we have

$$
\begin{aligned}
|||(A+B)^p|||^{1/p} &= \max_{X\in\Sigma} |||(A+B)X||| \\
&\leqslant \max_{X\in\Sigma} |||AX||| + \max_{X\in\Sigma} |||BX||| \\
&= |||A^p|||^{1/p} + |||B^p|||^{1/p}.
\end{aligned}
$$

This is the desired inequality. □

Remark 3.18. Theorem 3.16 when norm is replaced by trace is the main result in [9].

Acknowledgement

This work was done when the first named author was at the Center for Linear Structures and Combinatorics of University of Lisbon as a postdoctoral fellow (ref. BPD-CI/01/02) and was supported by Fundação para a Ciência e Tecnologia. He thanks them for the support and hospitality. Both the authors thank Professor T. Ando who read the earlier version of this paper and gave valuable suggestions.

References

[1] T. Ando, Operator-Theoretic Methods for Matrix Inequalities, Unpublished Notes, 1998.
[2] T. Ando, X. Zhan, Norm inequalities related to operator monotone functions, Math. Ann. 315 (1999) 771–780.
[3] J.S. Aujla, Some norm inequalities for completely monotone functions, SIAM J. Matrix Anal. Appl. 22 (2000) 569–573.
[4] R. Bhatia, Matrix Analysis, Springer Verlag, New York, 1997.
[5] R. Bhatia, F. Kittaneh, Norm inequalities for positive operators, Lett. Math. Phys. 43 (1998) 225–231.
[6] F. Hiai, X. Zhan, Inequalities involving unitarily invariant norms and operator monotone functions, Linear Algebra Appl. 341 (2002) 151–169.
[7] O. Hirzallah, F. Kittaneh, Norm inequalities for weighted power means of operators, Linear Algebra Appl. 341 (2002) 181–193.
[8] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[9] J.R. Magnus, A representation theorem for $(\mathrm{tr}A^p)^{1/p}$, Linear Algebra Appl. 95 (1987) 127–134.
[10] C.P. Niculescu, Convexity according to the geometric mean, Math. Inequal. Appl. 3 (2000) 155–167.
[11] A.W. Roberts, D.E. Varberg, Convex Functions, Academic Press, New York, London, 1973.