Convex functions and matrix inequalities

Jaspal Singh Aujla
Department of Mathematics
National Institute of Technology
Jalandhar-144011, Punjab, INDIA
email: aujlajs@yahoo.com, aujlajs@nitj.ac.in

Abstract

Let f be a convex function defined on an interval I, $0 \leq \alpha \leq 1$ and let A, B be $n \times n$ complex Hermitian matrices with spectrum in I. It is proved 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 it is proved that the eigenvalues of $f(\alpha A + (1-\alpha)B)$ are weakly log-majorized by the eigenvalues of $f(A)^{\alpha} f(B)^{1-\alpha}$. If $I = [0, \infty)$, $f(0) \leq 0$ and f is monotone, then it is proved that there exits unitaries U, V such that $Uf(A)U^* + Vf(B)V^* \leq f(A+B)$. As applications we shall obtain generalizations of the famous Golden-Thomson trace inequality, a representation theorem and a harmonic-geometric mean inequality. Some related inequalities are also discussed.

AMS classification: 47A30; 47B15; 15A60

Keywords: convex function; weak majorization; unitarily invariant norm

1. INTRODUCTION

Throughout $\mathcal{M}_n$ shall denote the set of $n \times n$ complex matrices and $\mathcal{H}_n$ shall denote the set of all Hermitian matrices in $\mathcal{M}_n$. We shall denote by $\mathcal{S}_n$, the set of all positive semidefinite matrices in $\mathcal{M}_n$. The set of all positive definite matrices in $\mathcal{M}_n$ shall be denoted by $\mathcal{P}_n$. Let I be an interval in $\mathbf{R}$. We shall denote by $\mathcal{H}_n(I)$, the set of all Hermitian matrices in $\mathcal{M}_n$ whose spectrum is contained in I.

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

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

for all $0 \leq \alpha \leq 1$ 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

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

and is called log-concave if

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

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

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

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

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

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

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

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

and the Schatten p-norms defined as

$$||A||_p=\Big(\sum_{j=1}^{n}(s_j(A))^p\Big)^{1/p}$$

$1\leq p<\infty$, where $s_1(A)\geq s_2(A)\geq\cdots\geq s_n(A)$ 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 $|||diag(1,0,\ldots,0)|||=1$. The spectral norm (or operator norm) is given by $||A||=s_1(A)$. An $A\in\mathcal{M}_n$ is called a contraction if $||A||\leq 1$.

Throughout I shall denote an arbitrary interval (unless specified otherwise) in $\mathbf{R}$ and $|||\cdot|||$ shall denote an arbitrary unitarily invariant norm on $\mathcal{M}_n$. For (column) vectors $x,y\in\mathcal{C}^n$ their inner product is denoted by $\langle x,y\rangle=y^*x$. For an $A\in\mathcal{M}_n$, $\lambda_j(A),\ 1\leq j\leq n$ will always denote the eigenvalues of A arranged in the decreasing order whereas $s_j(A),\ 1\leq j\leq n$ will always denote the singular values of A arranged in the decreasing order. We shall use the notation $\lambda(A)$ to denote the row vector $(\lambda_1(A),\lambda_2(A),\ldots,\lambda_n(A))$.

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

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

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

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

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

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

(ii) (eigenvalue inequalities)

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

$$\overset{\text{def}}{\Longleftrightarrow} \quad B \le U^*AU \quad \exists \text{ unitary } U \in \mathcal{M}_n,$$

(iii) (weak log-majorization)

$$\lambda(B) \prec_{wlog} \lambda(A) \quad \overset{\text{def}}{\Longleftrightarrow} \quad \prod_{j=1}^{k} \lambda_j(B) \le \prod_{j=1}^{k} \lambda_j(A) \quad (k = 1, 2, \ldots, n).$$

(iv) (weak majorization)

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

Trivially we can see

$$B \le A \quad \Longrightarrow \quad \lambda(B) \le \lambda(A) \quad \Longrightarrow \quad \lambda(B) \prec_w \lambda(A).$$

$$\lambda(A), \lambda(B) > 0, \quad \lambda(B) \prec_{wlog} \lambda(A) \quad \Longrightarrow \lambda(B) \prec_w \lambda(A).$$

f increasing on I , $A, B \in \mathcal{H}_n(I)$, $\lambda(B) \le \lambda(A) \quad \Longrightarrow \quad \lambda(f(B)) \le \lambda(f(A))$,

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

In Section 2, we shall prove that for a convex function f on I

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

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

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

for all $A \in \mathcal{H}_n(I)$ and for all contractions $X \in \mathcal{M}_n$. If in addition the function f is also increasing (or decreasing), it is proved that

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

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

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

for all $A \in \mathcal{H}_n(I)$ and for all contractions $X \in \mathcal{M}_n$. In Section 3, for a log-convex function f on I , we shall prove that

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

for all $A, B \in \mathcal{H}_n(I)$ and $0 \leq \alpha \leq 1$. In this section we shall also prove a representation theorem. In Section 4, we shall prove matrix sub-additive inequalities for convex functions.

2. CONVEX FUNCTIONS

The following lemmas will be used to prove the main results in this section. The reader may refer to [6] for their proofs.

Lemma 2.1. [6, page 281] Let $A \in \mathcal{H}_n(I)$ and f be a convex function on I . Then for every unit vector $x \in \mathcal{C}^n$,

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

Lemma 2.2. [6, page 35] Let $A \in \mathcal{H}_n$. 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 the orthonormal vectors $u_1, u_2, \ldots, u_k$.

Lemma 2.3. [6, page 93] Let $A, B \in \mathcal{M}_n$. Then

$$||A||_{(k)} \ \leq \ ||B||_{(k)}$$

$k = 1, 2, \ldots, n$ if and only if

$$|||A||| \ \leq \ |||B|||,$$

for all unitarily invariant norms $|||\cdot|||$.

Theorem 2.4. 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 $A, B \in \mathcal{H}_n(I)$ and $0 \le \alpha \le 1$. If further $0 \in I$ and $f(0) \le 0$ then

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

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

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 $f(\lambda_1) \ge f(\lambda_2) \ge \cdots \ge f(\lambda_n)$. Let $k = 1, 2, \ldots, n$. Then

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

using convexity of f, Lemma 2.1 and Lemma 2.2 respectively. This proves

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

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 $f(\lambda_1) \ge f(\lambda_2) \ge \cdots \ge f(\lambda_n)$. Since $f(0) \le 0$, to prove the desired inequality we can assume that $||Xu_j|| \neq 0,\ j = 1, 2, \ldots, n$. Then

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

using convexity of f, the condition $f(0) \leq 0$, Lemma 2.1 and Lemma 2.2 respectively. Thus

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

This completes a proof. $\square$

The following corollary which supplements the results of Ando, Bhatia, Kittaneh and Zhan in [2,7]:

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

for $0 \leq r \leq 1$ and

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

for $r \geq 1,\ A, B \in \mathcal{S}_n$, was proved in [3]. The proof follows on taking $f(t) = t^r$, $r \leq 0$ and $I = (0, \infty)$ in Theorem 2.4.

Corollary 2.5. Let $A, B \in \mathcal{P}_n$. Then

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

for all $r \leq 0$.

The following corollary follows on using Theorem 2.4 and Lemma 2.3.

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

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

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

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

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

Remark 2.7. Corollary 2.6 may not be true if f is not nonnegative. To see this one may take $f(t) = -\log t$. This is convex on $(0, \infty)$. Let $0 \leq \alpha \leq 1$. It is easy to find $s, t \in (0, \infty)$ such that the inequality

$$|f(\alpha s + (1-\alpha)t)| \leq |\alpha f(s) + (1-\alpha)f(t)|$$

does not hold.

Remark 2.8. For $A, B \in \mathcal{H}_n$, the inequality (see [6, page 294])

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

is equivalent to

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

$m = 1, 2, \ldots$. The inequality,

$$|||\ |A+B|^r||| \leq 2^{r-1}|||\ |A|^r + |B|^r|||, \tag{3}$$

$r \geq 1$, which follows on choosing the nonnegative convex function $f(t) = |t|^r$, $r \geq 1$, on $(-\infty, \infty)$ in Corollary 2.6, as a special case provides an analogue of the above inequality for even powers. Another particular case of Corollary 2.6 when $f(t) = t^r$, $r \geq 1$ is Theorem 1 in [8].

If in addition, in Theorem 2.4 we also 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

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

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

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

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

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

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

It is known [6, page 58] that the eigenvalue $\lambda_j(H)$ admits the following max-min characterization:

$$\lambda_j(H) = \max_{\dim\mathcal{M}=j} \min\{\langle Hx,x\rangle \;;\; ||x||=1, x\in\mathcal{M}\} \tag{4}$$

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

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

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

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

By convexity of f and Lemma 2.1, we get

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

Now using formula (4), we have

$$\lambda_j\Big(f(\alpha A+(1-\alpha)B)\Big) \;\le\; \lambda_j\Big(\alpha f(A)+(1-\alpha)f(B)\Big).$$

This completes a proof of the first assertion.

Now suppose $0 \in I$ and $f(0) \leq 0$. Since $f(0) \leq 0$, we can assume that $||Xx|| \neq 0$ for all unit vectors $x \in \mathcal{C}^n$. We have as above

$$\lambda_j(f(X^*AX)) = \max_{\dim \mathcal{M}=j} \min\{f(\langle X^*AXx, x\rangle); ||x|| = 1, x \in \mathcal{M}\}.$$

Using convexity of f, the condition $f(0) \leq 0$ and Lemma 2.1, we get

$$\begin{aligned} f(\langle X^*AXx, x\rangle) &= f\Big(||Xx||^2\Big\langle A\frac{Xx}{||Xx||}, \frac{Xx}{||Xx||}\Big\rangle + (1-||Xx||^2)\cdot 0\Big) \\ &\leq ||Xx||^2 f\Big(\Big\langle A\frac{Xx}{||Xx||}, \frac{Xx}{||Xx||}\Big\rangle\Big) + (1-||Xx||^2)f(0) \\ &\leq ||Xx||^2\Big\langle f(A)\frac{Xx}{||Xx||}, \frac{Xx}{||Xx||}\Big\rangle \\ &= \langle X^*f(A)Xx, x\rangle. \end{aligned}$$

By (4) we get

$$\lambda_j(f(X^*AX)) \leq \lambda_j(X^*f(A)X).$$

This completes a proof of the second assertion. $\square$

Remark 2.10. Theorem 2.9 may not be true if f is not increasing (or decreasing). To see this one may take $f(t) = |t|$, $t \in (-\infty, \infty)$, $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$.

Remark 2.11. We would like to remark here that the inequality in Corollary 2.5 is sharp whereas inequalities (1) and (2) are not sharp. Taking the convex function $f(t) = t^r$, $r \geq 1$ in Theorem 2.9, we get

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

for all $r \geq 1$, which in turn gives a sharp upper bound for inequality (2). Now let $0 \leq r \leq 1$. Applying Theorem 2.9 to the decreasing convex function $g(t) = -t^r$, we get

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

This provides a sharp lower bound for inequality (1). Taking the decreasing convex function $f(t) = t^r$, $r \leq 0$ in Theorem 2.9, we get

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

which gives a stronger result than Corollary 2.5.

3. LOG-CONVEX FUNCTIONS

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

Lamma 3.1. [1, page 56] Let $A, B \in \mathcal{P}_n$ and $0 < r < 1$. Then

$$\lambda\Big(\tfrac{1}{r}\log(A^{r/2}B^rA^{r/2})\Big) \prec_w \lambda(\log(A^{1/2}BA^{1/2})).$$

The following lemma is known as Trotter's formula.

Lemma 3.2. [1, page 57] Let $A, B \in \mathcal{P}_n$. Then

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

The next lemma follows from Lemma 3.1 and Lemma 3.2.

Lemma 3.3. Let $A, B \in \mathcal{P}_n$. 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

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

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

Proof. Let f be log-convex on I. Then the function $\log f(t)$ is convex function on I . Therefore by Theorem 2.4 and Lemma 3.3, we get

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

This implies

$$\prod_{j=1}^{k} \lambda_j(f(\alpha A + (1-\alpha)B)) \leq \prod_{j=1}^{k} \lambda_j(f(A)^\alpha f(B)^{1-\alpha}), \quad 1 \leq k \leq n,$$

that is,

$$\lambda(f(\alpha A+(1-\alpha)B)) \prec_{wlog} \lambda(f(A)^{\alpha/2}f(B)^{1-\alpha}f(A)^{\alpha/2}) .$$

Since $\lambda_j(f(A)^{\alpha/2}f(B)^{1-\alpha}f(A)^{\alpha/2}) = \lambda_j(f(A)^{\alpha}f(B)^{1-\alpha})$, we get

$$\lambda(f(\alpha A+(1-\alpha)B)) \prec_{wlog} \lambda(f(A)^{\alpha}f(B)^{1-\alpha}).$$

This completes a proof. $\square$

Since for any $X \in \mathcal{M}_n$, we have

$$\sum_{j=1}^{k} |\lambda_j(X)| \leq \sum_{j=1}^{k} s_j(X),$$

$(k = 1, 2, \ldots, n)$, [6, page 42], by Lemma 2.3 we get a proof of the following corollary.

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

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

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

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

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

Proof. Let $p = \max\{||A||, ||B||\}$. Then $-pI \leq A, B \leq pI$. The function $f(t) = a^t$ is log-convex function 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 $0 \leq \alpha \leq 1$. Now by taking $\alpha = \frac{1}{2}$ and then replacing A by $2A$ and B by $2B$ in the above inequality, we get the desired result. $\square$

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

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

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

Corollary 3.8. (See [10, page 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 $0 \le \alpha \le 1$ and $A, B \in \mathcal{H}_n$.

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

Corollary 3.9. Let $A, B \in \mathcal{P}_n$ and $0 \le \alpha \le 1$. Then

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

for all $r \ge 0$.

Proof. Let $p = \max\{||A||, ||A^{-1}||, ||B||, ||B^{-1}||\}$. Then $-pI \le A, A^{-1}, B, B^{-1} \le pI$ 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

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

The second inequality follows, since

$$\sum_{j=1}^{k} \lambda_j(A^{\alpha r} B^{(1-\alpha)r}) \le \sum_{j=1}^{k} s_j(A^{\alpha r} B^{(1-\alpha)r}),$$

$k = 1, 2, \ldots, n$. This completes the proof. $\square$

Remark 3.10. For an increasing log-convex function f

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

can not be improved as

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

In fact, let $A, B \in \mathcal{H}_n$ and $f(t) = e^t$. By Theorem 3.4, we have

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

But

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

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

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

Therefore the ordering

$$\lambda(\exp(\alpha A + (1-\alpha) B)) \leq \lambda(\exp(\alpha A) \exp((1-\alpha) B))$$

does not hold.

Remark 3.11. Let f be a log-concave function on I. Then one might conjecture that

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

for all $A, B \in \mathcal{H}_n(I)$ and $0 \leq \alpha \leq 1$. However this fails. To see it one may take $f(t) = t^6$, $I = (0, \infty)$, $\alpha = \frac{1}{2}$, $A = \begin{pmatrix} 6 & -5 \\ -5 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 9 & -1 \\ -1 & 1 \end{pmatrix}$.

For a proof of the next lemma the reader is refered to [6, page 267].

Lemma 3.12. Let $A, B \in \mathcal{P}_n$ and $0 \leq \alpha \leq 1$. Then

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

Next we prove a representation theorem.

Theorem 3.13. Let $p, q > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and $A \in \mathcal{P}_n$. Then

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

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

Proof. By Lemma 3.12, we have

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

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

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

using that $|||X^q||| = 1$ if $X \in \Sigma$. The equality occurs in the above inequality if we take $X^q = \frac{A^p}{|||A^p|||}$. This completes a proof. $\square$

The following corollary is the well known Minkowski's inequality (see [6, page 88]) for unitarily invariant norms.

Corollary 3.14. Let $A, B \in \mathcal{P}_n$ and $p > 1$. Then

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

Proof. Let $q = \frac{p}{p-1}$. Then $\frac{1}{p} + \frac{1}{q} = 1$. Therefore by Theorem 3.13, we have

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

This is the desired inequality. $\square$

4. SUPPER-ADDITIVE INEQUALITIES

In this section we shall prove supper-additive inequalities for convex functions.

Theorem 4.1. Let f be a nonnegative convex function on $I = [0, \infty)$ with $f(0) = 0$ and $A, B \in \mathcal{S}_n$. Then there exists unitary matrices U and V such that

$$Uf(A)U^* + Vf(B)V^* \le f(A+B).$$

Proof. Let $A, B \in \mathcal{S}_n$. We can assume that $A+B$ is invertible. Then

$$A = A^{1/2}(A+B)^{-1/2}(A+B)(A+B)^{-1/2}A^{1/2} = X(A+B)X^*$$

and

$$B = B^{1/2}(A+B)^{-1/2}(A+B)(A+B)^{-1/2}B^{1/2} = Y(A+B)Y^*$$

where $X = A^{1/2}(A+B)^{-1/2}$ and $Y = B^{1/2}(A+B)^{-1/2}$ are contractions. Since for any $T \in \mathcal{M}_n$, T^*T and TT^* are unitarily congruent, we have by Theorem 2.9

$$\begin{aligned}\lambda(f(A)) &= \lambda(f(X(A+B)X^*) \\ &\leq \lambda(Xf(A+B)X^*) \\ &= \lambda((f(A+B))^{1/2}X^*X(f(A+B))^{1/2})).\end{aligned}$$

This implies there exists a unitary matrix U such that

$$Uf(A)U^* \leq (f(A+B))^{1/2}X^*X(f(A+B))^{1/2}. \tag{5}$$

Similarly there exists a unitary matrix V such that

$$Vf(B)V^* \leq (f(A+B))^{1/2}Y^*Y(f(A+B))^{1/2}. \tag{6}$$

Adding (5) and (6) we get

$$Uf(A)U^* + Vf(B)V^* \leq f(A+B)$$

since $X^*X + Y^*Y = I_n$. This completes a proof. $\square$

We would like to remark here that the inequality

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

is not true. To see this one may take $f(t) = t^2, A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. We have the following analogous result for the concave functions which have a similar proof.

Theorem 4.2. Let f be an nonnegative concave function on $I = [0, \infty)$ with $f(0) = 0$ and $A, B \in \mathcal{S}_n$. Then there exists unitary matrices U and V such that

$$f(A+B) \leq Uf(A)U^* + Vf(B)V^*.$$

References

[1] T. Ando, Operator-Theoretic Methods for Matrix Inequalities, Unpublished Notes, (1998).

[2] T. Ando and 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] J. S. Aujla and J. C. Bourin, Eigenvalue inequalities for convex and log-convex functions, Linear Alg. Appl., 626 (2007) 25-35.

[5] J. S. Aujla and F. C. Silva, Weak majorization inequalities and convex functions, Linear Alg. Appl., 369 (2003) 217-233.

[6] R. Bhatia, Matrix Analysis, Springer Verlag, New York, 1997.

[7] R. Bhatia and F. Kittaneh, Norm inequalities for positive operators, Letters Math. Phys., 63 (1998) 225-231.

[8] F. Hiai and X. Zhan, Inequalities involving unitarily invariant norms and operator monotone functions, Linear Alg. Appl., 361 (2002) 151-169.

[9] O. Hirzallah and F. Kittaneh, Norm inequalities for weighted power means of operators, Linear Alg. Appl., 361 (2002) 181-193.

[10] R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.

[11] R. A. Horn and R. Mathias, Cauchy-Schwarz inequalities associated with positive semidefinite matrices, Linear Alg. Appl., 163 (1990) 63-82.

[12] J. R. Magnus, A representation theorem for $(trA^p)^{1/p}$, Linear Alg. Appl., 95 (1987) 127-136.

[13] C. P. Niculescu, Convexity according to the geometric mean, Math. Ineq. Appl., 3 (2000) 155-167.