J. Math. Anal. Appl. 361 (2010) 262–269

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Improved Young and Heinz inequalities for matrices

Fuad Kittaneh [a,*], Yousef Manasrah [b]

[a] Department of Mathematics, University of Jordan, Amman, Jordan
[b] Department of Mathematics, Al-Zaytoonah University, Amman, Jordan

ARTICLE INFO

Article history:
Received 16 May 2009
Available online 2 September 2009
Submitted by J.A. Ball

Keywords:
Unitarily invariant norm
Young inequality
Heinz inequality
Positive semidefinite matrix

ABSTRACT

We give refinements of the classical Young inequality for positive real numbers and we use these refinements to establish improved Young and Heinz inequalities for matrices.

1. Introduction

Let $M_n(\mathbb{C})$ be the space of $n \times n$ complex matrices and let $|||\cdot|||$ denote any unitarily invariant (or symmetric) norm on $M_n(\mathbb{C})$. So, $|||UAV||| = |||A|||$ for all $A \in M_n(\mathbb{C})$ and for all unitary matrices $U, V \in M_n(\mathbb{C})$. For $A \in M_n(\mathbb{C})$, the Hilbert–Schmidt norm, the trace norm, and the spectral norm of A are defined by $\|A\|_2 = (\sum_{j=1}^{n} s_j^2(A))^{1/2}$, $\|A\|_1 = \sum_{j=1}^{n} s_j(A)$, and $\|A\| = s_1(A)$, respectively, where $s_1(A) \geqslant \cdots \geqslant s_n(A)$ are the singular values of A, that is, the eigenvalues of the positive semidefinite matrix $|A| = (A^*A)^{1/2}$.

Note that $\|A\|_2 = (\operatorname{tr}|A|^2)^{1/2}$ and $\|A\|_1 = \operatorname{tr}|A|$, where tr is the usual trace functional. It is evident that these norms are unitarily invariant, and it is known that each unitarily invariant norm is a symmetric guage function of singular values [4, p. 91].

The classical Young inequality for two scalars is the ν-weighted arithmetic–geometric mean inequality, which is a fundamental relation between two nonnegative real numbers. This inequality says that if $a, b \geqslant 0$ and $0 \leqslant \nu \leqslant 1$, then

$$a^{\nu} b^{1-\nu} \leqslant \nu a + (1-\nu) b \tag{1.1}$$

with equality if and only if $a = b$. If $\nu = \frac{1}{2}$, we obtain the arithmetic–geometric mean inequality

$$\sqrt{ab} \leqslant \frac{a+b}{2}. \tag{1.2}$$

If $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, then the inequality (1.1) can be written as

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}. \tag{1.3}$$

* Corresponding author.
E-mail addresses: fkitt@ju.edu.jo (F. Kittaneh), Y.manasrah@Alzaytoonah.edu.jo (Y. Manasrah).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2009.08.059

A matrix version proved in [2] says that if $A, B \in M_n(\mathbb{C})$ are positive semidefinite, then

$$s_j(AB) \leqslant s_j\left(\frac{A^p}{p} + \frac{B^q}{q}\right) \tag{1.4}$$

for $j = 1, \ldots, n$.

The special case where $p = q = 2$ has been obtained earlier in [6]. Young's inequality in operator algebras has been considered in [8] and references therein.

It follows immediately from (1.4) that if $A, B \in M_n(\mathbb{C})$ are positive semidefinite and $0 \leqslant \nu \leqslant 1$, then a trace version of Young's inequality holds:

$$\operatorname{tr}\left|A^\nu B^{1-\nu}\right| \leqslant \operatorname{tr}\big(\nu A + (1-\nu)B\big). \tag{1.5}$$

A determinant version of Young's inequality is also known [10, p. 467]:

$$\det\big(A^\nu B^{1-\nu}\big) \leqslant \det\big(\nu A + (1-\nu)B\big). \tag{1.6}$$

Bhatia and Parthasarathy [7] and Kosaki [12] proved that if $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite and if $0 \leqslant \nu \leqslant 1$, then

$$\left\|A^\nu X B^{1-\nu}\right\|_2 \leqslant \left\|\nu AX + (1-\nu)XB\right\|_2. \tag{1.7}$$

It should be mentioned here that for $\nu \neq \frac{1}{2}$, the inequality (1.7) may not hold for other unitarily invariant norms.

The Heinz means are defined as follows:

$$H_\nu(a,b) = \frac{a^\nu b^{1-\nu} + a^{1-\nu} b^\nu}{2}$$

for $0 \leqslant \nu \leqslant 1$ and $a, b \geqslant 0$. It follows immediately from the inequalities (1.1) and (1.3) that the Heinz means interpolate between the geometric mean and the arithmetic mean:

$$\sqrt{ab} \leqslant \frac{a^\nu b^{1-\nu} + a^{1-\nu} b^\nu}{2} \leqslant \frac{a+b}{2}. \tag{1.8}$$

Bhatia and Davis proved in [5] that if $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite and if $0 \leqslant \nu \leqslant 1$, then

$$\left|\!\left|\!\left|A^{1/2} X B^{1/2}\right|\!\right|\!\right| \leqslant \left|\!\left|\!\left|\frac{A^\nu X B^{1-\nu} + A^{1-\nu} X B^\nu}{2}\right|\!\right|\!\right| \leqslant \left|\!\left|\!\left|\frac{AX + XB}{2}\right|\!\right|\!\right|. \tag{1.9}$$

The second inequality in (1.9) is known as Heinz inequality [4, p. 265].

Recently, Audenaert [3] gave a singular value inequality for Heinz means of matrices as follows:

If $A, B \in M_n(\mathbb{C})$ are positive semidefinite and $0 \leqslant \nu \leqslant 1$, then

$$s_j\big(A^\nu B^{1-\nu} + A^{1-\nu} B^\nu\big) \leqslant s_j(A+B) \tag{1.10}$$

for $j = 1, \ldots, n$.

The inequality (1.10) entails the second inequality in (1.9) for the case $X = I$. Recall that unitarily invariant norms are increasing functions of singular values.

In this paper, we refine Young's inequality (1.1) and the second inequality in (1.8). These refinements enable us to establish improved Young and Heinz inequalities for matrices. In particular, we will improve the trace and the determinant inequalities (1.5) and (1.6), and we will improve the Heinz inequality (the second inequality in (1.9)) for the Hilbert–Schmidt norm. As an application of our improved Heinz inequality, we will investigate the equality condition in the inequality (1.10). Though we confine our discussion to matrices, by slight modification the norm inequalities we obtain here can be extended to operators on an infinite-dimensional complex separable Hilbert space. In this setting, when we consider $|\!|\!|T|\!|\!|$, we are implicitly assuming that the operator T belongs to the norm ideal associated with $|\!|\!|\cdot|\!|\!|$.

2. Refinements of the scalar Young's inequality

We begin this section with the following refinement of the scalar Young's inequality (1.1).

Theorem 2.1. *If $a, b \geqslant 0$ and $0 \leqslant \nu \leqslant 1$, then*

$$a^\nu b^{1-\nu} + r_0(\sqrt{a} - \sqrt{b})^2 \leqslant \nu a + (1-\nu)b, \tag{2.1}$$

where $r_0 = \min\{\nu, 1-\nu\}$.

Proof. If $\nu = \frac{1}{2}$, the inequality (2.1) becomes an equality. Assume that $\nu < \frac{1}{2}$. Then, by the inequality (1.1), we have

$$\begin{aligned}\nu a + (1-\nu)b - \nu(\sqrt{a} - \sqrt{b})^2 &= 2\nu\sqrt{ab} + (1-2\nu)b \\ &\geqslant (ab)^\nu b^{1-2\nu} \\ &= a^\nu b^{1-\nu},\end{aligned}$$

and so

$$\nu a + (1-\nu)b \geqslant \nu(\sqrt{a} - \sqrt{b})^2 + a^\nu b^{1-\nu}.$$

If $1 - \nu < \frac{1}{2}$, then

$$\begin{aligned}\nu a + (1-\nu)b - (1-\nu)(\sqrt{a} - \sqrt{b})^2 &= (2\nu - 1)a + 2(1-\nu)\sqrt{ab} \\ &\geqslant a^{2\nu-1}(ab)^{1-\nu} \\ &= a^\nu b^{1-\nu},\end{aligned}$$

and so

$$\nu a + (1-\nu)b \geqslant (1-\nu)(\sqrt{a} - \sqrt{b})^2 + a^\nu b^{1-\nu}.$$

Hence,

$$a^\nu b^{1-\nu} + r_0(\sqrt{a} - \sqrt{b})^2 \leqslant \nu a + (1-\nu)b.$$

This completes the proof. □

As a direct consequence of Theorem 2.1, we have

$$a^\nu b^{1-\nu} + a^{1-\nu} b^\nu + 2r_0(\sqrt{a} - \sqrt{b})^2 \leqslant a + b, \tag{2.2}$$

and so

$$H_\nu(a,b) + r_0(\sqrt{a} - \sqrt{b})^2 \leqslant \frac{a+b}{2}.$$

Corollary 2.2. *If* $a, b \geqslant 0$ *and* $0 \leqslant \nu \leqslant 1$, *then*

$$\left(a^\nu b^{1-\nu} + a^{1-\nu} b^\nu\right)^2 + 2r_0(a-b)^2 \leqslant (a+b)^2, \tag{2.3}$$

where $r_0 = \min\{\nu, 1-\nu\}$.

Proof. By Theorem 2.1, we have

$$\begin{aligned}(a+b)^2 - \left(a^\nu b^{1-\nu} + a^{1-\nu} b^\nu\right)^2 &= a^2 + b^2 - a^{2\nu} b^{2(1-\nu)} - a^{2(1-\nu)} b^{2\nu} \\ &= \nu a^2 + (1-\nu)b^2 - a^{2\nu} b^{2(1-\nu)} + (1-\nu)a^2 + \nu b^2 - a^{2(1-\nu)} b^{2\nu} \\ &\geqslant r_0(a-b)^2 + r_0(a-b)^2 \\ &= 2r_0(a-b)^2. \qquad \square\end{aligned}$$

Hirzallah and Kittaneh [9] obtained a related refinement of the scalar Young's inequality as follows:

$$\left(a^\nu b^{1-\nu}\right)^2 + r_0^2(a-b)^2 \leqslant \left(\nu a + (1-\nu)b\right)^2. \tag{2.4}$$

Remark. If we replace a by a^2 and b by b^2, the inequality (2.1) can be written in the form

$$\left(a^\nu b^{1-\nu}\right)^2 + r_0(a-b)^2 \leqslant \nu a^2 + (1-\nu)b^2. \tag{2.5}$$

When comparing the inequality (2.5) with the inequality (2.4), it is easy to observe that the left-hand side and the right-hand side in the inequality (2.5) are greater than or equal to the corresponding sides in the inequality (2.4), respectively. It should be noticed here that neither (2.4) nor (2.5) is uniformly better than the other.

3. Improved Young and Heinz inequalities for matrices

Based on the refined Young inequality (2.4), it has been shown in [9] that if $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite and if $0 \leqslant \nu \leqslant 1$, then

$$\left\| A^\nu X B^{1-\nu} \right\|_2^2 + r_0^2 \|AX - XB\|_2^2 \leqslant \left\| \nu AX + (1-\nu) XB \right\|_2^2,$$

which is an improvement of the matrix Young inequality (1.7) for the Hilbert–Schmidt norm. In the next theorem, we give refinements of both the trace and the determinant versions of Young's inequality based on the refined Young inequality (2.1). To do this, we need the following lemma (see, e.g., [4, p. 94]).

Lemma 3.1. *Let $A, B \in M_n(\mathbb{C})$. Then*

$$\sum_{j=1}^n s_j(AB) \leqslant \sum_{j=1}^n s_j(A) s_j(B).$$

Theorem 3.2. *Let $A, B \in M_n(\mathbb{C})$ be positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$\operatorname{tr} \left| A^\nu B^{1-\nu} \right| + r_0 (\sqrt{\operatorname{tr} A} - \sqrt{\operatorname{tr} B})^2 \leqslant \operatorname{tr}\big(\nu A + (1-\nu) B\big).$$

If A and B are positive definite, then

$$\det\big(A^\nu B^{1-\nu}\big) + r_0^n \det(A + B - 2A \# B) \leqslant \det\big(\nu A + (1-\nu) B\big),$$

where $r_0 = \min\{\nu, 1-\nu\}$ and $A \# B = B^{1/2}(B^{-1/2} A B^{-1/2})^{1/2} B^{1/2}$ is the geometric mean of A and B.

Proof. By the inequality (2.1), we have

$$\nu s_j(A) + (1-\nu) s_j(B) \geqslant s_j^\nu(A) s_j^{1-\nu}(B) + r_0 \big(s_j^{1/2}(A) - s_j^{1/2}(B)\big)^2$$

for $j = 1, \dots, n$. Thus, by Lemma 3.1 and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\operatorname{tr}\big(\nu A + (1-\nu) B\big) &= \nu \operatorname{tr} A + (1-\nu) \operatorname{tr} B \\
&= \sum_{j=1}^n \big(\nu s_j(A) + (1-\nu) s_j(B)\big) \\
&\geqslant \sum_{j=1}^n s_j\big(A^\nu\big) s_j\big(B^{1-\nu}\big) + r_0 \left(\sum_{j=1}^n s_j(A) + \sum_{j=1}^n s_j(B) - 2 \sum_{j=1}^n s_j^{1/2}(A) s_j^{1/2}(B) \right) \\
&\geqslant \sum_{j=1}^n s_j\big(A^\nu B^{1-\nu}\big) + r_0 \left(\operatorname{tr} A + \operatorname{tr} B - 2 \left(\sum_{j=1}^n s_j(A) \right)^{1/2} \left(\sum_{j=1}^n s_j(B) \right)^{1/2} \right) \\
&= \operatorname{tr} \left| A^\nu B^{1-\nu} \right| + r_0 (\sqrt{\operatorname{tr} A} - \sqrt{\operatorname{tr} B})^2.
\end{aligned}$$

This completes the proof of the trace inequality.

To prove the determinant inequality, note that by the inequality (2.1), we have

$$\nu s_j\big(B^{-1/2} A B^{-1/2}\big) + (1-\nu) \geqslant s_j^\nu\big(B^{-1/2} A B^{-1/2}\big) + r_0 \big(s_j^{1/2}\big(B^{-1/2} A B^{-1/2}\big) - 1\big)^2$$

for $j = 1, \dots, n$. Thus,

$$\begin{aligned}
\det\big(\nu B^{-1/2} A B^{-1/2} + (1-\nu) I\big) &= \prod_{j=1}^n \big(\nu s_j\big(B^{-1/2} A B^{-1/2}\big) + 1 - \nu\big) \\
&\geqslant \prod_{j=1}^n \big[s_j^\nu\big(B^{-1/2} A B^{-1/2}\big) + r_0 \big(s_j^{1/2}\big(B^{-1/2} A B^{-1/2}\big) - 1\big)^2\big] \\
&\geqslant \prod_{j=1}^n s_j^\nu\big(B^{-1/2} A B^{-1/2}\big) + r_0^n \prod_{j=1}^n \big(s_j^{1/2}\big(B^{-1/2} A B^{-1/2}\big) - 1\big)^2 \\
&= \det\big(B^{-1/2} A B^{-1/2}\big)^\nu + r_0^n \det\big(\big(B^{-1/2} A B^{-1/2}\big)^{1/2} - I\big)^2.
\end{aligned}$$

Consequently,

$$\det\left(A^{\nu}B^{1-\nu}\right)+r_0^n\det(A+B-2A\,\#\,B)\leqslant\det\left(\nu A+(1-\nu)B\right).\qquad\square$$

Remark. Ando's singular value inequality (1.4) entails the norm inequality

$$\left|\left|\left|A^{\nu}B^{1-\nu}\right|\right|\right|\leqslant\left|\left|\left|\nu A+(1-\nu)B\right|\right|\right|.\tag{3.1}$$

So, our Theorem 3.2 improves this inequality for the trace norm:

$$\left\|A^{\nu}B^{1-\nu}\right\|_1+r_0\left(\sqrt{\|A\|_1}-\sqrt{\|B\|_1}\right)^2\leqslant\left\|\nu A+(1-\nu)B\right\|_1.\tag{3.2}$$

This is false for the spectral norm. To see this, let $A=\left[\begin{smallmatrix}2&0\\0&3\end{smallmatrix}\right]$, $B=\left[\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right]$, and $\nu=\frac{1}{2}$.

In the next result, we give an improved arithmetic–geometric mean inequality for the Hilbert–Schmidt norm.

Theorem 3.3. *Let $A, B, X\in M_n(\mathbb{C})$ such that A and B are positive semidefinite. Then*

$$2\left\|A^{1/2}XB^{1/2}\right\|_2+\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\leqslant\|AX+XB\|_2.$$

Proof. For $Y, Z\in M_n(\mathbb{C})$, the formula $\langle Y, Z\rangle=\operatorname{tr}YZ^*$ defines an inner product on $M_n(\mathbb{C})$. So, by the fact that $\operatorname{tr}YZ=\operatorname{tr}ZY$, we have $\langle AX, XB\rangle=\langle XB, AX\rangle$. Using the Cauchy–Schwarz inequality, we also have

$$\begin{aligned}\langle AX, XB\rangle&=\operatorname{tr}AXBX^*\\&=\operatorname{tr}A^{1/2}XB^{1/2}B^{1/2}X^*A^{1/2}\\&=\left\|A^{1/2}XB^{1/2}\right\|_2^2\\&\leqslant\|AX\|_2\|XB\|_2.\end{aligned}$$

Now,

$$\begin{aligned}&\|AX+XB\|_2^2-\left(\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2+2\left\|A^{1/2}XB^{1/2}\right\|_2\right)^2\\&\quad=\|AX\|_2^2+\|XB\|_2^2+2\langle AX, XB\rangle-4\left\|A^{1/2}XB^{1/2}\right\|_2^2\\&\qquad-\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^4-4\left\|A^{1/2}XB^{1/2}\right\|_2\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\\&\quad=\|AX\|_2^2+\|XB\|_2^2+2\langle AX, XB\rangle-4\langle AX, XB\rangle-\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^4\\&\qquad-4\left\|A^{1/2}XB^{1/2}\right\|_2\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\\&\quad=\|AX\|_2^2+\|XB\|_2^2-2\langle AX, XB\rangle-\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^4-4\left\|A^{1/2}XB^{1/2}\right\|_2\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\\&\quad\geqslant\|AX\|_2^2+\|XB\|_2^2-2\|AX\|_2\|XB\|_2-\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^4-4\left\|A^{1/2}XB^{1/2}\right\|_2\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\\&\quad=\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\left[\left(\sqrt{\|AX\|_2}+\sqrt{\|XB\|_2}\right)^2-\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2-4\left\|A^{1/2}XB^{1/2}\right\|_2\right]\\&\quad=4\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\left(\sqrt{\|AX\|_2\|XB\|_2}-\left\|A^{1/2}XB^{1/2}\right\|_2\right)\\&\quad\geqslant 0.\end{aligned}$$

Hence,

$$\|AX+XB\|_2\geqslant\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2+2\left\|A^{1/2}XB^{1/2}\right\|_2.$$

This completes the proof. $\square$

Our improved Heinz inequality for the Hilbert–Schmidt norm can be stated as follows:

Theorem 3.4. *Let $A, B, X\in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0\leqslant\nu\leqslant 1$, then*

$$\left\|A^{\nu}XB^{1-\nu}+A^{1-\nu}XB^{\nu}\right\|_2+2r_0\left(\sqrt{\|AX\|_2}-\sqrt{\|XB\|_2}\right)^2\leqslant\|AX+XB\|_2,$$

where $r_0=\min\{\nu, 1-\nu\}$.

Proof. First assume that A and B are positive definite matrices. The general case will follow from the special one by a continuity argument. Since $A, B > 0$, it follows by the spectral theorem that there are unitary matrices $U, V \in M_n(\mathbb{C})$ such that $A = UDU^*$ and $B = VEV^*$, where $D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $E = \operatorname{diag}(\mu_1, \mu_2, \dots, \mu_n)$, and $\lambda_i, \mu_i > 0$, $i = 1, 2, \dots, n$. If $Y = U^*XV = [y_{ij}]$, then

$$\begin{aligned} A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu &= U\big(D^\nu YE^{1-\nu} + D^{1-\nu}YE^\nu\big)U^* \\ &= U\big[\big(\lambda_i^\nu \mu_j^{1-\nu} + \lambda_i^{1-\nu}\mu_j^\nu\big)y_{ij}\big]U^*. \end{aligned}$$

So,

$$\big\|A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu\big\|_2 = \left(\sum_{i,j=1}^n \big(\lambda_i^\nu \mu_j^{1-\nu} + \lambda_i^{1-\nu}\mu_j^\nu\big)^2 |y_{ij}|^2\right)^{1/2}.$$

Let $\psi(\nu) = \|A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu\|_2$, $0 \leqslant \nu \leqslant 1$. Then ψ is a continuous convex function on $[0, 1]$ (see [4, p. 265]). Moreover, ψ is twice differentiable on $(0, 1)$. If $f(\nu) = \|AX + XB\|_2 - \|A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu\|_2$, then $f(\nu) = f(1-\nu)$, $f(0) = f(1) = 0$, and one can infer from [4, p. 265] that f is concave on $[0, 1]$, $f(\frac{1}{2}) \geqslant 0$, and $f'(\frac{1}{2}) = 0$.

Define $g(\nu) = \frac{1}{\min\{\nu, 1-\nu\}} f(\nu)$ for $0 < \nu < 1$. Then g can be rewritten as follows:

$$g(\nu) = \begin{cases} \frac{f(\nu)}{\nu}, & 0 < \nu \leqslant \frac{1}{2}, \\ \frac{f(1-\nu)}{1-\nu}, & \frac{1}{2} \leqslant \nu < 1. \end{cases}$$

So,

$$g'(\nu) = \begin{cases} \frac{\nu f'(\nu) - f(\nu)}{\nu^2}, & 0 < \nu < \frac{1}{2}, \\ \frac{-(1-\nu)f'(1-\nu) + f(1-\nu)}{(1-\nu)^2}, & \frac{1}{2} < \nu < 1. \end{cases}$$

Consider the function $h(\nu) = \nu f'(\nu) - f(\nu)$, $0 \leqslant \nu \leqslant 1$. Then $h(0) = 0$ and $h'(\nu) = \nu f''(\nu) \leqslant 0$, which means that $h(\nu) \leqslant 0$. Thus, $\nu f'(\nu) \leqslant f(\nu)$ for $0 \leqslant \nu \leqslant \frac{1}{2}$. By a similar argument, we can show that $f(1-\nu) \geqslant (1-\nu)f'(\nu)$ for $\frac{1}{2} \leqslant \nu \leqslant 1$. Therefore, g is decreasing on $(0, \frac{1}{2})$ and increasing on $(\frac{1}{2}, 1)$. Since $g'_-(\frac{1}{2}) = -4f(\frac{1}{2})$ and $g'_+(\frac{1}{2}) = 4f(\frac{1}{2})$, it follows from the continuity of g and the symmetry of g about $\nu = \frac{1}{2}$ that g attains its minimum at $\nu = \frac{1}{2}$. Hence, $g(\nu) \geqslant g(\frac{1}{2})$, and so by Theorem 3.3, we have

$$\begin{aligned} g(\nu) &\geqslant 2\big(\|AX + XB\|_2 - 2\big\|A^{1/2}XB^{1/2}\big\|_2\big) \\ &\geqslant 2\big(\sqrt{\|AX\|_2} - \sqrt{\|XB\|_2}\big)^2. \end{aligned}$$

Thus,

$$\|AX + XB\|_2 - \big\|A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu\big\|_2 \geqslant 2\min\{\nu, 1-\nu\}\big(\sqrt{\|AX\|_2} - \sqrt{\|XB\|_2}\big)^2.$$

To prove the general positive semidefinite case, assume that $A_\epsilon = A + \epsilon I$ and $B_\epsilon = B + \epsilon I$, where ϵ is an arbitrary positive real number. Then A_ϵ and B_ϵ are positive definite, and so by the above special case, we get

$$\big\|A_\epsilon^\nu XB_\epsilon^{1-\nu} + A_\epsilon^{1-\nu}XB_\epsilon^\nu\big\|_2 + 2r_0\big(\sqrt{\|A_\epsilon X\|_2} - \sqrt{\|XB_\epsilon\|_2}\big)^2 \leqslant \|A_\epsilon X + XB_\epsilon\|_2.$$

The desired inequality now follows by letting $\epsilon \to 0$. □

Another improvement of the Heinz inequality for the Hilbert–Schmidt norm can be stated as follows:

Theorem 3.5. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$\big\|A^\nu XB^{1-\nu} + A^{1-\nu}XB^\nu\big\|_2^2 + 2r_0\|AX - XB\|_2^2 \leqslant \|AX + XB\|_2^2,$$

where $r_0 = \min\{\nu, 1-\nu\}$.

Proof. The proof of this result can be accomplished by utilizing the inequality (2.3), the spectral theorem for positive semidefinite matrices, and the unitary invariance of the Hilbert–Schmidt norm. □

It is known [1] that for $\nu \neq \frac{1}{2}$ and $|||\cdot||| \neq \|\cdot\|_2$, the X-version of the inequality (3.1) is not true. That is, for $\nu \neq \frac{1}{2}$ and $|||\cdot||| \neq \|\cdot\|_2$, the inequality $|||A^\nu XB^{1-\nu}||| \leqslant |||\nu AX + (1-\nu)XB|||$ does not hold. However, it has been shown in [12] that

$$|||A^{\nu}XB^{1-\nu}||| \leqslant \nu|||AX||| + (1-\nu)|||XB|||.$$

In view of the inequality (2.1), we can improve this inequality and generalize the inequality (3.2). To achieve this, we need the following lemma [11].

Lemma 3.6. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$|||A^{\nu}XB^{1-\nu}||| \leqslant |||AX|||^{\nu}|||XB|||^{1-\nu}.$$

Theorem 3.7. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$|||A^{\nu}XB^{1-\nu}||| + r_0\big(\sqrt{|||AX|||} - \sqrt{|||XB|||}\big)^2 \leqslant \nu|||AX||| + (1-\nu)|||XB|||,$$

where $r_0 = \min\{\nu, 1-\nu\}$.

Proof. By Lemma 3.6 and the inequality (2.1), we have

$$\begin{aligned}|||A^{\nu}XB^{1-\nu}||| + r_0\big(\sqrt{|||AX|||} - \sqrt{|||XB|||}\big)^2 &\leqslant |||AX|||^{\nu}|||XB|||^{1-\nu} + r_0\big(\sqrt{|||AX|||} - \sqrt{|||XB|||}\big)^2 \\ &\leqslant \nu|||AX||| + (1-\nu)|||XB|||. \qquad \square\end{aligned}$$

Specializing Theorem 3.7 to the trace norm and letting $X = I$, we retain the inequality (3.2).

In the same vein, the triangle inequality, Lemma 3.6, and the inequality (2.3) yield the following result related to Theorem 3.4.

Theorem 3.8. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$|||A^{\nu}XB^{1-\nu} + A^{1-\nu}XB^{\nu}||| + 2r_0\big(\sqrt{|||AX|||} - \sqrt{|||XB|||}\big)^2 \leqslant |||AX||| + |||XB|||,$$

where $r_0 = \min\{\nu, 1-\nu\}$.

Specializing Theorem 3.8 to the trace norm and letting $X = I$, we obtain the following improved Heinz inequality for the trace norm.

Corollary 3.9. *Let $A, B \in M_n(\mathbb{C})$ be positive semidefinite. If $0 \leqslant \nu \leqslant 1$, then*

$$\|A^{\nu}B^{1-\nu} + A^{1-\nu}B^{\nu}\|_1 + 2r_0\big(\sqrt{\|A\|_1} - \sqrt{\|B\|_1}\big)^2 \leqslant \|A+B\|_1.$$

As applications of Theorem 3.5, we give necessary and sufficient conditions for equality to hold in Heinz inequality (the second inequality in (1.9)) for the Hilbert–Schmidt norm, and in the singular value inequality (1.10).

Corollary 3.10. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $0 < \nu < 1$, then*

$$\|A^{\nu}XB^{1-\nu} + A^{1-\nu}XB^{\nu}\|_2 = \|AX + XB\|_2$$

if and only if $AX = XB$.

Proof. If $AX = XB$, then by the spectral theorem, we have $A^{\nu}X = XB^{\nu}$ and $A^{1-\nu}X = XB^{1-\nu}$. Thus, $A^{\nu}XB^{1-\nu} + A^{1-\nu}XB^{\nu} = A^{\nu}A^{1-\nu}X + XB^{1-\nu}B^{\nu} = AX + XB$.

Conversely, assume that $\|A^{\nu}XB^{1-\nu} + A^{1-\nu}XB^{\nu}\|_2 = \|AX + XB\|_2$. Then it follows from Theorem 3.5 that $\|AX - XB\|_2 = 0$, and so $AX = XB$. $\square$

Corollary 3.11. *Let $A, B \in M_n(\mathbb{C})$ be positive semidefinite. If $0 < \nu < 1$, then*

$$s_j\big(A^{\nu}B^{1-\nu} + A^{1-\nu}B^{\nu}\big) = s_j(A+B)$$

for $j = 1, \ldots, n$ if and only if $A = B$.

Proof. If $A = B$, then clearly $A^{\nu}B^{1-\nu} + A^{1-\nu}B^{\nu} = A + B$.

Conversely, assume that $s_j(A^{\nu}B^{1-\nu} + A^{1-\nu}B^{\nu}) = s_j(A+B)$ for $j = 1, \ldots, n$. Then $\|A^{\nu}B^{1-\nu} + A^{1-\nu}B^{\nu}\|_2 = \|A+B\|_2$, and so it follows from the case $X = I$ of Corollary 3.10 that $A = B$. $\square$

References

[1] T. Ando, Majorization and inequalities in matrix theory, Linear Algebra Appl. 199 (1994) 17–67.
[2] T. Ando, Matrix Young inequality, Oper. Theory Adv. Appl. 75 (1995) 33–38.
[3] K.M.R. Audenaert, A singular value inequality for Heinz means, Linear Algebra Appl. 422 (2007) 279–283.
[4] R. Bhatia, Matrix Analysis, Springer-Verlag, 1997.
[5] R. Bhatia, C. Davis, More matrix forms of the arithmetic–geometric mean inequality, SIAM J. Matrix Anal. Appl. 14 (1993) 132–136.
[6] R. Bhatia, F. Kittaneh, On the singular values of a product of operators, SIAM J. Matrix Anal. Appl. 11 (1990) 272–277.
[7] R. Bhatia, K.R. Parthasarathy, Positive definite functions and operator inequalities, Bull. London Math. Soc. 32 (2000) 214–228.
[8] D.R. Farenick, M.S. Manjegani, Young's inequality in operator algebras, J. Ramanujan Math. Soc. 20 (2005) 107–124.
[9] O. Hirzallah, F. Kittaneh, Matrix Young inequalities for the Hilbert–Schmidt norm, Linear Algebra Appl. 308 (2000) 77–84.
[10] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[11] F. Kittaneh, Norm inequalities for fractional powers of positive operators, Lett. Math. Phys. 27 (1993) 279–285.
[12] H. Kosaki, Arithmetic–geometric mean and related inequalities for operators, J. Funct. Anal. 156 (1998) 429–451.