MONOGRAPHS IN INEQUALITIES 4

Recent Developments of Mond-Pečarić Method in Operator Inequalities

Inequalities for bounded selfadjoint operators on a Hilbert space, II

Masatoshi Fujii, Jadranka Mićić Hot, Josip Pečarić and Yuki Seo

ELEMENT

Recent Developments of Mond-Pečarić Method in Operator Inequalities

Inequalities for bounded selfadjoint operators on a Hilbert space, II

Masatoshi Fujii
Department of Mathematics
Osaka Kyoiku University
Osaka, Japan

Jadranka Mićić Hot
Faculty of Mechanical Engineering and Naval Architecture
University of Zagreb
Zagreb, Croatia

Josip Pečarić
Faculty of Textile Technology
University of Zagreb
Zagreb, Croatia

Yuki Seo
Department of Mathematics Education
Osaka Kyoiku University
Osaka, Japan

ELEMENT
Zagreb, 2012

MONOGRAPHS IN INEQUALITIES 4

Recent Developments of Mond-Pečarić Method in Operator Inequalities

Inequalities for bounded selfadjoint operators on a Hilbert space, II

Masatoshi Fujii, Jadranka Mićić Hot, Josip Pečarić and Yuki Seo

Consulting Editors

Ivan Perić
Faculty of Food Technology and Biotechnology
University of Zagreb
Zagreb, Croatia

Anita Matković
Faculty of Electrical Engineering, Mechanical
Engineering and Naval Architecture
University of Split
Split, Croatia

1st edition

Printed in Croatia.

A CIP catalogue record for this book is available from the National and University Library in Zagreb under 830846.

ISBN 978-953-197-575-9

Now when the little Dwarf heard *that he was to dance* a second time before the Infanta, and *by her own express command*, he was so proud that he ran out into the garden, kissing the white rose in an absurd ecstasy of pleasure, and making the most uncouth and clumsy gestures of delight.

"The Birthday of the Infanta"

by OSCAR WILDE

What inspired us to write this book

In the last five years, following the publishing of the first book about applying the Mond-Pečarić method in operator inequalities, many new results were obtained by using said method. That has inspired us write a new book. We have chosen important and interesting chapters, which were (mostly) published in many mathematical journals and presented at international conferences.

Preface

In the field of operator theory, an inequality due to L. V. Kantorovich provides an example of "sailing upstream", regarding the usual course of development of converse inequalities. However, we do not understand how L. V. Kantorovich interpreted the meaning of that inequality. Actually, the inequality only appeared as a Lemma to solve a certain problem. Those who have immortalized that inequality as "the Kantorovich inequality" are Greub and Rheinboldt. Based on an inequality due to Kantorovich, they gave a beautiful and simple formulation in terms of operators as follows: If A is a positive operator on a Hilbert space H such that $mI \leq A \leq MI$ for some scalars $0 < m \leq M$, then

$$\langle Ax, x\rangle\langle A^{-1}x, x\rangle \leq \frac{(M+m)^2}{4Mm}$$

holds for every unit vector $x \in H$.

Afterwards, in the course of a generalization by Ky Fan, and a converse of the arithmetic-geometric mean inequality by Specht, Mond and Pečarić give a definitive meaning to "the Kantorovich inequality". Namely, the Kantorovich inequality is a special case of the converse of Jensen's inequality: Under the same conditions as above the inequality

$$\langle A^{-1}x, x\rangle \leq \frac{(M+m)^2}{4Mm}\langle Ax, x\rangle^{-1}$$

holds for every unit vector $x \in H$ and it estimates the upper bounds of the ratio in Jensen's inequality for $f(t) = t^{-1}$. In the 1990's, Mond and Pečarić formulate directly the converse of various Jensen type inequalities. It might be said that, owing to their approach, the position of the Kantorovich inequality in the operator theory becomes clear for the first time. Furthermore, in the background of the Kantorovich inequality, they find the viewpoint for the converse of means, that is to say, the Kantorovich inequality is the converse of the arithmetic-harmonic means inequality: Under the same conditions as above the inequality

$$\langle Ax, x\rangle \leq \frac{(M+m)^2}{4Mm}\langle A^{-1}x, x\rangle^{-1}$$

holds for every unit vector $x \in H$.

In order to carry out a converse evaluation, a considerable amount of laborious manual calculation is required, including a complicated calculation depending on each particular case. In a long research series, Mond and Pečarić establish the method which gives the converse to Jensen's inequality associated with convex functions. This principle yields a rich harvest in a field of operator inequalities. We call it the Mond-Pečarić method for convex functions. One of the most important features of the Mond-Pečarić method is that it offers a totally new viewpoint in the field of operator inequalities: Let Φ be a normalized positive linear mapping on $B(H)$ and f an operator convex function on an interval I. Then Davis-Choi-Jensen's inequality asserts that

$$f(\Phi(A)) \le \Phi(f(A)) \tag{$\star$}$$

holds for every self-adjoint operator A on a Hilbert space H whose spectrum is contained in I. The operator convexity plays an essential role in the result above, that is, $(\star)$ would be false if we replace operator convexity by general convexity. We have no relation whatsoever between $f(\Phi(A)$ and $\Phi(f(A))$ under the operator ordering, but even so the Mond-Pečarić method brings us the following estimate:

$$\frac{1}{K(m,M,f)}\Phi(f(A)) \le f(\Phi(A)) \le K(m,M,f)\Phi(f(A))$$

where

$$K(m,M,f) = \max\left\{\frac{1}{f(t)}\left(\frac{f(M)-f(m)}{M-m}(t-m)+f(m)\right) : t \in [m,M]\right\}.$$

This book is devoted to the recent developments of the Mond-Pečarić method in the field of self-adjoint operators on a Hilbert space.

This book consists of eleven chapters:

In Chapter 1 we give a very brief and quick review of the basic facts about a Hilbert space and (bounded linear) operators on a Hilbert space, which will recur throughout the book.

In Chapter 2 we tell the history of the Kantorovich inequality, and describe how the Kantorovich inequality develops in the field of operator inequalities. In such context, the method for convex functions established by Mond and Pečarić (commonly known as "the Mond-Pečarić method") has outlined a more complete picture of that inequality in the field of operator inequalities. We discuss ratio and difference type converses of operator versions of Jensen's inequality. These constants in terms of spectra of given self-adjoint operators have many interesting properties and are connected with a closed relation, and play an essential role in the remainder of this book.

In Chapter 3 we explain fundamental operator inequalities related to the Furuta inequality. The base point is the Löwner-Heinz inequality. It induces weighted geometric means, which serves as an excellent technical tool. The chaotic order $\log A \ge \log B$ is conceptually important in the late discussion.

In Chapter 4 we study the order preserving operator inequality in another direction which differs from the Furuta inequality. We investigate the Kantorovich type inequalities related to the operator ordering and the chaotic one.

In Chapter 5 as applications of the Mond-Pečarić method for convex functions, we discuss inequalities involving the operator norm. Among others, we show a converse of the Araki-Cordes inequality, the norm inequality of several geometric means and a complement of the Ando-Hiai inequality. Also, we discuss Hölder's inequality and its converses in connection with the operator geometric mean.

In Chapter 6 we define the geometric mean of n operators due to Ando-Li-Mathias and Lowson-Lim. We present an alternative proof of the power convergence of the symmetrization procedure on the weighted geometric mean due to Lawson and Lim. We show a converse of the weighted arithmetic-geometric mean inequality of n operators.

In Chapter 7 we give some differential-geometrical structure of operators. The space of positive invertible operators of a unital C^*-algebra has the natural structure of a reductive homogenous manifold with a Finsler metric. Then a pair of points A and B can be joined by a unique geodesic $A \#_t B$ for $t \in [0,1]$ and we consider estimates of the upper bounds for the difference between the geodesic and extended interpolation paths in terms of the spectra of positive operators.

In Chapter 8 we give some properties of Mercer's type inequalities. A variant of Jensen's operator inequality for convex functions, which is a generalization of Mercer's result, is proved. We show a monotonicity property for Mercer's power means for operators, and a comparison theorem for quasi-arithmetic means for operators.

In Chapter 9 a general formulation of Jensen's operator inequality for some non-unital fields of positive linear mappings is given. Next, we consider different types of converse inequalities. We discuss the ordering among power functions in a general setting. We get the order among power means and some comparison theorems for quasi-arithmetic means. We also give a refined calculation of bounds in converses of Jensen's operator inequality.

In Chapter 10 we give Jensen's operator inequality without operator convexity. We observe this inequality for $n-$tuples of self-adjoint operators, unital $n-$tuples of positive linear mappings and real valued convex functions with conditions on the operators bounds. In the present context, we also give an extension and a refinement of Jensen's operator inequality. As an application we get the order among quasi-arithmetic operator means.

In Chapter 11 we observe some operator versions of Bohr's inequality. Using a general result involving matrix ordering, we derive several inequalities of Bohr's type. Furthermore, we present an approach to Bohr's inequality based on a generalization of the parallelogram theorem with absolute values of operators. Finally, applying Jensen's operator inequality we get a generalization of Bohr's inequality.

Notation

$\mathbb{R}$	the set of all real numbers
$\mathbb{C}$	the set of all complex numbers
λ, μ, ν, etc.	scalars
H, K, L, etc.	Hilbert spaces over $\mathbb{C}$
x, y, z, etc.	vectors in H
$\langle x, y \rangle$	the inner product of two vectors x and y
$\|\|x\|\|$	the norm of a vector x
$B(H)$	the C*-algebra of all bounded linear operators on a Hilbert space H
A, B, C, etc.	linear operators in $(H \to H)$
$\|\|A\|\|$	the operator norm of an operator A
$\|A\|$	the absolute value of an operator A
I_H	the identity operator in $B(H)$
O	the zero operator
$\mathsf{Sp}(A)$	the spectrum of an operator A
$\ker A$	the kernel of an operator A
$\operatorname{ran} A$	the range of an operator A

$[A]$	the range projection of an operator A
$r(A)$	the spectral radius of an operator A
$A \geq 0$	a positive operator, $\langle Ax, x\rangle \geq 0$ for all $x \in H$
$A > 0$	a strictly positive operator, A is positive and invertible
$A \geq B$	the usual operator ordering among operators A and B
$A \gg B$	the chaotic ordering among operators $A > 0$ and $B > 0$
$A \nabla_t B$	the weighted arithmetic operator mean
$A[n,t](A_1,\ldots,A_n)$	the weighted arithmetic operator mean of n operators
$A \,!_t\, B$	the weighted harmonic operator mean
$H[n,t](A_1,\ldots,A_n)$	the weighted harmonic operator mean of n operators
$A \,\#_t\, B$	the weighted geometric operator mean
$A \natural_t B$	the binary operation of $A > 0$ and B for $t \notin [0,1]$
$G[n,t](A_1,\ldots,A_n)$	the weighted geometric operator mean of n operators
$A \diamondsuit_t B$	the weighted chaotically geometric operator mean
$\mathrm{e}^{A[n,t](\log A_1,\ldots,\log A_n)}$	the weighted chaotically geometric operator mean of n operators
$S(A\|B)$	the relative operator entropy of A and B, $A,B > 0$
$A\, m_{r,t}\, B$	the interpolational path from A to B, $A,B > 0$
$D_\alpha(A,B)$	α-operator divergence, $A,B > 0$
$d(A,B)$	the Thompson metric on the convex cone of positive invertible operators A and B

$\mathscr{A}, \mathscr{B}, \mathscr{C}$ etc.	unital C*-algebras
$\mathbf{1}$	the identity element in unital C*-algebra
Φ, Ψ, Ω, etc.	positive linear mappings on C*-algebras
$P_k[\mathscr{A}, \mathscr{B}]$	the set of all fields $(\Phi_t)_{t\in T}$ of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$, such that $\int_T \Phi_t(\mathbf{1})\, d\mu(t) = k\mathbf{1}$ for some scalar $k > 0$
$M_r(\mathbf{A}, \mathbf{\Phi})$	the power operator mean of order $r \in \mathbb{R}$
$M_\varphi(\mathbf{A}, \mathbf{\Phi})$	the quasi-arithmetic operator mean generated by a function φ
$\widetilde{M}_r(\mathbf{A}, \mathbf{\Phi})$	Mercer's power operator mean of order $r \in \mathbb{R}$
$\widetilde{M}_\varphi(\mathbf{A}, \mathbf{\Phi})$	the quasi-arithmetic operator mean of Mercer's type generated by a function φ

Contents

Chapter 1

Preliminaries

In this chapter, we review the basic concepts of a Hilbert space and (bounded linear) operators on a Hilbert space, which will recur throughout the book.

1.1 Hilbert space and operators

Definition 1.1 *A complex vector space H is called an inner product space if to each pairs of vectors x and y in H is associated a complex number $\langle x,y\rangle$, called the inner product of x and y, such that the following rules hold:*

(i) *For $x,y \in H$, $\langle x,y\rangle = \overline{\langle y,x\rangle}$, where the bar denotes complex conjugation.*

(ii) *If x,y and $z \in H$ and $\alpha,\beta \in \mathbb{C}$, then $\langle \alpha x+\beta y,z\rangle = \alpha\langle x,z\rangle + \beta\langle y,z\rangle$.*

(iii) *$\langle x,x\rangle \geq 0$ for all $x \in H$ and equal to zero if and only if x is the zero vector.*

Theorem 1.1 (SCHWARZ INEQUALITY) *Let H be an inner product space. If x and $y \in H$, then*

$$|\langle x,y\rangle|^2 \leq \langle x,x\rangle\langle y,y\rangle \tag{1.1}$$

and the equality holds if and only if x and y are linearly dependent.

Proof. If $y = 0$, then the inequality (1.1) holds. Suppose that $y \neq 0$ and put

$$e = \frac{1}{\sqrt{\langle y, y\rangle}} y.$$

Then we have

$$\begin{aligned} 0 &\leq \langle x - \langle x, e\rangle e, x - \langle x, e\rangle e\rangle \\ &= \langle x, x\rangle - \overline{\langle x, e\rangle}\langle x, e\rangle - \langle x, e\rangle\langle e, x\rangle + |\langle x, e\rangle|^2\langle e, e\rangle \\ &= \langle x, x\rangle - 2|\langle x, e\rangle|^2 + |\langle x, e\rangle|^2 \\ &= \langle x, x\rangle - |\langle x, e\rangle|^2 \end{aligned}$$

and hence $|\langle x, e\rangle|^2 \leq \langle x, x\rangle$. Therefore it follows that $|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$.

If the equality holds in the inequality above, then we have $x - \langle x, e\rangle e = 0$, and so x and y are linearly dependent. Conversely, if x and y are linearly dependent, that is, there exists a constant $\alpha \in \mathbb{C}$ such that $x = \alpha y \neq 0$, then it follows that

$$|\langle x, y\rangle|^2 = |\langle \alpha y, y\rangle|^2 = |\alpha|^2|\langle y, y\rangle|^2 = \langle \alpha y, \alpha y\rangle\langle y, y\rangle = \langle x, x\rangle\langle y, y\rangle.$$

We can prove it in the case of $y = \alpha x$ in the same way. □

Let H be an inner product space. Put

$$\|x\| = \sqrt{\langle x, x\rangle} \qquad \text{for all } x \in H.$$

Then it follows that $\|\cdot\|$ is a norm on H:

(i) Positivity: $\|x\| \geq 0$ and $x = 0$ if and only if $\|x\| = 0$.

(ii) Homogeneity: $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{C}$.

(iii) Triangular inequality: $\|x + y\| \leq \|x\| + \|y\|$.

In fact, positivity and homogeneity are obvious by Definition 1.1. Triangular inequality follows from

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

by Schwarz's inequality (Theorem 1.1). Therefore, $\|x\|$ is a norm on H.

Definition 1.2 *If an inner product space H is complete with respect to the norm derived from the inner product, then H is said to be a Hilbert space.*

Some examples of Hilbert spaces will now be given.

Example 1.1 *The space* $\mathbb{C}^n$ *of all n-tuples of complex numbers with the inner product between* $x = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ *and* $y = (\beta_1, \beta_2, \cdots, \beta_n)$ *given by*

$$\langle x, y \rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$$

is a Hilbert space.

Example 1.2 *The space* l_2 *of all sequences of complex numbers* $(\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots)$ *with*

$$\sum_{i=1}^{\infty} |\alpha_i|^2 < \infty$$

and the inner product between $x = (\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots)$ *and* $y = (\beta_1, \beta_2, \cdots, \beta_n, \cdots)$ *given by*

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \alpha_i \overline{\beta_i}$$

is a Hilbert space.

A linear operator A on a Hilbert space H is said to be bounded if there exists $c > 0$ such that $\|Ax\| \leq c\|x\|$ for all $x \in H$. Let us define $\|A\|$ by

$$\|A\| = \inf\{c > 0 : \|Ax\| \leq c\|x\| \quad \text{for all } x \in H.\}$$

Then $\|A\|$ is said to be the operator norm of A. By definition,

$$\|Ax\| \leq \|A\|\|x\| \qquad \text{for all } x \in H.$$

In fact, for each $x \neq 0$, $\|Ax\| \leq c\|x\|$ implies $\frac{\|Ax\|}{\|x\|} \leq c$. Taking the inf of c, we have $\frac{\|Ax\|}{\|x\|} \leq \|A\|$.

We begin by adopting the word "operator" to mean a bounded linear operator.

$B(H)$ will now denote the algebra of all bounded linear operators on a Hilbert space $H \neq \{0\}$ and I_H stands for the identity operator.

The following lemma shows some characterizations of the operator norm.

Lemma 1.1 *For any operator* $A \in B(H)$*, the following formulae hold:*

$$\begin{aligned} \|A\| &= \sup\{\|Ax\| \ : \ \|x\| = 1,\ x \in H\} \\ &= \sup\left\{\frac{\|Ax\|}{\|x\|} \ : \ x \neq 0,\ x \in H\right\} \\ &= \sup\{|\langle Ax, y\rangle| \ : \ \|x\| = \|y\| = 1,\ x, y \in H\} \end{aligned}$$

Proof. Put

$$\gamma_1 = \sup\{\|Ax\| : \|x\| = 1\} \quad \text{and} \quad \gamma_2 = \sup\left\{\frac{\|Ax\|}{\|x\|} : x \neq 0\right\}.$$

For $\|x\| = 1$, we have $\|Ax\| \le \|A\|\|x\| = \|A\|$ and hence $\gamma_1 \le \|A\|$. For $x \ne 0$, we have

$$\frac{\|Ax\|}{\|x\|} = \|A\frac{x}{\|x\|}\| \le \gamma_1$$

and hence $\gamma_2 \le \gamma_1$. For an arbitrary $\varepsilon > 0$, there exists a nonzero vector $x \in H$ such that $(\|A\| - \varepsilon)\|x\| < \|Ax\|$ and hence

$$\|A\| - \varepsilon < \frac{\|Ax\|}{\|x\|} \le \gamma_2.$$

This fact implies $\|A\| \le \gamma_2$. Therefore we have $\|A\| = \gamma_1 = \gamma_2$.

Put

$$\gamma_3 = \sup\{|\langle Ax, y\rangle| : \|x\| = 1, \|y\| = 1\}.$$

Since $|\langle Ax, y\rangle| \le \|Ax\|\|y\| = \|Ax\| \le \gamma_1$ for $\|x\| = \|y\| = 1$, we have $\gamma_3 \le \gamma_1$. Conversely, for $Ax \ne 0$, we have

$$\|Ax\| = |\langle Ax, \frac{Ax}{\|Ax\|}\rangle| \le \gamma_3$$

and hence $\gamma_1 \le \gamma_3$. Therefore the proof is complete. □

Theorem 1.2 *The following properties hold for $A, B \in B(H)$:*

(i) *If $A \ne O$, then $\|A\| > 0$,*

(ii) $\|\alpha A\| = |\alpha|\|A\|$ *for all $\alpha \in \mathbb{C}$,*

(iii) $\|A + B\| \le \|A\| + \|B\|$,

(iv) $\|AB\| \le \|A\|\|B\|$.

Proof.

(i) If $A \ne O$, then there exists a nonzero vector $x \in H$ such that $Ax \ne 0$. Hence $0 < \|Ax\| \le \|A\|\|x\|$, therefore $\|A\| > 0$.

(ii) If $\alpha = 0$, then $\|\alpha A\| = \|O\| = 0 = |\alpha|\|A\|$. If $\alpha \ne 0$, then

$$\begin{aligned} \|\alpha A\| &= \sup\{\|(\alpha A)x\| : \|x\| = 1\} \\ &= \sup\{|\alpha|\|Ax\| : \|x\| = 1\} \\ &= |\alpha| \sup\{\|Ax\| : \|x\| = 1\} = |\alpha|\|A\|. \end{aligned}$$

(iii) If $\|x\| = 1$, then $\|(A + B)x\| = \|Ax + Bx\| \le \|Ax\| + \|Bx\| \le \|A\| + \|B\|$, therefore we have

$$\|A + B\| = \sup\{\|(A + B)x\| : \|x\| = 1\} \le \|A\| + \|B\|.$$

(iv) If $\|x\| = 1$, then $\|(AB)x\| = \|A(Bx)\| \le \|A\|\|Bx\| \le \|A\|\|B\|$, therefore we have

$$\|AB\| = \sup\{\|(AB)x\| : \|x\| = 1\} \le \|A\|\|B\|.$$

□

Theorem 1.3 (RIESZ REPRESENTATION THEOREM) *For each bounded linear functional f from H to $\mathbb{C}$, there exists a unique $y \in H$ such that*

$$f(x) = \langle x, y \rangle \qquad \text{for all } x \in H.$$

Moreover, $\|f\| = \|y\|$.

Proof. Define $\mathscr{M} = \{x \in H : f(x) = 0\}$. Then $\mathscr{M}$ is closed. If $\mathscr{M} = H$, then $f = 0$ and we can choose $y = 0$. If $\mathscr{M} \neq H$, then $\mathscr{M}^{\perp} \neq \{0\}$. For $x_0 \in \mathscr{M}^{\perp} \backslash \{0\}$, we have $f(x_0) \neq 0$. Since

$$f(x - \frac{f(x)}{f(x_0)} x_0) = f(x) - \frac{f(x)}{f(x_0)} f(x_0) = 0 \qquad \text{for all } x \in H,$$

it follows that $x - \frac{f(x)}{f(x_0)} x_0 \in \mathscr{M}$. Hence we have

$$\langle x - \frac{f(x)}{f(x_0)} x_0, x_0 \rangle = 0$$

and $\langle x, x_0 \rangle = \frac{f(x)}{f(x_0)} \|x_0\|^2$. If we put $y = \frac{\overline{f(x_0)}}{\|x_0\|^2} x_0$, then we have $f(x) = \langle x, y \rangle$ for all $x \in H$.

For the uniqueness, suppose that $f(x) = \langle x, y \rangle = \langle x, z \rangle$ for all $x \in H$. In this case, $\langle x, y - z \rangle = 0$ for all $x \in H$ implies $y - z = 0$.

Finally,

$$|f(x)| = |\langle x, y \rangle| \leq \|x\| \|y\|$$

implies $\|f\| \leq \|y\|$. Conversely,

$$\|y\|^2 = |\langle y, y \rangle| = |f(y)| \leq \|f\| \|y\|$$

implies $\|y\| \leq \|f\|$. Therefore, we have $\|f\| = \|y\|$. □

For a fixed $A \in B(H)$, a functional on H defined by

$$x \mapsto \langle Ax, y \rangle \in \mathbb{C}$$

is bounded linear on H. By the Riesz representation theorem, there exists a unique $y^* \in H$ such that

$$\langle Ax, y \rangle = \langle x, y^* \rangle \qquad \text{for all } x \in H.$$

We now define

$$A^* : y \mapsto y^*,$$

the mapping A^* being called the adjoint of A. In summary,

$$\langle Ax, y \rangle = \langle x, A^* y \rangle \qquad \text{for all } x, y \in H.$$

Theorem 1.4 *The adjoint operation is closed in $B(H)$ and moreover*

(i) $\|A^*\| = \|A\|$,

(ii) $\|A^* A\| = \|A\|^2$.

Proof.
(i): For $y_1, y_2 \in H$ and $\alpha_1, \alpha_2 \in \mathbb{C}$,

$$\begin{aligned}\langle Ax, \alpha_1 y_1 + \alpha_2 y_2 \rangle &= \overline{\alpha_1}\langle Ax, y_1\rangle + \overline{\alpha_2}\langle Ax, y_2\rangle \\ &= \overline{\alpha_1}\langle x, A^* y_1\rangle + \overline{\alpha_2}\langle x, A^* y_2\rangle \\ &= \langle x, \alpha_1 A^* y_1 + \alpha_2 A^* y_2\rangle \qquad \text{for all } x \in H.\end{aligned}$$

This implies $A^*(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 A^* y_1 + \alpha_2 A^* y_2$ and A^* is linear. Next,

$$\begin{aligned}\|A^* y\| &= \sup\{|\langle x, A^* y\rangle| : \|x\| = 1\} \\ &= \sup\{|\langle Ax, y\rangle| : \|x\| = 1\} \\ &\leq \sup\{\|Ax\|\|y\| : \|x\| = 1\} = \|A\|\|y\|,\end{aligned}$$

hence A^* is bounded and $\|A^*\| \leq \|A\|$. Therefore, the adjoint operation is closed in $B(H)$. Since $(A^*)^* = A$, we have

$$\|A\| = \|(A^*)^*\| \leq \|A^*\|$$

and hence $\|A^*\| = \|A\|$.
(ii): Since $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^*Ax, x\rangle \leq \|A^*A\|\|x\|^2$ for every $x \in H$, we have $\|A\|^2 \leq \|A^*A\|$.

On the other hand, (i) gives $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$. Hence the equality

$$\|A^*A\| = \|A\|^2$$

holds for every $A \in B(H)$. □

1.2 Self-adjoint operators

We present relevant classes of operators:

Definition 1.3 *An operator $A \in B(H)$ is said to be*

(i) *self-adjoint or Hermitian if $A = A^*$,*

(ii) *positive if $\langle Ax, x\rangle \geq 0$ for x in H,*

(iii) *unitary if $A^*A = AA^* = I_H$,*

(iv) *isometry if $A^*A = I_H$,*

(v) *projection if $A = A^* = A^2$.*

The following theorem gives characterizations of self-adjoint operators.

Theorem 1.5 *If $A \in B(H)$, the following three statements are mutually equivalent.*

(i) *A is self-adjoint.*

(ii) $\langle Ax, y\rangle = \langle x, Ay\rangle$ *for all $x, y \in H$.*

(iii) $\langle Ax, x\rangle \in \mathbb{R}$ *for all $x \in H$.*

Proof.
(i) $\Longleftrightarrow$ (ii): If A is self-adjoint, then $\langle Ax, y\rangle = \langle x, A^* y\rangle = \langle x, Ay\rangle$. Conversely suppose that (ii) holds. Since $\langle x, A^* y\rangle = \langle x, Ay\rangle$ for all $x, y \in H$, we have $A^* y = Ay$, so that $A = A^*$.
(ii) $\Longleftrightarrow$ (iii): If we put $y = x$ in (ii), then

$$\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle},$$

so $\langle Ax, x\rangle$ is real. Thus (ii) implies (iii). Finally, suppose that (iii) holds. For each x and $y \in H$, if we put $w = x + y$, then $\langle Aw, w\rangle$ is real, or $\langle Aw, w\rangle = \langle w, Aw\rangle$. Expanding $\langle A(x+y), x+y\rangle = \langle x+y, A(x+y)\rangle$, we have

$$\langle Ax, y\rangle + \langle Ay, x\rangle = \langle x, Ay\rangle + \langle y, Ax\rangle$$

and $\mathrm{Im}\langle Ax, y\rangle = \mathrm{Im}\langle x, Ay\rangle$. Replacing x by ix, we have $\mathrm{Re}\langle Ax, y\rangle = \mathrm{Re}\langle x, Ay\rangle$. Therefore it follows that $\langle Ax, y\rangle = \langle x, Ay\rangle$. Thus (iii) implies (ii). □

The spectrum of an operator A is the set

$$\mathsf{Sp}(A) = \{\lambda \in \mathbb{C} : A - \lambda I_H \text{ is not invertible in } B(H)\}.$$

The spectrum $\mathsf{Sp}(A)$ is nonempty and compact. An operator A on a Hilbert space H is bounded below if there exists $\varepsilon > 0$ such that $\|Ax\| \geq \varepsilon \|x\|$ for every $x \in H$. As a useful criterion for the invertibility of an operator, it is well known that A is invertible if and only if both A and A^* are bounded below.

The spectral radius $r(A)$ of an operator A is defined by

$$r(A) = \sup\{|\alpha| : \alpha \in \mathsf{Sp}(A)\}.$$

Then we have the following relation between the operator norm and the spectral radius.

Theorem 1.6 *For an operator A, the spectral radius is not greater than the operator norm:*

$$r(A) \leq \|A\|.$$

Proof. If $|\alpha| > \|A\|$, then $I_H - \alpha^{-1}A$ is invertible and hence $A - \alpha I_H$ is so. Therefore we have $\alpha \notin \mathsf{Sp}(A)$ and this implies $r(A) \leq \|A\|$. □

Let A be a self-adjoint operator on a Hilbert space H. We define

$$m_A = \inf_{\|x\|=1} \langle Ax, x\rangle \quad \text{and} \quad M_A = \sup_{\|x\|=1} \langle Ax, x\rangle. \tag{1.2}$$

Theorem 1.7 *For a self-adjoint operator A,* $\mathsf{Sp}(A)$ *is real and* $\mathsf{Sp}(A) \subseteq [m_A, M_A]$.

Proof. If $\lambda = \alpha + i\beta$ with α, β real and $\beta \neq 0$, then we must show that $A - \lambda I_H$ is invertible. Put $B = \frac{1}{\beta}(A - \alpha I_H)$. Since B is self-adjoint and $B - iI_H = \frac{1}{\beta}(A - \lambda I_H)$, it follows that $A - \lambda I_H$ is invertible if and only if $B - iI_H$ is invertible. For every $x \in H$, we have

$$\begin{aligned} \|(B \pm iI_H)x\|^2 &= \|Bx\|^2 - i\langle x, Bx\rangle + i\langle Bx, x\rangle + \|x\|^2 \\ &= \|Bx\|^2 + \|x\|^2 \geq \|x\|^2, \end{aligned}$$

so $B - iI_H$ and $(B - iI_H)^*$ are bounded below. Therefore $B - iI_H$ is invertible, and hence the spectrum of a self-adjoint operator is real.

Next, to prove $\mathsf{Sp}(A) \subset [m_A, M_A]$, it is enough to show that $\lambda > M_A$ implies $\lambda \notin \mathsf{Sp}(A)$. If $\lambda > M_A$ and $\varepsilon = \lambda - M_A > 0$, then

$$\begin{aligned} \langle(\lambda I_H - A)x, x\rangle &= \lambda\langle x, x\rangle - \langle Ax, x\rangle \geq \lambda\langle x, x\rangle - M_A\langle x, x\rangle \\ &= \varepsilon\langle x, x\rangle \geq 0 \qquad \text{by the definition of } M_A. \end{aligned}$$

Hence it follows that $\|(A - \lambda I_H)x\| \geq \varepsilon\|x\|$ for every $x \in H$, so, $A - \lambda I_H$ is bounded below. Since $A - \lambda I_H$ is self-adjoint, it follows that $A - \lambda I_H$ is invertible and $\lambda \notin \mathsf{Sp}(A)$. □

Definition 1.4 *Let A and B be self-adjoint operators on H. We write $A \geq B$ if $A - B$ is positive, i.e. $\langle Ax, x\rangle \geq \langle Bx, x\rangle$ for every $x \in H$. In particular, we write $A \geq 0$ if A is positive, $A > 0$ if A is positive and invertible.*

Now, we review the continuous functional calculus. A rudimentary functional calculus for an operator A can be defined as follows: For a polynomial $p(t) = \sum_{j=0}^{k} \alpha_j t^j$, define

$$p(A) = \alpha_0 I_H + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_k A^k.$$

The mapping $p \to p(A)$ is a homomorphism from the algebra of polynomials to the algebra of operators. The extension of this mapping to larger algebras of functions is really significant in operator theory.

Let A be a self-adjoint operator on a Hilbert space H. Then the Gelfand mapping establishes a $*$-isometrically isomorphism Φ between C*-algebra $C(\mathsf{Sp}(A))$ of all continuous functions on $\mathsf{Sp}(A)$ and C*-algebra $C^*(A)$ generated by A and the identity operator I_H on H as follows: For $f, g \in C(\mathsf{Sp}(A))$ and $\alpha, \beta \in \mathbb{C}$

(i) $\Phi(\alpha f + \beta g) = \alpha\Phi(f) + \beta\Phi(g)$,

(ii) $\Phi(fg) = \Phi(f)\Phi(g)$ and $\Phi(\overline{f}) = \Phi(f)^*$,

(iii) $\|\Phi(f)\| = \|f\| \left(:= \sup_{t \in \mathsf{Sp}(A)} |f(t)| \right)$,

(iv) $\Phi(f_0) = I_H$ and $\Phi(f_1) = A$, where $f_0(t) = 1$ and $f_1(t) = t$.

With this notation, we define

$$f(A) = \Phi(f)$$

for all $f \in C(\mathrm{Sp}(A))$ and we call it the continuous functional calculus for a self-adjoint operator A. It is an extension of $p(A)$ for a polynomial p. The continuous functional calculus is applicable.

Theorem 1.8 *Let A be a self-adjoint operator on H.*

(i) $f \in C(\mathrm{Sp}(A))$ *and* $f \geq 0$ *implies* $f(A) \geq 0$.

(ii) $f, g \in C(\mathrm{Sp}(A))$ *and* $f \geq g$ *implies* $f(A) \geq g(A)$.

(iii) $A \geq 0$ *and* $f_{1/2}(t) = \sqrt{t}$ *implies* $f_{1/2}(A) = A^{1/2}$.

(iv) $f_s(t) = |t|$ *implies* $f_s(A) = |A|$.

Proof.
(i) Since $f \geq 0$, we can choose $g = \sqrt{f} \in C(\mathrm{Sp}(A))$ and $f = g^2 = \overline{g}g$. Hence we have $f(A) = g(A)^* g(A) \geq 0$.
(ii) follows from (i).
(iii) Since $A \geq 0$, it follows from Theorem 1.7 that $f_{1/2}(t) = \sqrt{t} \in C(\mathrm{Sp}(A))$. Also, $f_1 = f_{1/2}^2$ implies $A = f_1(A) = f_{1/2}(A)^2$. By (i), we have $f_{1/2}(A) \geq 0$ and hence $f_{1/2}(A) = A^{1/2}$.
(iv) $f_s^2 = f_1^2$ implies $f_s(A)^2 = A^2 = |A|^2$. Since $f_s(A) \geq 0$, we have $f_s(A) = |A|$. □

We remark that the absolute value of an operator A is defined by $|A| = (A^*A)^{1/2}$.

Theorem 1.9 *An operator A is positive if and only if there is an operator B such that $A = B^*B$.*

Proof. If A is positive, take $B = \sqrt{A}$. If $A = B^*B$, then $\langle Ax, x\rangle = \langle B^*Bx, x\rangle = \|Bx\|^2 \geq 0$ for every $x \in H$. This yields that A is positive. □

Theorem 1.10 (GENERALIZED SCHWARZ'S INEQUALITY) *If A is positive, then*

$$|\langle Ax, y\rangle|^2 \leq \langle Ax, x\rangle \langle Ay, y\rangle$$

for every $x, y \in H$.

Proof. It follows from Theorem 1.1 that

$$|\langle Ax, y\rangle|^2 = |\langle A^{1/2}x, A^{1/2}y\rangle|^2 \leq \|A^{1/2}x\|^2 \|A^{1/2}y\|^2 = \langle Ax, x\rangle \langle Ay, y\rangle.$$

□

Theorem 1.11 *Let A be a self-adjoint operator on H. Then*

(i) $m_A I_H \leq A \leq M_A I_H$,

(ii) $\|A\| = \max\{|m_A|, |M_A|\} = \sup\{|\langle Ax, x\rangle| : \|x\| = 1\}$,
where m_A and M_A are defined by (1.2)*.*

Proof. The assertion (i) is clear by definition of m_A and M_A.
Next, put $K = \max\{|m_A|, |M_A|\}$. It is easily checked that

$$K = \sup\{|\langle Ax, x\rangle| : \ \|x\| = 1\} \le \|A\|.$$

By (i), we have

$$-K\|x\|^2 \le m\|x\|^2 \le \langle Ax, x\rangle \le M\|x\|^2 \le K\|x\|^2.$$

For each $x, y \in H$, since

$$|\langle A(x+y), x+y\rangle| \le K\|x+y\|^2 \quad \text{and} \quad |\langle A(x-y), x-y\rangle| \le K\|x-y\|^2,$$

it follows that

$$|\langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle| \le K(\|x+y\|^2 + \|x-y\|^2).$$

By the parallelogram identity, we have

$$4|\mathrm{Re}\langle Ax, y\rangle| \le 2K(\|x\|^2 + \|y\|^2). \tag{1.3}$$

Put $y = \frac{\|x\|}{\|Ax\|}Ax$ for $Ax \neq 0$. Then $\|x\| = \|y\|$ and $\mathrm{Re}\langle Ax, y\rangle = \|x\|\|Ax\|$. Therefore, by (1.3) we have

$$\|Ax\| \le K\|x\|. \tag{1.4}$$

If $Ax = 0$, then (1.4) holds automatically. Hence we have $\|A\| \le K$. Therefore we have $\|A\| = K$. □

Corollary 1.1 *If A is a self-adjoint operator, then $r(A) = \|A\|$ and $\|A^n\| = \|A\|^n$ for $n \in \mathbb{N}$.*

Proof. By Theorem 1.11, it follows that $r(A) = \|A\|$. By the spectral mapping theorem, we have $p(\mathsf{Sp}(A)) = \mathsf{Sp}(p(A))$ for polynomial p. Therefore, we have $\|A\|^n = r(A)^n = r(A^n) = \|A^n\|$. □

1.3 Spectral decomposition theorem

We shall introduce the spectral decomposition theorem for self-adjoint, bounded linear operators on a Hilbert space H. To show it, we need the following notation and lemma.

Definition 1.5 *If A is an operator on a Hilbert space H, then the kernel of A, denoted by* $\ker A$*, is the closed subspace $\{x \in H : Ax = 0\}$, and the range of A, denoted by* $\mathrm{ran}\, A$*, is the subspace $\{Ax : x \in H\}$.*

Lemma 1.2 *If A is an operator on a Hilbert space H, then*

$$\ker A = (\operatorname{ran} A^*)^{\perp} \quad \textit{and} \quad \ker A^* = (\operatorname{ran} A)^{\perp}.$$

Proof. If $x \in \ker A$, then $\langle A^*y, x\rangle = \langle y, Ax\rangle = 0$ for all $y \in H$, and hence x is orthogonal to ran A^*. Conversely, if x is orthogonal to ran A^*, then $\langle Ax, y\rangle = \langle x, A^*y\rangle = 0$ for all $y \in H$, which implies $Ax = 0$. Therefore, $x \in \ker A$ and hence $\ker A = (\operatorname{ran} A^*)^{\perp}$. We have the second relation by replacing A by A^*. □

Definition 1.6 *A family of projections $\{e(\lambda) : \lambda \in \mathbb{R}\}$ is said to be a resolution of the identity if the following properties hold:*

(i) $\lambda < \lambda' \quad \Longrightarrow \quad e(\lambda) \le e(\lambda')$,

(ii) $e(-\infty) = O$ *and* $e(\infty) = I_H$,

(iii) $e(\lambda + 0) = e(\lambda) \qquad (-\infty < \lambda < \infty)$,
where $e(\lambda + 0) = \text{s-}\lim_{\mu \to \lambda + 0} e(\mu)$.

Theorem 1.12 *Let A be a self-adjoint operator on a Hilbert space H and $m = m_A, M = M_A$ as defined by* (1.2)*. Then there exists a resolution of the identity $\{e(\lambda) : \lambda \in \mathbb{R}\}$ such that*

$$A = \int_{m-0}^{M} \lambda \, \mathrm{d}e(\lambda), \quad e(m-0) = 0 \quad \textit{and} \quad e(M) = I_H.$$

In particular,

$$\langle Ax, x\rangle = \int_{m-0}^{M} \lambda \, \mathrm{d}\langle e(\lambda)x, x\rangle \qquad \textit{for every } x \in H. \tag{1.5}$$

Proof. We prove only (1.5). Put $e(\lambda) = \operatorname{proj}(\ker((A - \lambda I_H)^+))$ for $\lambda \in \mathbb{R}$, where $A^+ = (|A| + A)/2$. Then it follows that $\{e(\lambda) : \lambda \in \mathbb{R}\}$ is a resolution of the identity and $e(m-0) = 0$, $e(M) = I_H$:
(i) Let $\lambda < \lambda'$. Since $A - \lambda I_H \ge A - \lambda' I_H$, we have $(A - \lambda I_H)^+ \ge (A - \lambda' I_H)^+ \ge 0$. If $(A - \lambda I_H)^+ x = 0$, then

$$0 = \langle (A - \lambda I_H)^+ x, x\rangle \ge \langle (A - \lambda' I_H)^+ x, x\rangle \ge 0$$

and hence $(A - \lambda' I_H)x = 0$. Therefore, we have $\ker((A - \lambda I_H)^+) \subset \ker((A - \lambda' I_H)^+)$ and this implies $e(\lambda) \le e(\lambda')$.
(ii) If $x \in \operatorname{ran}(e(\lambda)) = \ker((A - \lambda I_H)^+)$, then $(A - \lambda I_H)^+ x = 0$ implies $(A - \lambda I_H)x = -(A - \lambda I_H)^- x$ and hence

$$\langle (A - \lambda I_H)x, x\rangle = -\langle (A - \lambda I_H)^- x, x\rangle \le 0.$$

Therefore we have $\langle Ax, x\rangle \le \lambda \|x\|^2$.
(iii) If $x \in \operatorname{ran}(I_H - e(\lambda)) = (\ker((A - \lambda I_H)^+))^{\perp}$, then $(A - \lambda I_H)^- x \in \ker((A - \lambda I_H)^+)$ because $(A - \lambda I_H)^+ (A - \lambda I_H)^- = 0$. Hence $\langle (A - \lambda I_H)^- x, x\rangle = 0$ and $\langle (A - \lambda I_H)x, x\rangle =$

$\langle (A-\lambda I_H)^+ x,x\rangle \geq 0$. Therefore we have $\langle Ax,x\rangle \geq \lambda \|x\|^2$. If the equality holds, then $\langle (A-\lambda I_H)^+ x,x\rangle = 0$ and hence $(A-\lambda I_H)x = 0$. Therefore we have $x \in \ker((A-\lambda I_H)^+)$ and hence $x = 0$. Summing up, $x \in \operatorname{ran}(I_H - e(\lambda))$, $x \neq 0$ implies $\langle Ax,x\rangle > \lambda \|x\|^2$.

(iv) If $\lambda < m$ and $x \in \operatorname{ran}(e(\lambda))$, then it follows from (ii) that $m\|x\|^2 \leq \langle Ax,x\rangle \leq \lambda \|x\|^2$ and hence $x = 0$. Therefore we have $e(\lambda) = O$, so that $e(m-0) = O$.

(v) If $\lambda \geq M$ and $x \in \operatorname{ran}(I_H - e(\lambda))$, then it follows from (iii) that $\lambda \|x\|^2 \leq \langle Ax,x\rangle \leq M\|x\|^2$ and hence $x = 0$. Therefore we have $I_H - e(\lambda) = O$, so that $e(\lambda) = I_H$. In particular, we have $e(M) = I_H$.

(vi) If $\lambda < m$ or $\lambda \geq M$, then it follows from (iv), (v) that $e(\lambda) = e(\lambda - 0)$. Suppose that $m \leq \lambda < M$. Put $P = e(\lambda - 0) - e(\lambda)$. For $\lambda < \lambda' < M$, we have $\operatorname{ran}(P) \subset \operatorname{ran}(e(\lambda') - e(\lambda))') - e(\lambda)) = \operatorname{ran}(e(\lambda')) \cap \operatorname{ran}(I_H - e(\lambda))$. Hence $x \in \operatorname{ran}(P)$ and $x \neq 0$ implies $\lambda \|x\|^2 < \langle Ax,x\rangle \leq \lambda' \|x\|^2$ by (ii) and (iii). As $\lambda' \to \lambda + 0$, we get $\lambda \|x\|^2 < \lambda \|x\|^2$, which is a contradiction. Therefore we have $\operatorname{ran}(P) = \{0\}$, so that $P = e(\lambda+0) - e(\lambda) = O$.

For all $\varepsilon > 0$, we choose $\delta > 0$ such that

$$\Delta : \alpha = \lambda_0 < \lambda_1 < \cdots < \lambda_n = \beta, \quad \xi_k \in [\lambda_{k-1}, \lambda_k] \quad k = 1, \cdots, n,$$

and

$$|\Delta| = \max\{\lambda_k - \lambda_{k-1} : k = 1, \cdots, n\} < \delta.$$

Since A commutes with $e(\lambda)$ for each $\lambda \in \mathbb{R}$, it follows that

$$A = \sum_{k=1}^{n} A\,(e(\lambda_k) - e(\lambda_{k-1}))\,.$$

For every $x \in H$, we have

$$\begin{aligned}
&\left| \langle Ax,x\rangle - \sum_{k=1}^{n} \xi_k \langle (e(\lambda_k) - e(\lambda_{k-1}))x,x\rangle \right| \\
&= \left| \sum_{k=1}^{n} \langle A(e(\lambda_k) - e(\lambda_{k-1}))x,x\rangle - \sum_{k=1}^{n} \xi_k \langle (e(\lambda_k) - e(\lambda_{k-1}))x,x\rangle \right| \\
&\leq \sum_{k=1}^{n} |\langle (A - \xi_k I)(e(\lambda_k) - e(\lambda_{k-1}))x, (e(\lambda_k) - e(\lambda_{k-1}))x\rangle| \\
&\leq \sum_{k=1}^{n} (\lambda_k - \lambda_{k-1}) \|(e(\lambda_k) - e(\lambda_{k-1}))x\|^2 \\
&\leq |\Delta| \|x\|^2 \leq \varepsilon.
\end{aligned}$$

Hence we have the desired result $\langle Ax,x\rangle = \int_{m-0}^{M} \lambda \, \mathrm{d}\langle e(\lambda)x,x\rangle$. □

Definition 1.7 *Let A be a self-adjoint operator on a Hilbert space H and $m = m_A, M = M_A$ as defined by* (1.2)*. For a real valued continuous function $f(\lambda)$ on $[m,M]$, a self-adjoint operator $f(A)$ is defined by*

$$f(A) = \int_{m-0}^{M} f(\lambda) \, \mathrm{d}e(\lambda).$$

In particular,

$$A^r = \int_{m-0}^{M} \lambda^r \, \mathrm{d}e(\lambda) \quad \textit{for all } r > 0 \textit{ and } \quad A^{\frac{1}{2}} = \int_{m-0}^{M} \lambda^{\frac{1}{2}} \, \mathrm{d}e(\lambda).$$

In the last part of this chapter, we present the polar decomposition for an operator.

Every complex number can be written as the product of a nonnegative number and a number of modulus one:

$$z = |z|\mathrm{e}^{i\theta} \quad \text{for a complex number } z.$$

We shall attempt a similar argument for operators on an infinite dimensional Hilbert space. Before considering this result, we need to introduce the notion of a partial isometry.

Definition 1.8 *An operator V on a Hilbert space H is a partial isometry if $\|Vx\| = \|x\|$ for $x \in (\ker V)^{\perp}$, which is called the initial space of V.*

We consider a useful characterization of partial isometries:

Lemma 1.3 *Let V be an operator on a Hilbert space H. The following are equivalent:*

(i) *V is a partial isometry.*

(ii) *V^* is a partial isometry.*

(iii) *V^*V is a projection.*

(iv) *VV^* is a projection.*

Moreover, if V is a partial isometry, then VV^* is the projection onto the range of V, while V^*V is the projection onto the initial space.

Proof. Suppose that V is a partial isometry. Since

$$\langle (I - V^*V)x, x\rangle = \langle x, x\rangle - \langle V^*Vx, x\rangle = \|x\|^2 - \|Vx\|^2 \quad \text{for } x \in H,$$

it follows that $I - V^*V$ is a positive operator. Now if x is orthogonal to ker V, then $\|Vx\| = \|x\|$ which implies that $\langle (I-V^*V)x, x\rangle = 0$. Since $\|(I-V^*V)^{1/2}x\|^2 = \langle (I-V^*V)x, x) = 0$, we have $(I - V^*V)x = 0$ or $V^*Vx = x$. Therefore, V^*V is the projection onto the initial space of V.

Conversely, if V^*V is a projection and x is orthogonal to ker V^*V, then $V^*Vx = x$. Therefore,

$$\|Vx\|^2 = \langle V^*Vx, x\rangle = \langle x, x\rangle = \|x\|^2,$$

and hence V preserves the norm on $(\ker V^*V)^{\perp}$. Moreover, if $V^*Vx = 0$, then $0 = \langle V^*Vx, x\rangle = \|Vx\|^2$ and consequently ker $V^*V =$ ker V. Therefore, V is a partial isometry, and hence (i) and (iii) are equivalent.

Similarly, we have the equivalence of (ii) and (iv).

Moreover, if V^*V is a projection, then $(VV^*)^2 = VV^*VV^* = VV^*$, since $V(V^*V) = V$. Therefore, VV^* is a projection, which completes the proof. □

We now obtain the polar decomposition for an operator.

Theorem 1.13 *If A is an operator on a Hilbert space H, then there exists a positive operator P and a partial isometry V such that $A = VP$. Moreover, V and P are unique if* $\ker P = \ker V$.

Proof. If we set $P = |A|$, then

$$\|Px\|^2 = \langle Px, Px\rangle = \langle P^*Px, x\rangle = \langle A^*Ax, x\rangle = \|Ax\|^2 \quad \text{for } x \in H.$$

Thus, if we define $\tilde{V}$ on ran P such that $\tilde{V}Px = Ax$, then $\tilde{V}$ is well defined and is isometric. Hence, $\tilde{V}$ can be extended uniquely to an isometry from clos(ran$P$) to H. If we further extend $\tilde{V}$ to H by defining it to be the zero operator on $(\text{ran}P)^\perp$, then the extended extended operator V is a partial isometry satisfying $A = VP$ and $\ker V = (\text{ran}P)^\perp = \ker P$ by Lemma 1.3.

We next consider uniqueness. Suppose $A = WQ$, where W is a partial isometry, Q is a positive operator, and $\ker W = \ker Q$. Then $P^2 = A^*A = QW^*WQ = Q^2$, since W^*W is the projection onto

$$(\ker W)^\perp = (\ker Q)^\perp = \text{clos}(\text{ran } Q).$$

Thus, by the uniqueness of the square root, we have $P = Q$ and hence $WP = VP$. Therefore, $W = V$ on ran P. But

$$(\text{ran } P)^\perp = \ker P = \ker W = \ker V$$

and hence $W = V$ on $(\text{ran } P)^\perp$. Therefore, $V = W$ and the proof is complete. □

Corollary 1.2 *If A is an operator on a Hilbert space H, then there exists a positive operator Q and a partial isometry W such that $A = QW$. Moreover, W and Q are unique if* ran $Q = (\ker Q)^\perp$.

Proof. By Theorem 1.13, we obtain a partial isometry V and a positive operator P such that $A^* = VP$. Taking adjoints we have $A = PV^*$, which is the form that we desire with $W = V^*$ and $Q = P$. Moreover, the uniqueness also follows from Theorem 1.13 since ran $W = (\ker Q)^\perp$ if and only if

$$\ker V = \ker W^* = (\text{ran } W)^\perp = (\ker Q)^{\perp\perp} = \ker P.$$

□

1.4 Notes

For our exposition we have used [276], [45], [143], [18].

Chapter 2

Kantorovich Inequality and Mond-Pečarić Method

This chapter tells the history of the Kantorovich inequality, and describes how the Kantorovich inequality has developed in the field of operator inequalities. In such context, so called "the Mond-Pečarić method" for convex functions established by Mond and Pečarić has outlined a more complete picture of that inequality in the field of operator inequalities.

2.1 History

The story of the Kantorovich inequality is a very interesting example how a mathematician creates mathematics. It provides a deep insight into how a principle raised from the Kantorovich inequality has developed in the field of operator inequalities on a Hilbert space, and perhaps, more importantly, it has initiated a new way of thinking and new methods in operator theory, noncommutative differential geometry, quantum information theory and noncommutative probability theory. We call this principle *the Mond-Pečarić method for convex functions*.

In 1959, Greub and Rheinboldt published the celebrated paper [132]. It is just the birth of the Kantorovich inequality. They stated that Kantorovich proved the following inequality.

Theorem K1 *If the sequence* $\{\gamma_k\}$ $(k = 1,2,\cdots)$ *of real numbers has the property*

$$0 < m \le \gamma_k \le M$$

and $\{\xi_k\}$ $(k = 1,2,\cdots)$ *denotes another sequence with* $\sum_{k=1}^{\infty} \xi_k^2 < \infty$, *then the inequality*

$$\sum_{k=1}^{\infty} \gamma_k \xi_k^2 \sum_{k=1}^{\infty} \frac{1}{\gamma_k} \xi_k^2 \le \frac{(M+m)^2}{4Mm} \left[\sum_{k=1}^{\infty} \xi_k^2 \right]^2 \tag{2.1}$$

holds.

It seems to be the first paper which introduced (2.1) to the world of mathematics. Moreover, they say that Kantorovich pointed out that (2.1) is a special case of the following inequality enunciated by G. Pólya and G. Szegö [253].

Theorem PS *If real numbers* a_k *and* b_k $(k = 1,\cdots,n)$ *fulfill the conditions*

$$0 < m_1 \le a_k \le M_1 \qquad \text{and} \qquad 0 < m_2 \le b_k \le M_2$$

respectively, then

$$1 \le \frac{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2}{[\sum_{k=1}^{n} a_k b_k]^2} \le \frac{(M_1M_2 + m_1m_2)^2}{4m_1m_2M_1M_2}. \tag{2.2}$$

To understand (2.1) in Theorem K1 well, if we put $\xi_k = 1/\sqrt{n}$ for $k = 1,\cdots,n$, then (2.1) implies

$$\frac{\gamma_1 + \cdots + \gamma_n}{n} \cdot \frac{\gamma_1^{-1} + \cdots + \gamma_n^{-1}}{n} \le \frac{(M+m)^2}{4Mm}. \tag{2.3}$$

Summing up, whenever γ_k's move in the closed interval $[m,M]$, the left-hand side of (2.3) does not absolutely exceed the constant $\frac{(M+m)^2}{4Mm}$. At present, the constant $\frac{(M+m)^2}{4Mm}$ is called the *Kantorovich constant*.

Greub and Rheinboldt moreover went ahead with the ideas of Kantorovich and proved the following theorem as a generalization of the Kantorovich inequality.

Theorem K2 *Given a self-adjoint operator* A *on a Hilbert space* H. *If* A *fulfills the condition*

$$mI_H \le A \le MI_H \quad \text{for some scalars } 0 < m \le M,$$

then

$$\langle x,x\rangle^2 \le \langle Ax,x\rangle\langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm}\langle x,x\rangle^2 \tag{2.4}$$

for all $x \in H$.

Though this formulation is very simple, how to generalize (2.1) might be not plain. In the case that A is matrix, then (2.4) can be expressed as follows: Put

$$A = \begin{pmatrix} \gamma_1 & & & & 0 \\ & \gamma_2 & & & \\ & & \ddots & & \\ & & & \gamma_n & \\ 0 & & & & \ddots \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \\ \vdots \end{pmatrix}.$$

Then

$$A^{-1} = \begin{pmatrix} \gamma_1^{-1} & & & & 0 \\ & \gamma_2^{-1} & & & \\ & & \ddots & & \\ & & & \gamma_n^{-1} & \\ 0 & & & & \ddots \end{pmatrix}$$

and we get

$$\langle Ax, x\rangle = \sum_{k-1}^{\infty} \gamma_k \xi_k^2 \quad \text{and} \quad \langle A^{-1}x, x\rangle = \sum_{k=1}^{\infty} \gamma_k^{-1} \xi_k^2.$$

We shall agree that (2.4) is called a generalization of the Kantorovich inequality (2.1).

Though Greub and Rheinboldt carefully cite the Kantorovich inequality, they do not tell anything about his motivation for considering the inequality (2.1). What is his motive for considering (2.1)? Thus, we shall attempt to investigate Kantorovich's original paper in this occasion. It is written in Russian and very old. We read the original paper in an English translation [156]. It seems that he was interested in the mathematical formulation of economics, as he provided a detailed commentary on how to carry out mathematical analysis in economic activities. Now, when we read [156] slowly and carefully, we find the inequality (2.1) in question, in the middle of the paper [156].

Lemma K *The inequality*

$$\sum_k \gamma_k u_k^2 \sum_k \gamma_k^{-1} u_k^2 \le \frac{1}{4}\left[\sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}}\right]^2 \left(\sum_k u_k^2\right)^2 \tag{2.5}$$

holds, m and M being the bounds of the numbers γ_k

$$0 < m \le \gamma_k \le M.$$

The coefficient in the right-hand side of (2.5) seems to be different from the one in (2.1). However, since

$$\frac{1}{4}\left[\sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}}\right]^2 = \frac{1}{4}\left[\frac{M+m}{\sqrt{Mm}}\right]^2 = \frac{(M+m)^2}{4Mm},$$

the constant of (2.5) coincides with one of (2.1). Following Kantorovich's original paper, we know that Kantorovich represents an upper bound as (2.5). Therefore the Kantorovich constant $\frac{(M+m)^2}{4Mm}$ is deformed by Greub and Rheinboldt. Examining the history of mathematics a little more, Henrici [141] pointed out that in the case of equal weights, the inequality (2.3) is due to Schweitzer [258] in 1914. How Kantorovich proved the inequality (2.5) in Lemma K is a very interesting matter:

Proof of Lemma K. We may prove it in the case of finite sums $\gamma_1 \le \gamma_2 \le \cdots \le \gamma_n$ and $\sum_{k=1}^n u_k^2 = 1$. We shall seek the maximum of

$$G = \sigma\tilde{\sigma} = \left(\sum_{k=1}^n \gamma_k u_k^2\right)\left(\sum_{k=1}^n \frac{1}{\gamma_k} u_k^2\right)$$

under the condition that $\sum_{k=1}^n u_k^2 = 1$. By using the method of Lagrange multipliers, if we equate to zero the derivatives of the function

$$F = G - \lambda\left(\sum_{k=1}^n u_k^2 - 1\right),$$

then we have

$$\frac{1}{2}\frac{\partial F}{\partial u_s} = \sigma\frac{1}{\gamma_s}u_s + \tilde{\sigma}\gamma_s u_s - \lambda u_s = 0, \quad \text{i.e. } u_s(\sigma + \tilde{\sigma}\gamma_s^2 - \lambda\gamma_s) = 0.$$

The second factor in the last expression, being a polynomial of the second degree in γ_s, can reduce to zero at not more than two values of s; let these be $s = k, l$. For the remaining values of s, u_s must be zero. But then

$$\begin{aligned} G_{\max} &= (\gamma_k u_k^2 + \gamma_l u_l^2)\left(\frac{1}{\gamma_k}u_k^2 + \frac{1}{\gamma_l}u_l^2\right) \\ &= \frac{1}{4}\left[\sqrt{\frac{\gamma_k}{\gamma_l}} + \sqrt{\frac{\gamma_l}{\gamma_k}}\right]^2 (u_k^2 + u_l^2)^2 - \frac{1}{4}\left[\sqrt{\frac{\gamma_k}{\gamma_l}} + \sqrt{\frac{\gamma_l}{\gamma_k}}\right]^2 (u_k^2 - u_l^2)^2 \\ &\le \frac{1}{4}\left[\sqrt{\frac{\gamma_k}{\gamma_l}} + \sqrt{\frac{\gamma_l}{\gamma_k}}\right]^2 \le \frac{1}{4}\left[\sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}}\right]^2. \end{aligned}$$

□

Why does Kantorovich need the inequality (2.1)? If we only read the paper due to Greub and Rheinboldt, we probably cannot fully understand those circumstances. However, having thoroughly read [156], we are able to explain the necessity of the Kantorovich inequality.

Kantorovich says that as is generally known, a significant part of the problems of mathematical physics – the majority of the linear problems of analysis – may be reduced to a problem of the extremum of quadratic functionals. This fact may be utilized, on the one hand for different theoretical investigations relating to these problems. On the other hand, it serves as a basis for direct methods of solving the problems named. A certain method of successive approximations for the solution of problems concerning the minimum of quadratic functionals, and of the linear problems connected with them, is elaborated – the method of steepest descent.

Let H be a real Hilbert space and A a self-adjoint (bounded linear) operator on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m \le M$.

We shall consider the method of steepest descent as it applies to the solution of the equation

$$L(x) = Ax - y = 0, \tag{2.6}$$

where x and y are in H. We introduce the quadratic functional

$$H(x) = \langle Ax, x\rangle - 2\langle y, x\rangle. \tag{2.7}$$

For a given $y \in H$, a vector $x_0 \in H$ is the solution of $L(x) = 0$ if and only if $x_0 \in H$ attains the minimum of $H(x)$.

Indeed, suppose that $x \in H$ satisfies $H(x) = \min_{u\in H} H(u)$. Then for each nonzero $z \in H$ and a real parameter $\alpha \in \mathbb{R}$, it follows that

$$H(x+\alpha z) - H(x) \geq 0$$

and this implies

$$\begin{aligned} H(x+\alpha z) - H(x) &= \langle Ax + \alpha Az, x + \alpha z\rangle - 2\langle y, x+\alpha z\rangle - H(x) \\ &= \alpha\,[\langle Ax, z\rangle + \langle Az, x\rangle] + \alpha^2\langle Az, z\rangle - 2\alpha\langle y, z\rangle \\ &= 2\alpha\langle Ax - y, z\rangle + \alpha^2\langle Az, z\rangle \geq 0. \end{aligned}$$

Since A is positive invertible, we have $\langle Az, z\rangle > 0$. Since the inequality above holds for all $\alpha \in \mathbb{R}$, we get $(Ax - y, z) = 0$ for all nonzero $z \in H$. Therefore we have $Ax - y = 0$ and hence $x \in H$ is the solution of $L(x) = 0$.

Conversely, suppose that $x \in H$ is the solution of $L(x) = Ax - y = 0$. Then

$$H(x+z) - H(x) = \langle Az, z\rangle + 2\langle Ax - y, z\rangle = \langle Az, z\rangle > 0 \tag{2.8}$$

for all nonzero $z \in H$. For each $y \in H$, if we put $z = y - x$ in (2.8), then we have $H(y) \geq H(x)$ and this implies $H(x) = \min_{y\in H} H(y)$.

In this way, if the problem of solving an equation (2.6) reduces to the problem of seeking the minimum of the functional (2.7), then this fact is named the variational principle of the equation.

In seeking the minimum of a functional (2.7) we shall employ the method of steepest descent. Now, we consider the following three procedures (0), (1) and (2):

(0) For a given initial vector $x_0 \in H$, we find a sequence $\{x_n\} \subset H$ such that

$$H(x_0) > H(x_1) > \cdots > H(x_n) > \cdots \to \min_{u\in H} H(u) = H(x).$$

(1) By induction, we construct a sequence $\{x_n\} \subset H$ such that

$$x_{n+1} = x_n + \alpha_n z_n$$

for $\alpha_n \in \mathbb{R}$ and $z_n \in H$.

(2) Moreover, we choose $\alpha_n \in \mathbb{R}$ such that

$$H(x_n + \alpha_n z_n) = \min_{t\in\mathbb{R}} H(x_n + t z_n). \tag{2.9}$$

The following lemma shows that the condition (0) implies the convergence of $\{x_n\}$.

Lemma 2.1 *Let x be the solution of $L(x) = Ax - y = 0$. If a sequence $\{x_n\}$ satisfies*

$$H(x_0) > H(x_1) > \cdots > H(x_n) > \cdots \to \min_{u \in H} H(u) = H(x),$$

then $x_n \to x$ as $n \to \infty$.

Proof.

$$\begin{aligned} H(x_n) - H(x) &= \langle Ax_n, x_n \rangle - 2\langle y, x_n \rangle - \langle Ax, x \rangle + 2\langle y, x \rangle \\ &= 2\langle Ax - y, x_n - x \rangle + \langle A(x_n - x), x_n - x \rangle \\ &= \langle A(x_n - x), x_n - x \rangle \geq m\|x_n - x\|^2, \end{aligned}$$

because $m\langle z,z\rangle \leq \langle Az,z\rangle \leq M\langle z,z\rangle$ for every $z \in H$ by the assumption. Therefore $\lim_{n\to\infty} H(x_n) = H(x)$ implies $\lim_{n\to\infty} x_n = x$. □

The following lemma determines the form of α_n.

Lemma 2.2 *If* (2.9) *holds, then*

$$\alpha_n = \frac{\langle z_n, z_n \rangle}{\langle Az_n, z_n \rangle}$$

where $z_n = y - Ax_n$.

Proof.

$$\begin{aligned} H(x_n + tz_n) &= \langle Az_n, z_n \rangle t^2 + 2(\langle Ax_n, z_n \rangle - \langle y, z_n \rangle)t + H(x_n) \\ &= \langle Az_n, z_n \rangle t^2 + 2\langle z_n, z_n \rangle t + H(x_n) \\ &= \langle Az_n, z_n \rangle \left(t - \frac{\langle z_n, z_n \rangle}{\langle Az_n, z_n \rangle} \right)^2 - \frac{\langle z_n, z_n \rangle^2}{\langle Az_n, z_n \rangle} + H(x_n) \end{aligned}$$

Therefore, $t = \frac{\langle z_n, z_n \rangle}{\langle Az_n, z_n \rangle}$ attains the minimum of $H(x_n + tz_n)$. □

By the proof of Lemma 2.2, we have

$$H(x_{n+1}) = H(x_n) - \frac{\langle z_n, z_n \rangle^2}{\langle Az_n, z_n \rangle} < H(x_n)$$

and hence we have

$$H(x_0) > H(x_1) > \cdots > H(x_n) > \cdots.$$

Theorem K4 *The successive approximations $\{x_n\} \subset H$ constructed by the method of steepest descent converge strongly to the solution of the equation* (2.6) *with the speed of a geometrical progression.*

Proof. Let x^* be the solution of equation (2.6) and $\Delta_n H = H(x_n) - H(x^*)$. It is obtained that the change $\Delta_n H$ of H in passing from x^* to x_n is

$$\Delta_n H = H(x_n) - H(x^*) = \langle A(x^* - x_n), x^* - x_n \rangle.$$

Also, since

$$z_n = y - Ax_n$$

and

$$z_{n+1} = y - Ax_{n+1} = z_n - \alpha_n A z_n,$$

it follows that

$$\Delta_n H = \langle A(x_n - x^*), x_n - x^* \rangle = \langle A^{-1} z_n, z_n \rangle$$

and

$$\Delta_{n+1} H = \langle A(x_{n+1} - x^*), x_{n+1} - x^* \rangle = \Delta_n H - 2\alpha_n \langle z_n, z_n \rangle + \alpha_n^2 \langle A z_n, z_n \rangle.$$

By the definition of α_n, we have

$$\begin{aligned} \frac{\Delta_n H - \Delta_{n+1} H}{\Delta_n H} &= \frac{2\alpha_n \langle z_n, z_n \rangle - \alpha_n^2 \langle A z_n, z_n \rangle}{\langle A^{-1} z_n, z_n \rangle} \\ &= \frac{\langle z_n, z_n \rangle^2}{\langle A z_n, z_n \rangle \langle A^{-1} z_n, z_n \rangle} \end{aligned} \tag{2.10}$$

We notice the form of a generalization of the Kantorovich inequality due to Greub-Rheinboldt in the last expression of (2.10).

For the estimation of this ratio let us make use of the spectral decomposition of an operator A:

$$A = \int_m^M \lambda de_\lambda \quad \text{and} \quad \langle A z_1, z_1 \rangle = \int_m^M \lambda \, \mathrm{d}\langle e_\lambda z_1, z_1 \rangle = \lim \sum \lambda \langle \Delta e_\lambda z_1, z_1 \rangle; \tag{2.11}$$

analogously

$$\langle z_1, z_1 \rangle = \lim \sum \langle \Delta e_\lambda z_1, z_1 \rangle \quad \text{and} \quad \langle A^{-1} z_1, z_1 \rangle = \lim \sum \frac{1}{\lambda} \langle \Delta e_\lambda z_1, z_1 \rangle. \tag{2.12}$$

Replacing in expression (2.10) the inner product by their approximate value as given by (2.11) and (2.12), we have

$$\begin{aligned} \frac{\Delta_n H - \Delta_{n+1} H}{\Delta_n H} &= \frac{[\sum \langle \Delta e_\lambda z_1, z_1 \rangle]^2}{\sum \lambda \langle \Delta e_\lambda z_1, z_1 \rangle \sum \frac{1}{\lambda} \langle \Delta e_\lambda z_1, z_1 \rangle} \\ &\geq \frac{4Mm}{(M+m)^2} > 0. \end{aligned}$$

The Kantorovich inequality is utilized here to estimate a lower bound!

The approximate equality here is correct with as small an error as one pleases, and we have therefore an exact inequality

$$\frac{\Delta_n H - \Delta_{n+1} H}{\Delta_n H} \geq \frac{4Mm}{(M+m)^2},$$

whence

$$\Delta_{n+1} H \leq \left(1 - \frac{4Mm}{(M+m)^2}\right) \Delta_n H = \left(\frac{M-m}{M+m}\right)^2 \Delta_n H.$$

Since $0 \le \frac{M-m}{M+m} < 1$, for a given initial vector x_0, we have

$$\lim_{n\to\infty} \Delta_n H = 0$$

so that $\lim_{n\to\infty} H(x_n) = H(x^*)$. By Lemma 2.1, we have $x_n \to x^*$ as $n \to \infty$ and this proves the assertion. □

The rapidity of convergence of the process is of the order of a geometric progression with ratio $q = (M-m)/(M+m)$.

It is surprising that the Kantorovich inequality is utilized in the linear problems of analysis. We cannot understand this fact by reading [132] only. Also, as mentioned above, we think that Kantorovich proved the following form: If an operator A on H is positive such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$, then

$$\frac{\langle x,x\rangle^2}{\langle Ax,x\rangle\langle A^{-1}x,x\rangle} \ge \frac{4}{\left[\sqrt{\frac{M}{m}}+\sqrt{\frac{m}{M}}\right]^2} \tag{2.13}$$

holds for every nonzero vector x in H.

Namely, the Kantorovich inequality is not only the form (2.1) shown in Lemma K, but also the form (2.13) of the operator version.

Now, the theorem denoted by K2 is a generalization of the Kantorovich inequality in the operator form, as it was derived by Greub and Rheinboldt. In fact, we easily see that (2.13) implies Theorem K2. Therefore, one could say that Kantorovich proved Theorem K2 in a certain sense. At this point, it is suitable to cite a relevant part of [132]:

> The subject of this paper is the proof of a generalized form of the inequality for linear, bounded and self-adjoint operators in Hilbert space. This generalized Kantorovich inequality proves to be equivalent to a similarly generalized form of the inequality which we shall call the generalized Pólya-Szegö inequality. Our generalized Kantorovich inequality is already implicitly contained in the paper of L.V.Kantorovich. However, its proof there involves the use of the theory of spectral decomposition for the operators in question. The proof we shall present here will proceed in a considerable simpler way.

Hence, from the underlined sentence we learn that the proof of Theorem K2 was essentially contained in [156]. Furthermore, we see that Greub and Rheinboldt prefer to avoid the spectral decomposition theorem in the proof, as they believe their own proof to be considerably simpler.

However, it turned out that their method of proof had a deep significance for mathematics. The impact of Theorem K2 could be compared to spreading of shock waves around the world of mathematics. Thus we present the proof of which Greub and Rheinboldt say that is simpler.

Proof of Theorem K2. The left hand side of the inequality follows directly from Schwarz's inequality

$$\begin{aligned}\langle x,x\rangle^2 &= \langle A^{1/2}x, A^{-1/2}x\rangle^2 \le \langle A^{1/2}x, A^{1/2}x\rangle\langle A^{-1/2}x, A^{-1/2}x\rangle \\ &= \langle Ax,x\rangle\langle A^{-1}x,x\rangle.\end{aligned}$$

We shall first prove the right hand side of (2.4) for finite dimensional space H. Then we will show that the proof for the general case can be reduced to that of the finite dimensional case.

Suppose that H is a finite dimensional space. Then the unit sphere $S \subset H$ is compact. Hence, considered on S, the continuous functional

$$f(x) = \frac{\langle Ax,x\rangle\langle A^{-1}x,x\rangle}{\langle x,x\rangle^2}$$

attains its maximum at a certain point, say $x_0 \in S$, i.e.

$$f(x_0) = \max_{x\in S} f(x) = \langle Ax_0,x_0\rangle\langle A^{-1}x_0,x_0\rangle.$$

With a fixed vector $y \in H$ and the real parameter t ($|t| < 1$) we consider the real valued function

$$g(t) = f(x_0 + ty).$$

This function $g(t)$ has a relative maximum at $t = 0$ and therefore we must necessarily have $g'(0) = 0$. Using the self-adjointness of A and A^{-1} we find

$$g'(0) = 2\langle Ax_0,y\rangle\langle A^{-1}x_0,x_0\rangle + 2\langle A^{-1}x_0,y\rangle\langle Ax_0,x_0\rangle - 4f(x_0)\langle x_0,y\rangle = 0$$

and thus

$$\langle \gamma Ax_0 + \mu A^{-1}x_0 - x_0, y\rangle = 0$$

holds for all $y \in H$, where

$$\gamma = \frac{1}{2\langle Ax_0,x_0\rangle} \quad \text{and} \quad \mu = \frac{1}{2\langle A^{-1}x_0,x_0\rangle}.$$

Consequently

$$x_0 = \gamma Ax_0 + \mu A^{-1}x_0.$$

Applying A and A^{-1} successively to this equation we find that

$$Ax_0 = \gamma A^2x_0 + \mu x_0 \quad \text{and} \quad A^{-1}x_0 = \gamma x_0 + \mu A^{-2}x_0$$

or

$$\left(A - \frac{1}{2\gamma}I_H\right)^2 x_0 = \frac{1-4\gamma\mu}{4\gamma^2}x_0 \text{ and } \left(A^{-1} - \frac{1}{2\mu}I_H\right)^2 x_0 = \frac{1-4\gamma\mu}{4\mu^2}x_0.$$

Taking into account the assumption $0 < mI_H \le A \le MI_H$, we have

$$4\gamma\mu\frac{m}{M} \le \left(1 + (1-4\gamma\mu)^{1/2}\right)^2 \le 4\gamma\mu\frac{M}{m}.$$

It follows

$$\left[4\gamma\mu\left(\frac{m}{M}+1\right)-2\right]^2 \le 4(1-4\gamma\mu) \le \left[4\gamma\mu\left(\frac{M}{m}+1\right)-2\right]^2$$

or

$$\frac{\gamma\mu}{M^2}\left[4\gamma\mu(M+m)^2-4mM\right]\le 0\le\frac{\gamma\mu}{m^2}\left[4\gamma\mu(M+m)^2-4mM\right]$$

and therefore

$$4\gamma\mu(M+m)^2-4mM=0.$$

On the other hand, since

$$4\gamma\mu=\frac{1}{\langle Ax_0,x_0\rangle\langle A^{-1}x_0,x_0\rangle},$$

we finally have

$$\langle Ax_0,x_0\rangle\langle A^{-1}x_0,x_0\rangle=\frac{(M+m)^2}{4Mm}, \tag{2.14}$$

which was to be proved. (2.14) shows furthermore that (at least in the finite dimensional case) the upper bound in (2.4) can not be improved.

We now remove the restriction of the finite-dimensionality of H. Let x_0 be a fixed vector of H and let $H_0\subset H$ be a finite dimensional subspace of H which contains three vectors x_0, Ax_0 and $A^{-1}x_0$. We denote by P the projection of H onto H_0. For the operator $B=PA$, we have $B(H_0)\subset H_0$ and

$$\langle Bx,y\rangle=\langle PAx,y\rangle=\langle PAPx,y\rangle=\langle x,PAPy\rangle=\langle x,By\rangle$$

for all $x,y\in H_0$. Hence, B is a self-adjoint operator on the space H_0. Furthermore, we find for $x\in H_0$

$$\langle Bx,x\rangle=\langle PAx,x\rangle=\langle Ax,Px\rangle=\langle Ax,x\rangle$$

and therefore in H_0

$$0<mI_{H_0}\le m'I_{H_0}\le B\le M'I_{H_0}\le MI_{H_0} \tag{2.15}$$

where

$$m'=\inf_{x\in H_0}\frac{\langle Bx,x\rangle}{\langle x,x\rangle}\qquad\text{and}\qquad M'=\sup_{x\in H_0}\frac{\langle Bx,x\rangle}{\langle x,x\rangle}.$$

Hence, we can apply the first part of the proof to the operator B in the finite dimensional space H_0. By doing that we obtain for all $x\in H_0$

$$\frac{\langle Bx,x\rangle\langle B^{-1}x,x\rangle}{\langle x,x\rangle^2}\le\frac{(M'+m')^2}{4m'M'}=\frac{1}{4}\left(\frac{M'}{m'}+\frac{m'}{M'}\right)+\frac{1}{2}. \tag{2.16}$$

From (2.15) we conclude that

$$1\le\frac{M'}{m'}\le\frac{M}{m}\qquad\text{and}\qquad\frac{M'}{m'}+\frac{m'}{M'}\le\frac{M}{m}+\frac{m}{M}. \tag{2.17}$$

This last inequality is a result of the fact that for $u\ge 1$ the function $f(u)=u+1/u$ is monotonically increasing. (2.16) and (2.17) together yield

$$\frac{\langle Bx,x\rangle\langle B^{-1}x,x\rangle}{\langle x,x\rangle^2}\le\frac{1}{4}\left(\frac{M}{m}+\frac{m}{M}\right)+\frac{1}{2}=\frac{(M+m)^2}{4mM}$$

for all $x \in H_0$. Since H_0 contains x_0, Ax_0 and $A^{-1}x_0$, we find

$$Bx_0 = PAx_0 = Ax_0 \qquad \text{and} \qquad x_0 = Px_0 = PAA^{-1}x_0 = BA^{-1}x_0.$$

The last relation implies $B^{-1}x_0 = A^{-1}x_0$ when one considers that the existence of B^{-1} in H_0 is a direct consequence of (2.15). Substituting we obtain finally

$$\langle Ax_0, x_0\rangle\langle A^{-1}x_0, x_0\rangle \le \frac{(M+m)^2}{4mM}\langle x_0, x_0\rangle^2$$

Since x_0 was arbitrary the theorem is hereby completely proved. □

Moreover, they showed the generalized Pólya-Szegö inequality, which is equivalent to the Kantorovich inequality:

Theorem 2.1 *Let A and B be commuting self-adjoint operators on a Hilbert space H such that*

$$0 < m_1 I_H \le A \le M_1 I_H \qquad \text{and} \qquad 0 < m_2 I_H \le B \le M_2 I_H.$$

Then

$$\langle Ax, Ax\rangle\langle Bx, Bx\rangle \le \frac{(M_1M_2 + m_1m_2)^2}{4m_1m_2M_1M_2}\langle Ax, Bx\rangle^2$$

for all $x \in H$.

Proof. It is rather obvious that the Kantorovich inequality is contained in Theorem 2.1. In fact, let C be any given self-adjoint operator with

$$0 < mI_H \le C \le MI_H.$$

We set $A = C^{1/2}$ and $B = (C^{-1})^{1/2}$. Since

$$0 < m^{1/2} I_H \le A \le M^{1/2} I_H \qquad \text{and} \qquad 0 < (M^{-1})^{1/2} I_H \le B \le (m^{-1})^{1/2} I_H,$$

it follows immediately from Theorem 2.1 that

$$\frac{\langle Cx, x\rangle\langle C^{-1}x, x\rangle}{\langle x, x\rangle^2} = \frac{\langle Ax, Ax\rangle\langle Bx, Bx\rangle}{\langle Ax, Bx\rangle^2} \le \frac{(M+m)^2}{4mm}$$

for all $x \in H$ and this is the statement of the Kantorovich inequality.

Next, we show that Theorem 2.1 is a consequence of Theorem K2.

From the commutativity of A and B, for the self-adjoint operator $C = AB^{-1}$ we have

$$0 < \frac{m_1}{M_2} I_H \le C \le \frac{M_1}{m_2} I_H.$$

Therefore, it follows from Theorem K2 that

$$\frac{\langle Cx, x\rangle\langle C^{-1}x, x\rangle}{\langle x, x\rangle^2} \le \frac{(M_1M_2 + m_1m_2)^2}{4m_1m_2M_1M_2}$$

for all $x \in H$. Put $x = (AB)^{1/2}y$, then we obtain $\langle Cx,x\rangle = \langle Ay,Ay\rangle, \langle C^{-1}x,x\rangle = \langle By,By\rangle$ and $\langle x,x\rangle = \langle Ay,By\rangle$. Substituting these relations, we get the statement of Theorem 2.1. □

The proof by Greub and Rheinboldt is very long, spanning over approximately five pages. We can feel the strictness of their proof, but, in contrast, Kantorovich's proof is simple and only half a page long. However, it was the formulation by Greub and Rheinboldt that brought the first wave of excitement into the world of mathematics. Owing to Greub and Rheinboldt, the work of Kantorovich has become an object of research in mathematics, in operator theory in particular. In their own words, their proof is simple. But, it is a proof on a grand scale, unexpected and fascinating. Based on a beautiful relation, this simple formulation may strike a chord in the heart of a mathematician. Many mathematicians concentrated their energies on the generalization of the Kantorovich inequality and on searching for an even simpler proof.

2.2 Generalizations and improvements

In 1960, one year after the publication of [132], Strang [272] shows the following generalization of the Kantorovich inequality for an arbitrary operator without conditions such as self-adjoiness and positivity.

Theorem 2.2 *If T is an arbitrary invertible operator on H, and $\|T\| = M$, $\|T^{-1}\|^{-1} = m$, then*

$$|\langle Tx,y\rangle\langle x,T^{-1}y\rangle| \le \frac{(M+m)^2}{4Mm}\langle x,x\rangle\langle y,y\rangle \qquad \text{for all } x,y \in H.$$

Furthermore, this bound is the best possible.

Proof. We consider the polar decomposition of T. Let $A = (T^*T)^{1/2}$. Then $U = TA^{-1}$ is unitary, and

$$\begin{aligned} |\langle Tx,y\rangle\langle x,T^{-1}y\rangle| &= |\langle UAx,x\rangle\langle x,A^{-1}U^{-1}y\rangle| = |\langle Ax,U^*y\rangle\langle A^{-1}x,U^*y\rangle| \\ &\le \left[\langle Ax,x\rangle\langle AU^*y,U^*y\rangle\langle A^{-1}x,x\rangle\langle A^{-1}U^*y,U^*y\rangle\right]^{1/2} \end{aligned} \tag{2.18}$$

by generalized Schwarz's inequality (Theorem 1.10). Since $\|A\| = \|(T^*T)^{1/2}\| = \|T\| = M$ and $\|A^{-1}\|^{-1} = \|T^{-1}\|^{-1} = m$, it follows that $mI_H \le A \le MI_H$. Therefore, by (2.4) in Theorem K2, we have

$$\begin{aligned} \text{RHS in (2.18)} &\le \left(\frac{(M+m)^2}{4Mm}\langle x,x\rangle^2 \cdot \frac{(M+m)^2}{4Mm}\langle U^*y,U^*y\rangle^2\right)^{1/2} \\ &= \frac{(M+m)^2}{4Mm}\langle x,x\rangle\langle y,y\rangle, \end{aligned}$$

by using $\langle U^*y,U^*y\rangle = \langle y,y\rangle$.

If H is finite dimensional, the bound is attained for $x = U^*y = u + v$, where u and v are unit eigenvectors of A corresponding to eigenvalues m and M. In a general case, the bound need not be attained. But a sequence $x_n = U^*y_n = u_n + v_n$, where $\|u_n\| = \|v_n\|$, $(e(m+1/n) - e(m-0))u_n = u_n$, $(e(M+0) - e(M-1/n))v_n = v_n$ shows on calculation that the bound is best possible. □

Also, Schopf [257] considered a generalization of the power in the Kantorovich inequality. Moving to the year 1996, there is the following extension due to Spain [270] which is totally different from the Kantorovich inequality. But it is surely an extension. It does not assume positivity, either. It is slightly long, but we will quote it:

The Kantorovich inequality says that if A is a positive operator on a Hilbert space H such that $mI_H \le A \le MI_H$ for some scalars $0 < m \le M$, then

$$4mM\langle A^{-1}x, x\rangle \le (m+M)^2 \frac{\|x\|^4}{\langle Ax, x\rangle}$$

holds for every vector x in H. If we replace x by $A^{\frac{1}{2}}x$, then

$$4mM\langle x, x\rangle \le (m+M)^2 \frac{\|A^{\frac{1}{2}}x\|^2}{\langle A^2x, x\rangle}.$$

This inequality may be viewed as a conversion of the special case

$$\langle Ax, x\rangle \le \|Ax\|\|x\|$$

of the Cauchy-Schwarz inequality, for it is equivalent to the inequality

$$2\sqrt{mM}\|Ax\|\|x\| \le (m+M)\langle Ax, x\rangle.$$

The methods of operator and spectral theory allow one to generalize the inequality to a wide class of operators on a Hilbert space.

Let Γ be any nonzero complex number, let $R = |\Gamma|$, and let $0 \le r \le R$.

Theorem 2.3 *Let A be an operator on H such that $\left|A - \Gamma[A]\right|^2 \le r^2[A]$, where $[A]$ is the range projection of A. Let $u \in B(K,H)$ be an operator such that $u^*[A]u$ is a projection. Then*

$$(R^2 - r^2)u^*A^*Au \le R^2(u^*A^*u)(u^*Au).$$

Proof. Since $u^*[A]u$ is a projection, we have

$$\begin{aligned}&\left|(R^2 - r^2)u^*[A]u - \overline{\Gamma}u^*Au\right|^2\\ &= (R^2 - r^2)^2u^*[A]u - (R^2 - r^2)\{\overline{\Gamma}u^*Au + \Gamma u^*A^*u\} + R^2(u^*A^*u)(u^*Au),\end{aligned}$$

while

$$\begin{aligned}&u^*\left(r^2[A] - \left|A - \Gamma[A]\right|^2\right)u\\ &= -(R^2 - r^2)u^*[A]u - u^*A^*Au + \overline{\Gamma}u^*Au + \Gamma u^*A^*u,\end{aligned}$$

and hence

$$\begin{aligned}&R^2u^*A^*uu^*Au-(R^2-r^2)u^*A^*Au\\&=\left|(R^2-r^2)u^*[A]u-\overline{\Gamma}u^*Au\right|^2+u^*\left(r^2[A]-\left|A-\Gamma[A]\right|^2\right)u.\end{aligned}$$

By the assumption of $|A-\Gamma[A]|^2\le r^2[A]$, we have

$$R^2(u^*A^*u)(u^*Au)-(R^2-r^2)u^*A^*Au\ge 0.$$

□

Corollary 2.1 *Let A be a positive operator on H such that A is invertible on its range, let $m=\min \mathsf{Sp}(A)\backslash\{0\}$ and $M=\max \mathsf{Sp}(A)=\|A\|$. Let $u\in B(K,H)$ be an operator such that $u^*[A]u$ is a projection. Then*

$$4Mmu^*A^2u\le (M+m)^2(u^*Au)^2.$$

Proof. In the situation of Theorem 2.3, we have

$$R=\Gamma=\frac{M+m}{2}\qquad\text{and}\qquad r=\frac{M-m}{2}.$$

By the assumption of A, it follows that

$$m[A]\le A\le M[A]$$

and hence $|A-\Gamma[A]|^2\le r^2[A]$. Therefore Corollary 2.1 follow from Theorem 2.3. □

Theorem 2.4 *Let A be an operator on H such that $|A-\Gamma[A]|^2\le r^2[A]$. Then*

$$(R^2-r^2)^{1/2}\|Ax\|\left\|[A]x\right\|\le R|\langle Ax,x\rangle|,\qquad x\in H.$$

If A is positive with $\mathsf{Sp}(A)\backslash\{0\}\subset[m,M]$ $(0<m<M)$, then

$$2\sqrt{Mm}\|Ax\|\left\|[A]x\right\|\le (m+M)\langle Ax,x\rangle\qquad \textit{for all } x\in H.$$

Proof. For $x\in H$ define $u_x:\mathbb{C}\mapsto H:\lambda\mapsto\lambda x$. Then, identifying $\mathbb{C}$ and $B(\mathbb{C})$ canonically,

$$u_x^*Au_x=\langle Ax,x\rangle\qquad\text{for } A\in B(H).$$

There is nothing to prove if $[A]x=0$, otherwise put $u=u_{x/\|[A]x\|}$. The first assertion follows from Corollary 2.1. The second assertion is a direct consequence of the the first. □

Remark 2.1 *The second assertion in Theorem 2.4 may be proved in one line:*

$$\begin{aligned}&(m+M)^2\langle Ax,x\rangle^2-4Mm\|Ax\|^2\|[A]x\|^2\\&=\left\{2mM\|[A]x\|^2-(m+M)\langle Ax,x\rangle\right\}^2\\&\qquad+4Mm\left\langle (M-A)(A-m)[A]x,[A]x\right\rangle\left\|[A]x\right\|^2\ge 0\end{aligned}$$

Generalizations of the Kantorovich inequality have made significant progress. The Mathematical Society was given a treat in the form of topics for the Kantorovich inequality for a while.

On the other hand, in pursuit of an even simpler proof, in such a flood of papers, Nakamura [237] instantly presents the following result in Proceedings of the Japan Academy. It was in 1960, just one year after the paper due to Greub and Rheinboldt was published. It is a simple visual proof by using the concavity of $f(t) = t^{-1}$.

Theorem 2.5 *For $0 < m < M$, the following inequality holds true:*

$$\int_m^M t\,d\mu(t) \cdot \int_m^M \frac{1}{t}\,d\mu(t) \leq \frac{(M+m)^2}{4Mm} \tag{2.19}$$

for any positive Stieltjes measure μ on $[m,M]$ with $\|\mu\| = 1$.

It is easy to see, by the Gelfand representation of the C*-algebra generated by A and the identity operator I, that Theorem 2.5 implies the Kantorovich inequality.

If Nakamura had the opportunity to read [156] in an English translation and if he asked the mathematical community for judgment on the inequality (2.19) and its overwhelmingly simple proof, then how would that turn out? In one possible outcome, mathematicians would mostly get the impression that it was very easy to prove that result and therefore the investigations related to the Kantorovich inequality would be brought to the end. For some reason, Nakamura's paper is overlooked in the mathematical world.
To the best of this author's knowledge, there is no evidence that anyone has ever cited Nakamura's paper. Instead, several improvements to proofs of the Kantorovich inequality have been independently developed in Europe.

The origin of the Kantorovich inequality might be the following case of finite sequences.

Theorem 2.6 *If the sequence $\{\gamma_i\}$ satisfies the conditions such that $m \leq \gamma_i \leq M$ for some scalars $0 < m \leq M$ and $i = 1,2,\cdots,n$, then*

$$(\xi_1\gamma_1 + \cdots + \xi_n\gamma_n)(\xi_1\gamma_1^{-1} + \cdots + \xi_n\gamma_n^{-1}) \leq \frac{(M+m)^2}{4Mm} \tag{2.20}$$

holds for every $\xi_i \geq 0$ such that $\xi_1 + \cdots + \xi_n = 1$.

First of all, we present a direct proof due to Henrici [141]:

Proof of Theorem 2.6. We may assume that $m < M$. Determine p_i and q_i from the equations

$$\gamma_i = p_iM + q_im \quad \text{and} \quad \gamma_i^{-1} = p_iM^{-1} + q_im^{-1} \quad \text{for} \quad i = 1,\cdots,n.$$

An easy computation shows that $p_i, q_i \geq 0$, $i = 1,2,\cdots,n$. Furthermore from

$$1 = (p_iM + q_im)(p_iM^{-1} + q_im^{-1}) = (p_i + q_i)^2 + p_iq_i\frac{(M-m)^2}{mM}$$

it follows that $p_i+q_i \le 1$. Setting $p=\sum \xi_i p_i$, $q=\sum \xi_i q_i$, we thus have $p+q=\sum \xi_i(p_i+q_i) \le \sum \xi_i = 1$. Hence using the arithmetic-geometric mean inequality,

$$
\begin{aligned}
&(\xi_1\gamma_1+\cdots+\xi_n\gamma_n)(\xi_1\gamma_1^{-1}+\cdots+\xi_n\gamma_n^{-1})\\
&=(pM+qm)(pM^{-1}+qm^{-1})=(p+q)^2+pq\frac{(M-m)^2}{Mm}\\
&\le (p+q)^2\left[1+\frac{(M-m)^2}{4Mm}\right]=(p+q)^2\frac{(M+m)^2}{4Mm}\le\frac{(M+m)^2}{4Mm}.
\end{aligned}
$$

Equality is attained in (2.20) if and only if the following two conditions are simultaneously fulfilled (we assume here $\xi_i>0, i=1,2,\cdots,n$ without loss of generalization):

(i) $p+q=1$. This implies that $p_i+q_i=1$ or $p_iq_i=0$ for $i=1,\cdots,n$. Thus, for equality every γ_i must equal either M or m.

(ii) $p+q=4pq$. This implies that $p=q$ or, $\sum_{\gamma_i=m}\xi_i=\sum_{\gamma_i=M}\xi_i$.

Thus, the weights attached to m and M must be the same. □

In comparison with Kantorovich's proof, Henrici's one relies on an algebraic calculation. Inspired by Henrici, Rennie [255] gives the following improved proof with functions in 1963:

Let f be a measurable function on the probability space such that $0<m\le f(x)\le M$. Integrating the inequality

$$\frac{(f(x)-m)(f(x)-M)}{f(x)}\le 0$$

gives

$$\int f(x)\mathrm{d}x+mM\int\frac{1}{f(x)}\mathrm{d}x\le m+M.$$

Put $u=mM\int\frac{1}{f(x)}\mathrm{d}x$, then we have

$$u\int f(x)\mathrm{d}x\le(m+M)u-u^2=-\left(u-\frac{M+m}{2}\right)^2+\frac{(M+m)^2}{4}\le\frac{(M+m)^2}{4},$$

which is the Kantorovich inequality:

$$\int\frac{1}{f(x)}\mathrm{d}x\int f(x)\mathrm{d}x\le\frac{(M+m)^2}{4mM}.$$

□

This is exactly a function version of the Kantorovich inequality due to Nakamura. Its emphatic brevity is surprising. Moreover, inspired by Rennie, Mond [209] gives the following improved proof with matrices in 1965:

Let A be a positive definite Hermitian matrix with eigenvalues $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_n>0$. Since three factors in the LHS of below inequality commute, we have

$$(A-\lambda_nI)(A-\lambda_1I)A^{-1}\le 0.$$

Therefore,

$$\langle Ax,x\rangle + \lambda_1\lambda_n\langle A^{-1}x,x\rangle \le \lambda_1+\lambda_n$$

for every unit vector x. If we put $u=\lambda_1\lambda_n\langle A^{-1}x,x\rangle$, then

$$\lambda_1\lambda_n\langle A^{-1}x,x\rangle\langle Ax,x\rangle = u\langle Ax,x\rangle \le (\lambda_1+\lambda_n)u-u^2 \le \frac{(\lambda_1+\lambda_n)^2}{4},$$

which implies the Kantorovich inequality:

$$\langle A^{-1}x,x\rangle\langle Ax,x\rangle \le \frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n}.$$

The proof of Mond may be considered one of the generalized Kantorovich inequality. But, we present a somewhere different proof by using the arithmetic-geometric mean inequality in [164, 144, 158]:

Since A is positive and $0 < mI_H \le A \le MI_H$, it follows that $MI_H - A \ge 0$ and $A-mI_H \ge 0$. The commutativity of $MI_H - A$ and $A-mI_H$ implies $(MI_H-A)(m^{-1}I_H-A^{-1}) \ge 0$. Hence

$$(M+m)I_H \ge MmA^{-1}+A$$

and

$$\langle (M+m)x,x\rangle \ge Mm\langle A^{-1}x,x\rangle + \langle Ax,x\rangle$$

holds for every unit vector $x\in H$. By using the arithmetic-geometric mean inequality

$$M+m = \langle (M+m)x,x\rangle \ge Mm\langle A^{-1}x,x\rangle + \langle Ax,x\rangle \ge 2\sqrt{Mm\langle A^{-1}x,x\rangle\langle Ax,x\rangle}.$$

Squaring both sides, we obtain the desired inequality

$$\langle Ax,x\rangle\langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm}.$$

Finally, we present an extremely simple idea due to Diaz and Metcalf [43]:

Lemma 2.3 *Let real numbers $a_k \ne 0$ and b_k $(k=1,2,\cdots,n)$ satisfy*

$$m \le \frac{b_k}{a_k} \le M. \tag{2.21}$$

Then

$$\sum_{k=1}^n b_k^2 + mM\sum_{k=1}^n a_k^2 \le (M+m)\sum_{k=1}^n a_kb_k.$$

The equality holds if and only if in each of the n inequalities (2.21)*, at least one of the equality signs holds, i.e. either $b_k=ma_k$ or $b_k=Ma_k$ (where the equation may vary with k).*

Proof. It follows from the hypothesis (2.21) that

$$0 \le \left(\frac{b_k}{a_k} - m\right)\left(M - \frac{b_k}{a_k}\right) a_k^2.$$

Thus, summing from $k = 1$ to $k = n$,

$$\begin{aligned} 0 &\le \sum_{k=1}^{n} (b_k - ma_k)(Ma_k - b_k) \qquad (2.22) \\ &= (M+m)\sum_{k=1}^{n} a_k b_k - \sum_{k=1}^{n} b_k^2 - mM \sum_{k=1}^{n} a_k^2, \end{aligned}$$

which gives the desired result. Clearly, the equality holds in (2.22) if and only if each term of the summation is zero. □

By using Lemma 2.3, we have

$$\begin{aligned} 0 &\le \left(\left(\sum_{k=1}^{n} b_k^2\right)^{1/2} - \left(mM\sum_{k=1}^{n} a_k^2\right)^{1/2}\right)^2 \\ &= \sum_{k=1}^{n} b_k^2 - 2\left(\sum_{k=1}^{n} b_k^2\right)^{1/2}\left(mM\sum_{k=1}^{n} a_k^2\right)^{1/2} + mM\sum_{k=1}^{n} a_k^2 \\ &\le (m+M)\sum_{k=1}^{n} a_k b_k - 2\sqrt{mM}\left(\sum_{k=1}^{n} b_k^2\right)^{1/2}\left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \end{aligned}$$

and hence

$$4mM\left(\sum_{k=1}^{n} b_k^2\right)\left(\sum_{k=1}^{n} a_k^2\right) \le (m+M)^2\left(\sum_{k=1}^{n} a_k b_k\right)^2$$

yields immediately the result of Pólya and Szegö (Theorem PS (2.2)).

Similarly, we have an operator version of Lemma 2.3:

Theorem 2.7 *Let A and B be self-adjoint operators such that $AB = BA$ and A^{-1} exists, and*

$$mI_H \le BA^{-1} \le MI_H \qquad \text{for some scalars } 0 < m \le M.$$

Then

$$B^2 + mMA^2 \le (m+M)AB. \qquad (2.23)$$

The equality holds in (2.23) *if and only if* $(MI_H - BA^{-1})(BA^{-1} - mI_H) = 0$.

By using Theorem 2.7, we have

$$\begin{aligned} 0 &\le \left\{\langle Bx, Bx\rangle^{1/2} - mM\langle Ax, Ax\rangle^{1/2}\right\}^2 \\ &= \langle Bx, Bx\rangle - 2\sqrt{mM}\langle Bx, Bx\rangle^{1/2}\langle Ax, Ax\rangle^{1/2} + mM\langle Ax, Ax\rangle \\ &\le (m+M)\langle ABx, x\rangle - 2\sqrt{mM}\langle Bx, Bx\rangle^{1/2}\langle Ax, Ax\rangle^{1/2} \end{aligned}$$

and hence

$$4mM\langle Bx,Bx\rangle\langle Ax,Ax\rangle \le (m+M)^2\langle ABx,x\rangle^2$$

yields immediately results of Greub and Rheinboldt (Theorem 2.1).

Comparing with the proofs of Kantorovich and Greub and Rheinboldt, only algebraic calculation seems to belong to a different age. However, when we can prove it plainly and simply, devising a new proof stops being an object of interest for mathematicians.

2.3 The Mond-Pečarić method

In this section, we present the principle of the Mond-Pečarić method for convex functions.

Mond and Pečarić rephrased the Kantorovich inequality as follows: The Kantorovich inequality says that if A is a positive operator such that $0 < mI_H \le A \le MI_H$, then

$$\langle Ax,x\rangle\langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm} \tag{2.24}$$

for every unit vector $x \in H$. Divideing both sides by $\langle Ax,x\rangle$, we get

$$\langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm}\langle Ax,x\rangle^{-1}. \tag{2.25}$$

Also, since $1 \le \langle Ax,x\rangle\langle A^{-1}x,x\rangle$, we may extend (2.25) into the following inequality:

$$\langle Ax,x\rangle^{-1} \le \langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm}\langle Ax,x\rangle^{-1}. \tag{2.26}$$

The first inequality of (2.26) is a special case of Jensen's inequality. In fact, if we put $f(t) = t^{-1}$, then

$$\left(\frac{a_1+\cdots+a_n}{n}\right)^{-1} \le \frac{a_1^{-1}+\cdots+a_n^{-1}}{n}$$

for all positive real numbers $a_1,\cdots,a_n$. Moreover, if $f(t)$ is a convex function on an interval $[m,M]$, then

$$f\left(\sum_{i=1}^n t_i x_i\right) \le \sum_{i=1}^n t_i f(x_i)$$

for every $x_1,\cdots,x_n \in [m,M]$ and every positive real number $t_1,\cdots,t_n$ with $\sum_{i=1}^n t_i = 1$. This inequality is called the classical Jensen's inequality. Moreover, an operator version of the classical Jensen's inequality holds:

Theorem 2.8 *Let A be a self-adjoint operator on H such that $mI_H \le A \le MI_H$ for some scalars $m \le M$ and f a real valued continuous convex function on $[m,M]$. Then*

$$f(\langle Ax,x\rangle) \le \langle f(A)x,x\rangle$$

holds for every unit vector $x \in H$.

Proof. Refer to [124, Theorem 1.2] for the proof. □

From this point of view, $\langle Ax,x\rangle^{-1} \le \langle A^{-1}x,x\rangle$ is considered as one form of Jensen's inequality. Namely, Mond and Pečarić noticed that

the Kantorovich inequality is the converse inequality of the so called Jensen's one for the function $f(t) = 1/t$.

Jensen's inequality is one of the most important inequalities in the functional analysis. Many generalizations are developed and many significant results are obtained by using Jensen's inequalitiy.

Here, let us consider a generalization of the Kantorovich inequality. Jensen's inequality for $f(t) = t^3$ yields

$$\langle Ax,x\rangle^3 \le \langle A^3x,x\rangle \qquad \text{for every unit vector } x \in H. \tag{2.27}$$

What is a converse of (2.27)? Unfortunately, it seems to be difficult to apply the same method as in the proof of the Kantorovich inequality. We need a new way of thinking. We recall Nakamura's article [237]. It was published too early, as it was ahead of its time and later on hardly anyone looked back at that paper. Thirty years later ideas similar to his had appeared in Eastern Europe. By then Nakamura had forgotten all about his principle, but it had taken root in Eastern Europe and would grow in time.

Thus, we shall recall the proof due to Nakamura: Let μ be a normalized positive Stieltjes measure on $[m,M]$. Let $y = g(t)$ a straight line joining the points $(m,1/m)$ and $(M,1/M)$. Since $1/t \le g(t)$, we have

$$\int_m^M \frac{1}{t} d\mu(t) \le \int_m^M g(t) d\mu(t) = \frac{M^{-1}+m^{-1}}{2}.$$

Multiply $\int_m^M t d\mu(t) = \frac{M+m}{2}$ to both sides,

$$\int_m^M t d\mu(t) \int_m^M \frac{1}{t} d\mu(t) \le \frac{M+m}{2} \cdot \frac{M^{-1}+m^{-1}}{2} = \frac{(M+m)^2}{4Mm}.$$

Applying it to a positive operator A with $\|A\| = M$ and $\|A^{-1}\|^{-1} = m$, we have just the Kantorovich inequality

$$\langle Ax,x\rangle\langle A^{-1}x,x\rangle \le \frac{(M+m)^2}{4Mm}$$

for every unit vector $x \in H$. We remark that the Kantorovich constant equals the arithmetic mean of m and M divided by the harmonic one:

$$\frac{(M+m)^2}{4Mm} = \frac{\frac{M+m}{2}}{\left(\frac{M^{-1}+m^{-1}}{2}\right)^{-1}}.$$

Namely, we know that Nakamura's proof is actually the origin of the so called the Mond-Pečarić method for convex functions by which the converses of Jensen's inequality are induced. Moreover, Ky Fan [48] proceeded with a generalization of the Kantorovich inequality for $f(t) = t^p$ with $p \in \mathbb{Z}$. Here, we shall present the principle of the Mond-Pečarić method for convex functions:

Theorem 2.9 *Let A be a self-adjoint operator on a Hilbert space H such that $mI_H \leq A \leq MI_H$ for some scalars $m < M$. If f is a convex function on $[m,M]$ such that $f > 0$ on $[m,M]$, then*

$$\langle f(A)x,x\rangle \leq K(m,M,f)f(\langle Ax,x\rangle)$$

for every unit vector $x \in H$, where

$$K(m,M,f) = \max\left\{\frac{1}{f(t)}\left(\frac{f(M)-f(m)}{M-m}(t-m)+f(m)\right) : m \leq t \leq M\right\}.$$

Proof. Since $f(t)$ is convex on $[m,M]$, we have

$$f(t) \leq \frac{f(M)-f(m)}{M-m}(t-m)+f(m) \qquad \text{for all } t \in [m,M].$$

Using the operator calculus, it follows that

$$f(A) \leq \frac{f(M)-f(m)}{M-m}(A-m)+f(m)I_H$$

and hence

$$\langle f(A)x,x\rangle \leq \frac{f(M)-f(m)}{M-m}(\langle Ax,x\rangle - m)+f(m)$$

for every unit vector $x \in H$. Divide both sides by $f(\langle Ax,x\rangle)$ (>0), and we get

$$\begin{aligned}\frac{\langle f(A)x,x\rangle}{f(\langle Ax,x\rangle)} &\leq \frac{\frac{f(M)-f(m)}{M-m}(\langle Ax,x\rangle - m)+f(m)}{f(\langle Ax,x\rangle)} \\ &\leq \max\left\{\frac{1}{f(t)}\left(\frac{f(M)-f(m)}{M-m}(t-m)+f(m)\right) : m \leq t \leq M\right\},\end{aligned}$$

since $m \leq \langle Ax,x\rangle \leq M$. Therefore, we have the desired inequality. □

Theorem 2.10 *Let A be a self-adjoint operator on a Hilbert space H such that $mI_H \leq A \leq MI_H$ for some scalars $m < M$. If f is a concave function on $[m,M]$ such that $f > 0$ on $[m,M]$, then*

$$\tilde{K}(m,M,f)f(\langle Ax,x\rangle) \leq \langle f(A)x,x\rangle \leq f(\langle Ax,x\rangle)$$

for every unit vector $x \in H$, where

$$\tilde{K}(m,M,f) = \min\left\{\frac{1}{f(t)}\left(\frac{f(M)-f(m)}{M-m}(t-m)+f(m)\right) : m \le t \le M\right\}.$$

In particular, if we put $f(t) = t^p$ for $p \in \mathbb{R}$ in Theorem 2.9 and 2.10, then we have the Hölder-McCarthy inequality and its converse:

Theorem 2.11 *Let A be a positive operator on a Hilbert space H such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$. Then*

$$\langle Ax,x\rangle^p \le \langle A^p x,x\rangle \le K(m,M,p)\langle Ax,x\rangle^p \quad \text{for } p \notin [0,1] \tag{2.28}$$

and

$$K(m,M,p)\langle Ax,x\rangle^p \le \langle A^p x,x\rangle \le \langle Ax,x\rangle^p \quad \text{for } p \in [0,1]$$

for every unit vector $x \in H$, where

$$K(m,M,p) = \frac{mM^p - Mm^p}{(p-1)(M-m)}\left(\frac{p-1}{p}\frac{M^p-m^p}{mM^p-Mm^p}\right)^p \tag{2.29}$$

for each $p \in \mathbb{R}$. The constant $K(m,M,p)$ is sharp in the sense that there exists a unit vector $z \in H$ such that

$$\langle A^p z,z\rangle = K(m,M,p)\langle Az,z\rangle^p.$$

Proof. We only show the sharpness of $K(m,M,p)$ in (2.28) for $p > 1$. Let $Ax = mx$, $Ay = My$, and $z = \alpha x + \beta y$, where $\|x\| = \|y\| = 1, |\alpha|^2 + |\beta|^2 = 1$, and $h = \frac{M}{m}$. Then we have

$$\langle A^p z,z\rangle = \langle \alpha m^p x + \beta M^p y, \alpha x + \beta y\rangle = |\alpha|^2 m^p + |\beta|^2 M^p$$

and

$$\langle Az,z\rangle^p = (|\alpha|^2 m + |\beta|^2 M)^p.$$

Therefore we want to obtain the unit vector z satisfying the following equality:

$$|\alpha|^2 m^p + |\beta|^2 M^p = K(m,M,p)(|\alpha|^2 m + |\beta|^2 M)^p,$$

that is,

$$m^p + |\beta|^2(M^p - m^p) = K(m,M,p)\{m + |\beta|^2(M-m)\}^p,$$

or equivalently

$$1 + |\beta|^2(h^p - 1) = K(m,M,p)\{1 + |\beta|^2(h-1)\}^p. \tag{2.30}$$

We can obtain a solution β of the above equation (2.30) as

$$\beta = \left(\frac{h^p - 1 - p(h-1)}{(p-1)(h-1)(h^p-1)}\right)^{\frac{1}{2}} < 1.$$

For example, we have $z = \frac{1}{\sqrt{M+m}}(\sqrt{M}x + \sqrt{m}y)$ for $p = 2$. □

If we put $p = -1$ in (2.28) of Theorem 2.11, then

$$K(m,M,-1) = \frac{(M+m)^2}{4Mm}$$

is the Kantorovich constant and hence

$$\langle A^{-1}x,x\rangle \leq \frac{(M+m)^2}{4Mm}\langle Ax,x\rangle^{-1}$$

for every unit vector $x \in H$. Thus, Theorem 2.11 is an extension of Kantorovich inequality and we call $K(m,M,p)$ the generalized Kantorovich constant. We introduce another definition of $K(m,M,p)$.

Definition 2.1 *The condition number $h = h(A)$ of an invertible operator A is defined by*

$$h(A) = \|A\|\|A^{-1}\|.$$

If a positive operator A satisfies the condition $mI_H \leq A \leq MI_H$ for some scalars $0 < m \leq M$, then it may be thought as $M = \|A\|$ and $m = \|A^{-1}\|^{-1}$, so that

$$h = h(A) = \frac{M}{m}.$$

Definition 2.2 *Let $h > 0$. The generalized Kantorovich constant $K(h,p)$ is defined by*

$$K(h,p) = \frac{h^p - h}{(p-1)(h-1)}\left(\frac{p-1}{p}\frac{h^p-1}{h^p-h}\right)^p \tag{2.31}$$

for any real number $p \in \mathbb{R}$ and $K(h,p)$ is sometimes briefly denoted by $K(p)$ briefly.

We remark that $K(m,M,p)$ just coincides with $K(h,p)$ by putting $h = \frac{M}{m}(\geq 1)$. We mention basic properties of $K(h,p)$:

Theorem 2.12 *Let $h > 0$ be given. Then the generalized Kantorovich constant $K(h,p)$ has the following properties:*

(i) $K(h,p) = K(h^{-1},p)$ *for all* $p \in \mathbb{R}$,

(ii) $K\left(h, \frac{1}{2}+p\right) = K\left(h, \frac{1}{2}-p\right)$ *for all $p \in \mathbb{R}$, that is, $K(h,p)$ is symmetric with respect to $p = 1/2$,*

(iii) $K(h,0) = K(h,1) = 1$ *and* $K(1,p) = 1$ *for all* $p \in \mathbb{R}$,

(iv) *$K(h,p)$ is an increasing function of p for $p > 1/2$ and a decreasing function of p for $p < 1/2$,*

(v) $K(h,p) > 0$ *for all* $p \in \mathbb{R}$ *and*

$$K(h,p)\begin{cases} \geq 1 & \text{if } p \notin (0,1) \\ \leq 1 & \text{if } p \in [0,1]\,. \end{cases} \tag{2.32}$$

Proof. Refer to [124, Theorem 2.54] for the proof. □

Next, we present the inversion formula of the generalized Kantorovich constant and a closed relation between the condition number and the generalized Kantorovich constant:

Theorem 2.13 *Let $h > 0$ be given. Then the generalized Kantorovich constant has the following properties:*

(i) $K\left(h^r, \frac{p}{r}\right)^{\frac{1}{p}} = K\left(h^p, \frac{r}{p}\right)^{-\frac{1}{r}}$ *for* $pr \neq 0$,

(ii) $K(h,p) \leq h^{p-1}$ *for all* $p \geq 1$ *and* $h > 1$.

Proof. Refer to [124, Theorem 2.54] for the proof. □

Now we present an important constant due to Specht. He estimated the upper bound of the arithmetic mean by the geometric one for positive numbers: For $x_1, \cdots, x_n \in [m,M]$ with $M \geq m > 0$,

$$\frac{x_1 + \cdots + x_n}{n} \leq \frac{(h-1)h^{\frac{1}{h-1}}}{e \log h} \sqrt[n]{x_1 \cdots x_n}, \tag{2.33}$$

where $h = \frac{M}{m} (\geq 1)$. It is well known that

$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n} \tag{2.34}$$

holds for positive numbers $x_1, x_2, \cdots, x_n$. Therefore, the Specht theorem (2.33) means a ratio type converse inequality of the arithmetic-geometric mean inequality (2.34).

So we define the following constant.

Definition 2.3 *Let $h > 0$ be given. The Specht ratio $S(h)$ is defined by*

$$S(h) = \frac{(h-1)h^{\frac{1}{h-1}}}{e \log h} \quad (h \neq 1) \qquad \text{and} \qquad S(1) = 1. \tag{2.35}$$

Now let us show an operator version of (2.33).

Theorem 2.14 *Let A be a positive operator such that $mI_H \leq A \leq MI_H$ for some scalars $0 < m \leq M$ and put $h = \frac{M}{m}$. Then*

$$\langle Ax, x \rangle \leq S(h) \exp \langle \log Ax, x \rangle \tag{2.36}$$

holds for every unit vector $x \in H$.

Proof. Refer to [124, Theorem 2.49] for the proof. □

If we put $f(t) = \exp(t)$ in Theorem 2.8 and Theorem 2.14, then we have the following result.

Theorem 2.15 *Let A be a self-adjoint operator such that $mI_H \le A \le MI_H$ for some scalars $m \le M$. Then*

$$\exp\langle Ax,x\rangle \le \langle \exp Ax,x\rangle \le S(\mathrm{e}^{M-m})\exp\langle Ax,x\rangle \tag{2.37}$$

holds for every unit vector $x \in H$.

We mention some basic properties of the Specht ratio $S(h)$.

Theorem 2.16 *Let $h > 0$ and $p \in \mathbb{R}$.*

(i) $S(1) = \lim_{h\to 1} S(h) = 1$.

(ii) $S(h) = S(h^{-1})$.

(iii) *A function $S(h)$ is strictly decreasing for $0 < h < 1$ and strictly increasing for $h > 1$.*

(iv) $\lim_{p\to 0} S(h^p)^{\frac{1}{p}} = 1$.

(v) $\lim_{p\to\infty} S(h^p)^{\frac{1}{p}} = h$ *for* $h > 1$ *and* $\lim_{p\to\infty} S(h^p)^{\frac{1}{p}} = h^{-1}$ *for* $0 < h < 1$.

Proof. Refer to [124, Lemma 2.47] for the proof. □

We show also a closed relation between the generalized Kantorovich constant and the Specht ratio.

Theorem 2.17 *Let $h > 0$ be given. Then*

(i) $\lim_{r\to 0} K\left(h^r, \frac{p}{r}\right) = S(h^p)$,

(ii) $\lim_{r\to 0} K\left(h^r, \frac{r+p}{r}\right) = S(h^p)$.

Proof. Refer to [124, Theorem 2.56] for the proof. □

Moreover, we have the following most crucial result on the generalized Kantorovich constant.

Theorem 2.18 *Let $h > 1$. Then*

$$S(h) = \mathrm{e}^{-K'(0)} = \mathrm{e}^{K'(1)},$$

where $K(p) = K(h,p)$ for all $p \in \mathbb{R}$.

Proof. Refer to [124, Theorem 2.57] for the proof. □

We notice that the Kantorovich inequality can be interpreted as a converse of Jensen's inequality for $f(t) = t^{-1}$:

$$\langle A^{-1}x,x\rangle \leq \frac{(M+m)^2}{4Mm}\langle Ax,x\rangle^{-1}.$$

We consider a difference type converse of the Kantorovich inequality:

Theorem 2.19 *Let A be a positive operator such that $mI_H \leq A \leq MI_H$ for some scalars $0 < m \leq M$. Then*

$$\langle A^{-1}x,x\rangle - \langle Ax,x\rangle^{-1} \leq \frac{(\sqrt{M}-\sqrt{m})^2}{Mm}$$

for every unit vector $x \in H$.

Proof. Refer to [124, Theorem 1.31] for the proof. □

In a similar way, we have the following result.

Theorem 2.20 *Let A be a self-adjoint operator such that $mI_H \leq A \leq MI_H$ for some scalars $m \leq M$. Then*

$$\langle A^2x,x\rangle - \langle Ax,x\rangle^2 \leq \frac{(M-m)^2}{4}$$

for every unit vector $x \in H$.

Proof. Refer to [124, Theorem 1.30] for the proof. □

It seems that a generalization of Theorem 2.19 and Theorem 2.20 is very difficult. However, as an application of the Mond-Pečarić method, we can show a difference type converse of Jensen's inequality for convex functions:

Theorem 2.21 *Let A be a self-adjoint operator such that $mI_H \leq A \leq MI_H$ for some scalars $m < M$ and f a real valued continuous convex function on $[m,M]$. Then*

$$0 \leq \langle f(A)x,x\rangle - f(\langle Ax,x\rangle) \leq \beta(m,M,f)$$

holds for every unit vector $x \in H$, where

$$\beta(m,M,f) = \max\left\{\frac{f(M)-f(m)}{M-m}(t-m)+f(m)-f(t) : t \in [m,M]\right\}.$$

Theorem 2.22 *Let A be a self-adjoint operator such that $mI_H \leq A \leq MI_H$ for some scalars $m\langle M$ and f a real valued continuous concave function on $[m,M]$. Then*

$$\overline{\beta}(m,M,f) \leq \langle f(A)x,x\rangle - f(\langle Ax,x\rangle) \leq 0$$

holds for every unit vector $x \in H$, where

$$\overline{\beta}(m,M,f) = \min\left\{\frac{f(M)-f(m)}{M-m}(t-m)+f(m)-f(t) : t \in [m,M]\right\}.$$

If we put $f(t) = t^p$ for $p \in \mathbb{R}$ in Theorem 2.21 and Theorem 2.22, then we have a difference type converse of the Hölder-McCarthy inequality.

Theorem 2.23 *Let A be a self-adjoint operator such that $mI_H \leq A \leq MI_H$ for some scalars $m < M$. Then*

$$0 \leq \langle A^p x, x\rangle - \langle Ax, x\rangle^p \leq C(m, M, p) \qquad \text{for all } p \notin [0,1]$$

and

$$C(m, M, p) \leq \langle A^p x, x\rangle - \langle Ax, x\rangle^p \leq 0 \qquad \text{for all } p \in [0,1]$$

for every unit vector $x \in H$, where

$$C(m, M, p) = (p-1)\left(\frac{M^p - m^p}{p(M-m)}\right)^{\frac{p}{p-1}} + \frac{Mm^p - mM^p}{M-m} \tag{2.38}$$

for any real number $p \in \mathbb{R}$.

We call $C(m,M,p)$ *the Kantorovich constant for the difference* . Let us collect the basic properties of $C(m,M,p)$:

Theorem 2.24 *Let $M > m > 0$ and $p \in \mathbb{R}$.*

(i) $C(m,M,p) = \frac{mM^p - Mm^p}{M-m}\{K(m,M,p)^{\frac{1}{p-1}} - 1\}$,

(ii) $0 \leq C(m,M,p) \leq M(M^{p-1} - m^{p-1})$ *for all $p > 1$,*

(iii) $C(m,M,1) = 0$.

Proof. Refer to [124, Lemma 2.59] for the proof. □

If we put $f(t) = \log t$, $\eta(t) = -t\log t$ in Theorem 2.22 and $f(t) = \exp(t)$ in Theorem 2.21, then we have the following results.

Theorem 2.25 *Let A be a positive operator such that $0 < mI_H \leq A \leq MI_H$ for some scalars $0 < m < M$. Then*

$$-\log S(h) \leq \langle \log Ax, x\rangle - \log\langle Ax, x\rangle \leq 0$$

and

$$-\log S(h)\langle Ax, x\rangle \leq \langle \eta(A)x, x\rangle - \eta(\langle Ax, x\rangle) \leq 0$$

for every unit vector $x \in H$.

Theorem 2.26 *Let A be a self-adjoint operator such that $mI_H \leq A \leq MI_H$ for some scalars $m < M$. Then*

$$0 \leq \langle \exp Ax, x\rangle - \exp\langle Ax, x\rangle \leq \left(\frac{M\mathrm{e}^m - m\mathrm{e}^M}{M-m} + \frac{\mathrm{e}^M - \mathrm{e}^m}{M-m}\log\left(\frac{\mathrm{e}^M - \mathrm{e}^m}{\mathrm{e}(M-m)}\right)\right)$$

for every unit vector $x \in H$.

We shall give an estimate of the difference between the arithmetic mean and the geometric one:

Corollary 2.2 *For positive numbers* $x_1, \cdots, x_n \in [m,M]$ *with* $M > m > 0$ *and* $h = \frac{M}{m}$,

$$\sqrt[n]{x_1 x_2 \cdots x_n} + D(m,M) \geq \frac{x_1 + x_2 + \cdots x_n}{n} \tag{2.39}$$

where

$$D(m,M) = \theta M + (1-\theta)m - M^{\theta} m^{1-\theta} \quad \textit{and} \quad \theta = \log\left(\frac{h-1}{\log h}\right)\frac{1}{\log h}. \tag{2.40}$$

We call $D(m,M)$ the *Mond-Shisha difference*. Notice that (2.39) represents a difference type converse inequality of the arithmetic-geometric mean inequality. Recall that the logarithmic mean $L(m,M)$ is defined for $M \geq m > 0$ as

$$L(m,M) = \frac{M-m}{\log M - \log m} \ (m < M) \quad \text{and} \quad L(m,m) = m. \tag{2.41}$$

Lemma 2.4 *The Mond-Shisha difference coincides with the following constant via the Specht ratio: If* $M > m > 0$ *and* $h = \frac{M}{m} > 1$, *then*

$$D(m^p, M^p) = L(m^p, M^p) \log S(h,p) \tag{2.42}$$

for all $p \in \mathbb{R}$.

Proof. Refer to [124, Lemma 2.51] for the proof. □

The following result is considered as a continuous version of Mond-Shisha result (2.39).

Theorem 2.27 *Let A be a positive operator on H satisfying* $MI_H \geq A \geq mI_H > 0$. *Put* $h = \frac{M}{m}$. *Then the difference between* $\langle Ax, x\rangle$ *and* $\exp\langle \log Ax, x\rangle$ *at a unit vector* $x \in H$ *is not greater than the Mond-Shisha difference:*

$$\langle Ax, x\rangle - \exp\langle \log Ax, x\rangle \leq D(m,M),$$

where $D(m,M)$ *is defined in (2.40) and the equality holds if and only if both m and M are eigenvalues of A and*

$$x = \sqrt{1 - \log\left(\frac{h-1}{\log h}\right)\frac{1}{\log h}}\, e_m + \sqrt{\log\left(\frac{h-1}{\log h}\right)\frac{1}{\log h}}\, e_M,$$

where e_m *and* e_M *are corresponding unit eigenvectors to m and M, respectively.*

Proof. Refer to [124, Theorem 2.52] for the proof. □

Finally, in a general situation, we state explicitly the heart of the Mond-Pečarić method:

Theorem 2.28 *Let $f : [m,M] \mapsto \mathbb{R}$ be a convex continuous function, J an interval such that $J \subset f([m,M])$ and A a self-adjoint operator such that $mI_H \le A \le MI_H$ for some scalars $m < M$. If $F(u,v)$ is a real function defined on $J \times J$, non-decreasing in u, then*

$$F\left[\langle f(A)x,x\rangle, f(\langle Ax,x\rangle)\right] \le \max_{t\in[m,M]} F\left[\frac{f(M)-f(m)}{M-m}(t-m)+f(m), f(t)\right]$$
$$= \max_{\theta\in[0,1]} F\left[\theta f(m)+(1-\theta)f(M), f(\theta m+(1-\theta)M)\right]$$

for every unit vector $x \in H$.

This book is dedicated to applications of the Mond-Pečarić method for convex functions. One of the most important points of the Mond-Pečarić method is to offer a totally new viewpoint in the field of operator theory.

2.4 Notes

The idea of the Mond-Pečarić method is firstly proposed by Nakamura [237] for $p = -1$ in 1960, afterwards by Ky Fan [48] for any integer $p \neq 0,1$ in 1966. Finally the principle of the Mond-Pečarić method as Theorem 2.28 is established explicitly by [214] for a vector version in 1993, and [216] for an operator version, [221] for Hansen-Pedersen version and [222] for multiple vector version.

Finally, we present the following A.N. Kolmogorov's word. He said in a lecture that

"Behind every theorem lies an inequality."

A.W. Marshall, I. Olkin and B.C. Arnold
Inequalities: Theory of Majorization
and Its Applications
Second Edition

Chapter 3

Order Preserving Operator Inequality

This chapter is devoted to explain fundamental operator inequalities related to the Furuta inequality. The base point is the Löwner-Heinz inequality. It induces weighted geometric means, which serves as an excellent technical tool. The chaotic order $\log A \geq \log B$ is conceptually important in the discussion below.

3.1 From the Löwner-Heinz inequality to the Furuta inequality

The non-commutativity of operators appears in the fact that the function $t \mapsto t^2$ is not order-preserving. That is, there is a pair of positive operators A and B such that $A \geq B$ and $A^2 \not\geq B^2$. The following is a quite familiar example;

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

This implies that the function $t \mapsto t^p$ is not order-preserving for $p > 1$ by assuming the following fact.

Theorem 3.1 (LÖWNER-HEINZ INEQUALITY (LH)) *The function $t \mapsto t^p$ is order-preserving for $0 \le p \le 1$, i.e.*

$$A \ge B \ge 0 \implies A^p \ge B^p.$$

The essence of the Löwner-Heinz inequality is the case $p = \frac{1}{2}$:

$$A \ge B \ge 0 \implies A^{\frac{1}{2}} \ge B^{\frac{1}{2}}.$$

It is rephrased as follows: For $A, B \ge 0$,

$$AB^2A \le I_H \implies A^{\frac{1}{2}}BA^{\frac{1}{2}} \le I_H.$$

The assumption $AB^2A \le I_H$ is equivalent to $\|AB\| \le 1$. Thus, noting the commutativity of the spectral radius, $r(XY) = r(YX)$, we have

$$\|A^{\frac{1}{2}}BA^{\frac{1}{2}}\| = r(A^{\frac{1}{2}}BA^{\frac{1}{2}}) = r(AB) \le \|AB\| \le 1.$$

The above discussion goes to Pedersen's proof of the Löwner-Heinz inequality. As a matter of fact, the following statement is true:
Let P be the set of all $p \in [0, \frac{1}{2}]$ such that $A \ge B \ge 0$ implies $A^{2p} \ge B^{2p}$. Then P is convex.

So suppose that $A^pB^{2p}A^p \le I_H$ and $A^qB^{2q}A^q \le I_H$, or equivalently $\|A^pB^p\| \le 1$ and $\|B^qA^q\| \le 1$. Then

$$\begin{aligned}\|A^{\frac{p+q}{2}}B^{p+q}A^{\frac{p+q}{2}}\| &= r(A^{\frac{p+q}{2}}B^{p+q}A^{\frac{p+q}{2}}) = r(A^{p+q}B^{p+q}) = r(A^pB^pB^qA^q) \\ &\le \|A^pB^p\|\|B^qA^q\| \le 1.\end{aligned}$$

This implies that if $2p, 2q \in P$, then $p+q \in P$, that is, P is convex.

Related to the case $p = \frac{1}{2}$ in the Löwner-Heinz inequality, Chan-Kwong conjectured that

$$A \ge B \ge 0 \implies (AB^2A)^{\frac{1}{2}} \le A^2.$$

Moreover, if it is true, then the following inequality holds:

$$A \ge B \ge 0 \implies (BA^2B)^{\frac{1}{2}} \ge B^2.$$

Here we cite a useful lemma on exponent.

Lemma 3.1 *For $p \in \mathbb{R}$, $(X^*A^2X)^p = X^*A(AXX^*A)^{p-1}AX$ for $A > 0$ and invertible X.*

Proof. It is easily checked that $Y^*(YY^*)^nY = Y^*Y(Y^*Y)^n$ for any $n \in \mathbb{N}$. This implies that $Y^*f(YY^*)Y = Y^*Yf(Y^*Y)$ for any polynomials f and so it holds for continuous functions f on a suitable interval. Hence we have the conclusion by applying it to $f(x) = x^p$ and $Y = AX$. □

Consequently, the Chan-Kwong conjecture is modified in the following sense: If it is true, then

$$A \ge B \ge 0 \implies (AB^2A)^{\frac{3}{4}} \le A^3.$$

As a matter of fact, we have

$$(AB^2A)^{\frac{3}{4}} = AB(BA^2B)^{-\frac{1}{4}}BA = AB((BA^2B)^{-\frac{1}{2}})^{\frac{1}{2}}BA$$
$$\leq ABB^{-1}BA = ABA \leq A^3.$$

Based on this consideration, the Furuta inequality was established.

Theorem 3.2 (FURUTA INEQUALITY (FI)) *If $A \geq B \geq 0$, then for each $r \geq 0$,*

$$(A^{\frac{r}{2}}A^pA^{\frac{r}{2}})^{\frac{1}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}} \tag{i}$$

and

$$(B^{\frac{r}{2}}A^pB^{\frac{r}{2}})^{\frac{1}{q}} \geq (B^{\frac{r}{2}}B^pB^{\frac{r}{2}})^{\frac{1}{q}} \tag{ii}$$

hold for $p \geq 0$ and $q \geq 1$ with

$$(1+r)q \geq p+r. \tag{$*$}$$

The domain $(*)$ is drawn as in Figure 3.1.

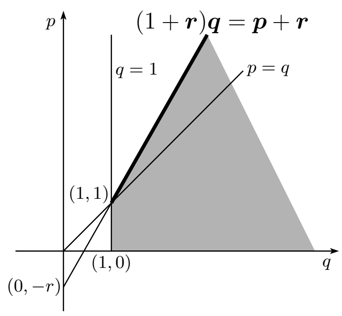


Figure 3.1: The domain $(*)$

It is a quite important information on (FI) that the domain defined by $(*)$ is the best possible in the sense that it cannot extend. It is proved by Tanahashi [277]:

If $p,q,r > 0$ satisfy either $(1+r)q < p+r$ or $q < 1$, then there exist $A,B > 0$ such that $A \geq B$ and

$$A^{\frac{p+r}{q}} \not\geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}}.$$

Professor Berberian said that Figure 3.1 is "Rosetta Stone" in (FI). Incidentally it is notable that Figure 3.1 is expressed by qp–axis: Berberian's interesting comment might contain it.

Proof of (FI). It suffices to show that if $A \geq B > 0$, then

$$(A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1+r}{p+r}} \leq A^{1+r}.$$

It is proved for arbitrary $p \geq 1$ by the induction on r. First of all, we take $r \in [0,1]$.

$$(A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1+r}{p+r}} = A^{\frac{r}{2}}B^{\frac{p}{2}}(B^{\frac{p}{2}}A^rB^{\frac{p}{2}})^{\frac{1-p}{p+r}}B^{\frac{p}{2}}A^{\frac{r}{2}}$$

$$\leq A^{\frac{r}{2}}B^{\frac{p}{2}}(B^{\frac{p}{2}}B^rB^{\frac{p}{2}})^{\frac{1-p}{p+r}}B^{\frac{p}{2}}A^{\frac{r}{2}} = A^{\frac{r}{2}}BA^{\frac{r}{2}} \leq A^{\frac{r}{2}}AA^{\frac{r}{2}} = A^{1+r}.$$

Next we suppose that it is true for some $r_1 > 0$, i.e.

$$B_1 = (A^{\frac{r_1}{2}}B^pA^{\frac{r_1}{2}})^{\frac{1+r_1}{p+r_1}} \leq A^{1+r_1} = A_1.$$

Then for $r \in (0,1]$

$$(A_1^{\frac{r}{2}}B_1^{\frac{p+r_1}{1+r_1}}A_1^{\frac{r}{2}})^{\frac{1+r}{p_1+r}} \leq A_1^{1+r},$$

where $p_1 = \frac{p+r_1}{1+r_1}$. Putting $s = r_1 + (1+r_1)r = (1+r_1)(1+r) - 1$, we have

$$(A^{\frac{s}{2}}B^pA^{\frac{s}{2}})^{\frac{1+s}{p+s}} \leq A^{1+s},$$

This means that it is true for $s \in [r_1, 1+2r_1]$. Hence the proof is complete. □

To make clear the structure of (FI), we give a mean theoretic approach to (FI).

The Löwner-Heinz inequality says that the function t^α is operator monotone for $\alpha \in [0,1]$. It induces the α-geometric operator mean defined for $\alpha \in [0,1]$ as

$$A \#_\alpha B = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha A^{\frac{1}{2}}$$

if $A > 0$, i.e. A is invertible, by the Kubo-Ando theory [165].

For the sake of convenience, we cite a useful lemma which we will use frequently in the below.

Lemma 3.2 *For $X, Y > 0$ and $a,\ b \in [0,1]$,*

(i) *monotonicity: $X \leq X_1$ and $Y \leq Y_1 \implies X \#_a Y \leq X_1 \#_a Y_1$,*

(ii) *transformer equality: $T^*XT \#_a T^*YT = T^*(X \#_a Y)T$ for invertible T,*

(iii) *transposition: $X \#_a Y = Y \#_{1-a} X$,*

(iv) *multiplicity: $X \#_{ab} Y = X \#_a (X \#_b Y)$.*

Proof. First of all, (iii) follows from Lemma 3.1, and (iv) does from a direct computation under the assumption of invertibility of operators.

To prove (i), we may assume that $X, Y > 0$. If $Y \leq Y_1$, then $X \#_a Y \leq X \#_a Y_1$ is assured by (LH) (and the formula of $\#_a$). Moreover the monotonicity of the other is shown by the use of (iii).

Finally (ii) is obtained by Jensen's inequality (JI) which is discussed in Theorem 3.45. We put $Z = X^{\frac{1}{2}}T$. Then it follows from (JI) that

$$\begin{aligned} T^*XT \#_a T^*YT &= Z^*Z \#_a T^*YT \\ &= |Z|\big(|Z|^{-1}T^*YT|Z|^{-1}\big)^a|Z| \\ &= |Z|\left(|Z|^{-1}Z^*\big(X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\big)Z|Z|^{-1}\right)^a|Z| \\ &\geq Z^*\big(X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\big)^aZ \\ &= T^*X^{\frac{1}{2}}\big(X^{-\frac{1}{2}}YX^{-\frac{1}{2}}\big)^aX^{\frac{1}{2}}T \\ &= T^*(X \#_a Y)T, \end{aligned}$$

because $Z|Z|^{-1} = V$ is the partial inequality in the polar decomposition of Z and so a contraction. □

By using the mean theoretic notation, the Furuta inequality has the following expression:

(FI) *If* $A \geq B > 0$, *then*

$$A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq A \quad \textit{for} \quad p \geq 1 \textit{ and } r \geq 0. \tag{3.1}$$

Related to this, we have to mention the following more precise expression of it. We say it a satellite inequality of (FI), simply (SF).

Theorem 3.3 (SATELLITE INEQUALITY (SF)) *If* $A \geq B > 0$, *then*

$$A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq B \leq A \quad \textit{for} \quad p \geq 1 \textit{ and } r \geq 0. \tag{3.2}$$

Proof. As the first stage, we assume that $0 \leq r \leq 1$. Then the monotonicity of $\#_\alpha$ $(\alpha \in [0,1])$ implies that

$$A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq B^{-r} \#_{\frac{1+r}{p+r}} B^p = B.$$

Next we assume that for some $r > 0$,

$$A \geq B > 0 \implies A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq B \leq A$$

holds for all $p \geq 1$. So we prove that it is true for $s = 1 + 2r$. Since $A \geq B > 0$ is assumed, we have

$$A^{-1} \#_{\frac{2}{p+1}} B^p \leq B,$$

so that

$$B_1 = (A^{\frac{1}{2}}B^pA^{\frac{1}{2}})^{\frac{2}{p+1}} \leq A^{\frac{1}{2}}BA^{\frac{1}{2}} \leq A^2 = A_1.$$

By the assumption, it follows that for $p_1 \geq 1$

$$A_1^{-r} \#_{\frac{1+r}{p_1+r}} B_1^p \leq B_1 \leq A^{\frac{1}{2}}BA^{\frac{1}{2}}.$$

Arranging this for $p_1 = \frac{p+1}{2}$, we have

$$A^{-2r} \#_{\frac{2(1+r)}{p+1+2r}} A^{\frac{1}{2}} B^p A^{\frac{1}{2}} \le B_1 \le A^{\frac{1}{2}} B A^{\frac{1}{2}}.$$

Furthermore multiplying $A^{-\frac{1}{2}}$ on both sides, it follows that for $s = 2r+1$

$$A^{-s} \#_{\frac{1+s}{p+s}} B^p \le B,$$

as desired. □

3.2 The Ando-Hiai inequality

Ando and Hiai proposed a log-majorization inequality, whose essential part is the following operator inequality. We say it the Ando-Hiai inequality, simply (AH).

Theorem 3.4 (ANDO-HIAI INEQUALITY (AH)) *If $A \#_\alpha B \le I_H$ for $A, B > 0$, then $A^r \#_\alpha B^r \le I_H$ for $r \ge 1$.*

Proof. It suffices to show that $A^r \#_\alpha B^r \le I_H$ for $1 \le r \le 2$. Put $p = r-1 \in [0,1]$ and $C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$. Then, since the assumption $A \#_\alpha B \le I_H$ is equivalent to $C^\alpha \le A^{-1}$ and so $C^{-\alpha} \ge A$, it follows from Lemma 3.1 that

$$\begin{aligned} A^{-\frac{1}{2}} B^r A^{-\frac{1}{2}} &= A^{-\frac{1}{2}} (A^{\frac{1}{2}} C A^{\frac{1}{2}})^r A^{-\frac{1}{2}} = C^{\frac{1}{2}} (C^{\frac{1}{2}} A C^{\frac{1}{2}})^p A^{-\frac{1}{2}} \\ &\le C^{\frac{1}{2}} (C^{\frac{1}{2}} C^{-\alpha} C^{\frac{1}{2}})^p C^{\frac{1}{2}} = C^{1+(1-\alpha)p}. \end{aligned}$$

Hence we have

$$\begin{aligned} A^r \#_\alpha B^r &= A^{\frac{1}{2}} (A^p \#_\alpha A^{-\frac{1}{2}} B^r A^{-\frac{1}{2}}) A^{\frac{1}{2}} \le A^{\frac{1}{2}} (C^{-\alpha p} \#_\alpha C^{1+(1-\alpha)p}) A^{\frac{1}{2}} \\ &= A^{\frac{1}{2}} C^{(1+p)\alpha - \alpha p} A^{\frac{1}{2}} = A^{\frac{1}{2}} C^\alpha A^{\frac{1}{2}} \le A^{\frac{1}{2}} A^{-1} A^{\frac{1}{2}} = I_H. \end{aligned}$$

□

Based on an idea of the Furuta inequality, we propose two variables version of the Ando-Hiai inequality:

Theorem 3.5 (GENERALIZED ANDO-HIAI INEQUALITY (GAH)) *For $A, B > 0$ and $\alpha \in [0,1]$, if $A \#_\alpha B \le I_H$, then*

$$A^r \#_{\frac{\alpha r}{\alpha r + (1-\alpha)s}} B^s \le I_H \quad \text{for} \quad r, s \ge 1.$$

It is obvious that the case $r = s$ in Theorem 3.5 is just the Ando-Hiai inequality.

Now we consider two one-sided versions of Theorem 3.5:

Proposition 3.1 *For $A, B > 0$ and $\alpha \in [0,\ 1]$, if $A \#_\alpha B \le I_H$, then*

$$A^r \#_{\frac{\alpha r}{\alpha r+1-\alpha}} B \le I_H \quad \textit{for} \quad r \ge 1.$$

Proposition 3.2 *For $A, B > 0$ and $\alpha \in [0,\ 1]$, if $A \#_\alpha B \le I_H$, then*

$$A \#_{\frac{\alpha}{\alpha+(1-\alpha)s}} B^s \le I_H \quad \textit{for} \quad s \ge 1.$$

Next we investigate relations among these propositions and Theorem 3.5.

Theorem 3.6 (1) *Propositions 3.1 and 3.2 are equivalent.*

(2) *Theorem 3.5 follows from Propositions 3.1 and 3.2.*

Proof.
(1) We first note the transposition formula $X \#_\alpha Y = Y \#_\beta X$ for $\beta = 1-\alpha$. Therefore Proposition 3.1 (for β) is rephrased as follows:

$$B \#_\beta A \le I_H \quad \Longrightarrow \quad B^s \#_{\frac{\beta s}{\beta s+\alpha}} A \le I_H \quad \text{for} \quad s \ge 1.$$

Using the transposition formula again, it coincides with Proposition 3.2 because

$$1 - \frac{\beta s}{\beta s + \alpha} = \frac{\alpha}{\beta s + \alpha} = \frac{\alpha}{(1-\alpha)s + \alpha}.$$

(2) Suppose that $A \#_\alpha B \le I_H$ and $r, s \ge 1$ are given. Then it follows from Proposition 3.1 that $A^r \#_{\alpha_1} B \le I_H$ for $\alpha_1 = \frac{\alpha r}{\alpha r+1-\alpha}$. We next apply Proposition 3.2 to it, so that we have

$$I_H \ge A^r \#_{\frac{\alpha_1}{\alpha_1+(1-\alpha_1)s}} B^s = A^r \#_{\frac{\alpha r}{\alpha r+(1-\alpha)s}} B^s,$$

as desired. □

We now point out that Proposition 3.1 is an equivalent expression of the Furuta inequality of the Ando-Hiai type:

Theorem 3.7 *The inequality in Proposition 3.1 is equivalent to the Furuta inequality.*

Proof. For a given $p \ge 1$, we put $\alpha = \frac{1}{p}$. Then $A \ge B (\ge 0)$ if and only if

$$A^{-1} \#_\alpha B_1 \le I_H, \text{ for } B_1 = A^{-\frac{1}{2}} B^p A^{-\frac{1}{2}}. \tag{3.3}$$

If $A \ge B > 0$, then (3.3) holds for $A, B > 0$, so that Proposition 3.1 implies that for any $r \ge 0$

$$I_H \ge A^{-(r+1)} \#_{\frac{\frac{r+1}{p}}{(1-\frac{1}{p})+\frac{r+1}{p}}} B_1 = A^{-(r+1)} \#_{\frac{1+r}{p+r}} B_1 = A^{-(r+1)} \#_{\frac{1+r}{p+r}} A^{-\frac{1}{2}} B^p A^{-\frac{1}{2}}.$$

Hence we have (FI);

$$A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq A.$$

Conversely suppose that (FI) is assumed. If $A^{-1} \#_\alpha B_1 \leq I_H$, then $A \geq (A^{\frac{1}{2}} B_1 A^{\frac{1}{2}})^\alpha = B$, where $p = \frac{1}{\alpha}$. So (FI) implies that for $r_1 = r - 1 \geq 0$

$$A \geq A^{-r_1} \#_{\frac{1+r_1}{p+r_1}} B^p = A^{-(r-1)} \#_{\frac{r}{p+r-1}} A^{\frac{1}{2}} B_1 A^{\frac{1}{2}}.$$

Since $\frac{r}{p+r-1} = \frac{\alpha r}{1+\alpha r-\alpha}$, we have Proposition 3.1. □

As in the discussion as above, Theorem 3.5 can be proved by showing Proposition 3.1. Finally we cite its proof. Since it is equivalent to the Furuta inequality, we have an alternative proof of it. It is done by the usual induction, whose technical point is a multiplicative property of the index $\frac{\alpha r}{(1-\alpha)+\alpha r}$ of # as appeared below.

Proof of Proposition 3.1. For convenience, we show that if $A^{-1} \#_\alpha B \leq I_H$, then

$$A^{-r} \#_{\frac{\alpha r}{(1-\alpha)+\alpha r}} B \leq I_H \quad \text{for} \quad r \geq 1. \tag{3.4}$$

Now the assumption says that

$$C^\alpha = (A^{\frac{1}{2}} B A^{\frac{1}{2}})^\alpha \leq A.$$

For any $\varepsilon \in (0,1]$, we have $C^{\alpha\varepsilon} \leq A^\varepsilon$ by the Löwner-Heinz inequality and so

$$\begin{aligned} A^{-(1+\varepsilon)} \#_{\frac{\alpha(1+\varepsilon)}{(1-\alpha)+\alpha(1+\varepsilon)}} B &= A^{-\frac{1}{2}} (A^{-\varepsilon} \#_{\frac{\alpha(1+\varepsilon)}{1+\alpha\varepsilon}} A^{\frac{1}{2}} B A^{\frac{1}{2}}) A^{-\frac{1}{2}} \\ \leq A^{-\frac{1}{2}} (C^{-\alpha\varepsilon} \#_{\frac{\alpha(1+\varepsilon)}{1+\alpha\varepsilon}} C) A^{-\frac{1}{2}} &= A^{-\frac{1}{2}} C^\alpha A^{-\frac{1}{2}} = A^{-1} \#_\alpha B \leq I_H. \end{aligned}$$

Hence we proved the conclusion (3.4) for $1 \leq r \leq 2$. So we next assume that (3.4) holds for $1 \leq r \leq 2^n$. Then the discussion of the first half ensures that

$$(A^{-r})^{r_1} \#_{\frac{\alpha_1 r_1}{(1-\alpha_1)+\alpha_1 r_1}} B \leq I_H \quad \text{for } 1 \leq r_1 \leq 2, \text{ where } \alpha_1 = \frac{\alpha r}{(1-\alpha)+\alpha r}.$$

Thus the multiplicative property of the index

$$\frac{\alpha_1 r_1}{(1-\alpha_1)+\alpha_1 r_1} = \frac{\alpha r r_1}{(1-\alpha)+\alpha r r_1}$$

shows that (3.4) holds for all $r \geq 1$. □

Here we consider an expression of (AH)-type for satellite of (FI): Suppose that $A^{-1} \#_\alpha B \leq I_H$ and put $\alpha = \frac{1}{p}$. It is equivalent to $C = (A^{\frac{1}{2}} B A^{\frac{1}{2}})^{\frac{1}{p}} \leq A$. So (SF) says that

$$A^{-r} \#_{\frac{1+r}{p+r}} C^p \leq C, \quad \text{or} \quad A^{-(r+1)} \#_{\frac{1+r}{p+r}} B \leq A^{-\frac{1}{2}} C A^{-\frac{1}{2}} = A^{-1} \#_{\frac{1}{p}} B.$$

Namely (SF) has an (AH)-type representation as follows:

Theorem 3.8 *Let A and B be positive invertible operators. Then*

$$A \#_\alpha B \le I_H \implies A^r \#_{\frac{\alpha r}{\alpha r+1-\alpha}} B \le A \#_\alpha B \ (\le I_H) \ \text{ for } \ r \ge 1.$$

As an application, we have the monotonicity of the operator function induced by (GAH):

Theorem 3.9 *If $A \#_\alpha B \le I$ for $A, B > 0$, then*

$$f(r,s) = A^r \#_{\frac{\alpha r}{\alpha r+(1-\alpha)s}} B^s$$

is decreasing for $r, \ s \ge 1$.

Proof. It suffices to show that f is decreasing for $r \ge 1$ because $f_{\alpha,A,B}(r,s) = f_{1-\alpha,B,A}(s,r)$. So we fix $s \ge 1$.

By (GAH), it follows that for each $r \ge 1$

$$f(r,s) = A^r \#_{\alpha_1} B \le I_H, \quad \text{where} \quad \alpha_1 = \frac{\alpha r}{\alpha r + (1-\alpha)s}.$$

For arbitrary $r_2 > r$, we put $r_1 = \frac{r_2}{r} > 1$. Then we have

$$f(r_2,s) = A^{r_2} \#_{\frac{\alpha r_2}{\alpha r_2+(1-\alpha)s}} B^s = (A^r)^{r_1} \#_{\frac{\alpha_1 r_1}{\alpha_1 r_1+(1-\alpha_1)}} B^s \le A^r \#_{\alpha_1} B^s = f(r,s)$$

by Theorem 3.8. □

3.3 The grand Furuta inequality

To compare (AH) with (FI), (AH) is arranged as a Furuta type operator inequality. As in the proof of (AH), its assumption is that

$$B_1 = C^\alpha = (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^\alpha \le A^{-1} = A_1.$$

Replacing $p = \alpha^{-1}$, it is reformulated that

$$A \ge B > 0 \implies A^r \ge \left(A^{\frac{r}{2}} (A^{-\frac{1}{2}} B^p A^{-\frac{1}{2}})^r A^{\frac{r}{2}} \right)^{\frac{1}{p}} \tag{†}$$

for $r \ge 1$ and $p \ge 1$.

Moreover, to make a simultaneous extension of both (FI) and (AH), Furuta added variables as in the case of (FI). Actually he paid his attention to $A^{-\frac{1}{2}}$ in (†), precisely, he replaced it to $A^{-\frac{t}{2}}$ $(t \in [0,1])$. Consequently he established so-called the grand Furuta inequality, simply (GFI). It is sometimes said to be a generalized Furuta inequality.

Theorem 3.10 (Grand Furuta inequality (GFI)) *If $A \geq B > 0$ and $t \in [0,1]$, then*

$$\left[A^{\frac{r}{2}}\left(A^{-\frac{t}{2}}B^pA^{-\frac{t}{2}}\right)^sA^{\frac{r}{2}}\right]^{\frac{1-t+r}{(p-t)s+r}} \leq A^{1-t+r}$$

holds for $r \geq t$ and $p,s \geq 1$.

It is easily seen that

(GFI) for $t=1,\ r=s \iff$ (AH)
(GFI) for $t=0,\ s=1 \iff$ (FI).

Proof of (GFI). We prove it by the induction on s. For this, we first prove it for $1 \leq s \leq 2$: Since $(X^*C^2X)^s = X^*C(CXX^*C)^{s-1}CX$ for arbitrary X and $C \geq 0$, and $0 \leq s-1 \leq 1$, (LH) implies that

$$\begin{aligned}A^{r/2}(A^{-t/2}B^pA^{-t/2})^sA^{r/2} &= A^{\frac{r-t}{2}}B^{\frac{p}{2}}(B^{\frac{p}{2}}A^{-t}B^{\frac{p}{2}})^{s-1}B^{\frac{p}{2}}A^{\frac{r-t}{2}}\\ &\leq A^{\frac{r-t}{2}}B^{\frac{p}{2}}(B^{\frac{p}{2}}B^{-t}B^{\frac{p}{2}})^{s-1}B^{\frac{p}{2}}A^{\frac{r-t}{2}} = A^{\frac{r-t}{2}}B^{(p-t)s+t}A^{\frac{r-t}{2}}.\end{aligned}$$

Furthermore it follows from (LH) and (FI) that

$$\left\{A^{r/2}\left(A^{-t/2}B^pA^{-t/2}\right)^sA^{r/2}\right\}^{\frac{1-t+r}{(p-t)s+r}} \leq \left\{A^{\frac{r-t}{2}}B^{(p-t)s+t}A^{\frac{r-t}{2}}\right\}^{\frac{1-t+r}{(p-t)s+r}} \leq A^{1-t+r}$$

by noting that $(p-t)s+t+(r-t) = (p-t)s+r$. Hence (GFI) is proved for $1 \leq s \leq 2$.

Next, under the assumption (GFI) holds for some $s \geq 1$, we now prove that (GFI) holds for $s+1$. Since (GFI) holds for s, we take $r=t$ in it. Thus we have

$$A \geq \left\{A^{t/2}\left(A^{-t/2}B^pA^{-t/2}\right)^sA^{t/2}\right\}^{\frac{1}{(p-t)s+t}}.$$

Put $C = \{A^{t/2}(A^{-t/2}B^pA^{-t/2})^sA^{t/2}\}^{\frac{1}{(p-t)s+t}}$, that is, $A \geq C$. By using that $s \geq 1$ if and only if $1 \leq \frac{s+1}{s} \leq 2$ and that (GFI) for $1 \leq s \leq 2$ has been proved, we obtain that

$$\begin{aligned}A^{1-t+r} &\geq \left\{A^{r/2}(A^{-t/2}C^{(p-t)s+t}A^{-t/2})^{\frac{s+1}{s}}A^{r/2}\right\}^{\frac{1-t+r}{\{(p-t)s+t-t\}(\frac{s+1}{s})+r}}\\ &= \left\{A^{r/2}(A^{-t/2}C^{(p-t)s+t}A^{-t/2})^{\frac{s+1}{s}}A^{r/2}\right\}^{\frac{1-t+r}{(p-t)(s+1)+r}}\\ &= \left\{A^{r/2}(A^{-t/2}\{A^{t/2}(A^{-t/2}B^pA^{-t/2})^sA^{t/2}\}A^{-t/2})^{\frac{s+1}{s}}A^{r/2}\right\}^{\frac{1-t+r}{(p-t)(s+1)+r}}\\ &= \left\{A^{r/2}(A^{-t/2}B^pA^{-t/2})^{s+1}A^{r/2}\right\}^{\frac{1-t+r}{(p-t)(s+1)+r}}.\end{aligned}$$

This means that (GFI) holds for $s+1$, and so the proof is complete. □

Next we point out that (GFI) for $t=1$ includes both: the Ando-Hiai and Furuta inequality.

Since the Ando-Hiai inequality is just (GFI; $t=1$) for $r=s$, it suffices to check that the Furuta inequality is contained in (GFI; $t=1$). As a matter of fact, it is just (GFI; $t=1$) for $s=1$.

Theorem 3.11 *Furuta inequality* (FI) *is equivalent to* (GFI) *for* $t = s = 1$.

Proof. We write down (GFI; $t = 1$) for $s = 1$: If $A \geq B > 0$, then

$$\left[A^{\frac{r}{2}}\left(A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}}\right)A^{\frac{r}{2}}\right]^{\frac{r}{p-1+r}} \leq A^r$$

for $p, r \geq 1$, or equivalently,

$$A^{-(r-1)} \#_{\frac{r}{p-1+r}} B^p \leq A$$

for $p, r \geq 1$. Replacing $r-1$ by r_1, (GFI; $t = 1$) for $s = 1$ is rephrased as follows: If $A \geq B > 0$, then

$$A^{-r_1} \#_{\frac{1+r_1}{p+r_1}} B^p \leq A$$

for $p \geq 1$ and $r_1 \geq 0$, which is nothing but the Furuta inequality. □

Furthermore Theorem 3.5, the generalized Ando-Hiai inequality, is understood as the case $t = 1$ in (GFI):

Theorem 3.12 (GFI; $t = 1$) *is equivalent to Theorem 3.5.*

Proof. (GFI; $t = 1$) is written as

$$A \geq B > 0 \implies \left[A^{\frac{r}{2}}\left(A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}}\right)^sA^{\frac{r}{2}}\right]^{\frac{r}{(p-1)s+r}} \leq A^r \quad (p, r, s \geq 1).$$

Here we put

$$\alpha = \frac{1}{p}, \quad B_1 = A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}}.$$

Then we have

$$A \geq B > 0 \iff A^{-1} \#_{\frac{1}{p}} A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}} \leq I_H \iff A^{-1} \#_\alpha B_1 \leq I_H$$

and for each $p, r, s \geq 1$

$$\begin{aligned} &[A^{\frac{r}{2}}(A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}})^sA^{\frac{r}{2}}]^{\frac{r}{(p-1)s+r}} \leq A^r \\ \iff &A^{-r} \#_{\frac{r}{(p-1)s+r}} (A^{-\frac{1}{2}}B^pA^{-\frac{1}{2}})^s \leq I_H \\ \iff &A^{-r} \#_{\frac{\alpha r}{\alpha r+(1-\alpha)s}} B_1^s \leq I_H. \end{aligned}$$

This shows the statement of Theorem 3.5 (GAH). □

Next we consider some variants of (GFI), which are useful in the discussion of Kantorovich type inequalities.

Theorem 3.13 *If* $A \geq B \geq 0$, *then*

$$A^{\frac{(p+t)s+r}{q}} \geq \left(A^{\frac{r}{2}}\left(A^{\frac{t}{2}}B^pA^{\frac{t}{2}}\right)^sA^{\frac{r}{2}}\right)^{\frac{1}{q}}$$

holds for all $p,\ t,\ s,\ r \geq 0$ *and* $q \geq 1$ *with* $(p+t+r)q \geq (p+t)s+r$ *and* $(1+t+r)q \geq (p+t)s+r$.

Proof. First of all, we may assume $p > 0$. Now the Furuta inequality says that

$$A_1 = A^{\frac{p+t}{q_1}} \geq B_1 = \left(A^{\frac{t}{2}} B^p A^{\frac{t}{2}}\right)^{\frac{1}{q_1}}$$

holds for $t \geq 0$, where $q_1 = \max\{1, \frac{p+t}{1+t}\}$. Applying the Furuta inequality again, we have

$$A_1^{\frac{p_1+r_1}{q}} \geq \left(A_1^{\frac{r_1}{2}} B_1^{p_1} A_1^{\frac{r_1}{2}}\right)^{\frac{1}{q}},$$

that is,

$$A^{\frac{(p+t)(p_1+r_1)}{qq_1}} \geq \left(A^{\frac{(p+t)r_1}{2q_1}} \left(A^{\frac{t}{2}} B^p A^{\frac{t}{2}}\right)^{\frac{p_1}{q_1}} A^{\frac{(p+t)r_1}{2q_1}}\right)^{\frac{1}{q}},$$

for all $p_1,\ r_1 \geq 0$ and $q \geq 1$ with $(1+r_1)q \geq p_1 + r_1$. So we take $p_1 = sq_1$ and $r_1 = \frac{rq_1}{p+t}$. Since $(1+r_1)q \geq p_1 + r_1$ is equivalent to the condition that $(p+t+r)q \geq (p+t)s + r$ and $(1+t+r)q \geq (p+t)s + r$, the statement is proved. □

In the remainder, we reconsider (GFI). For this, we cite it by the use of operator means. For convenience, we use the notation $\natural_s$ for the binary operation

$$A \natural_s B = A^{\frac{1}{2}} (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^s A^{\frac{1}{2}} \quad \text{for } s \notin [0,1],$$

whose formula is the same as $\#_s$.

GRAND FURUTA INEQUALITY (GFI)

$$A \geq B > 0, t \in [0,1] \implies A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) \leq A \ (r \geq t;\ p, s \geq 1)$$

This mean theoretic expression of (GFI) induces the following improvement of it.

SATELLITE OF THE GRAND FURUTA INEQUALITY (SGF)

$$A \geq B > 0, t \in [0,1] \implies A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) \leq B \ (r \geq t;\ p, s \geq 1)$$

Here we clarify that the case $t = 1$ is essential in (GFI), in which SGF) is quite meaningful. As a matter of fact, we prove that (SGF; t=1) implies (SGF) for every $t \in [0,1]$.

For the reader's convenience, we prove (SGF). For this, the following lemma is needed, which is a variational expression of (LH).

Lemma 3.3 *If $A \geq B > 0$, $t \in [0,1]$ and $1 \leq s \leq 2$, then*

$$A^t \natural_s C \leq B^t \natural_s C$$

holds for arbitrary $C > 0$, in particular,

$$A^t \natural_s B^p \leq B^{(p-t)s+t}$$

holds for $p \geq 1$.

Proof. Since $A^{-t} \leq B^{-t}$ by (LH), we have

$$A^t \natural_s C = C(C^{-1} \#_{s-1} A^{-t})C \leq C(C^{-1} \#_{s-1} B^{-t})C = B^t \natural_s C.$$

Similarly we have

$$A^t \natural_s B^p = B^p(B^{-p} \#_{s-1} A^{-t})B^p \leq B^p(B^{-p} \#_{s-1} B^{-t})B^p = B^{(p-t)s+t}.$$

□

Here we give a short comment on the first statement in the above lemma: Suppose that $A \geq B > 0$ and $t \in [0,1]$. Then

$$A^t \natural_s C \leq B^t \natural_s C$$

holds for arbitrary $C > 0$ and $1 \leq s \leq 2$. Then taking $C = B^t$ and $s = 2$, we have

$$A^t \natural_2 B^t \leq B^t \natural_2 B^t = B^t,$$

so that $B^t A^{-t} B^t \leq B^t$, or $A^t \geq B^t$. That is, it is equivalent to (LH).

More generally, we know the following fact.

Lemma 3.4 *If $A \geq B > 0$ and $t \in [0,1]$, then*

$$(A^t \natural_s B^p)^{\frac{1}{(p-t)s+t}} \leq B \leq A$$

holds for $p,s \geq 1$.

Proof. We fix $p \geq 1$ and $t \in [0,1]$. It follows from Lemma 3.3.5 and (LH) that

$$A \geq B > 0 \implies B_1 = (A^t \natural_s B^p)^{\frac{1}{(p-t)s+t}} \leq B \leq A \qquad (**)$$

for $s \in [1,2]$. So we assume that $(**)$ holds for some $s \geq 1$, and prove that

$$B_2 = (A^t \natural_{2s} B^p)^{\frac{1}{2(p-t)s+t}} \leq B_1 \leq B.$$

Actually we apply (†) to $B_1 \leq A$. Then we have

$$(A^t \natural_2 B_1^{p_1})^{\frac{1}{2(p_1-t)+t}} \leq B_1 \leq B, \text{ where } p_1 = (p-t)s+t,$$

and moreover

$$(A^t \natural_2 B_1^{p_1})^{\frac{1}{2(p_1-t)+t}} = \left[A^t \natural_2 \left(A^t \natural_s B^p\right)\right]^{\frac{1}{(p-t)2s+t}} = (A^t \natural_{2s} B^p)^{\frac{1}{(p-t)2s+t}} = B_2,$$

which completes the proof. □

Under this preparation, we can easily prove (SGF) by virtue of (SF) in Theorem 3.3:

Proof of (SGF). For given p,t,s, we use the same notation as above; $p_1 = (p-t)s+t$ and $B_1 = (A^t \natural_s B^p)^{\frac{1}{p_1}}$. Then Lemma 3.3.6 implies that $B_1 \le B \le A$. Hence it follows from (SF) for $B_1 \le A$ and $r_1 = r-t$ that

$$A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) = A^{r_1} \#_{\frac{1+r_1}{p_1+r_1}} B_1^{p_1} \le B_1 \le B.$$

□

It is shown that (SGF; $t=1$) is essential among (SGF; $t \in [0,1]$), in which (LH) completely works. That is,

Theorem 3.14 (SGF; $t=1$) *implies* (SGF; t) *for* $t \in [0,1]$.

Proof. Suppose that for $A \ge B > 0$,

$$A^{-r+1} \#_{\frac{r}{(p-1)s+r}} (A \natural_s B^p) \le B$$

holds for $r \ge 1$.

We fix arbitrary $t \in (0,1)$. As $A^t \ge B^t$ by (LH), we have

$$(A^t)^{-\frac{r}{t}+1} \#_{\frac{\frac{r}{t}}{(\frac{p}{t}-1)s+\frac{r}{t}}} (A^t \natural_s B^p) \le B^t$$

for $r \ge t$. It is arranged as

$$A^{-r+t} \#_{\frac{r}{(p-t)s+r}} (A^t \natural_s B^p) \le B^t,$$

or equivalently,

$$(A^t \natural_s B^p) \#_{\frac{(p-t)s}{(p-t)s+r}} A^{-r+t} \le B^t.$$

Therefore it follows from Lemma 3.3 that for $s \in [1,2]$

$$\begin{aligned} A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) &= (A^t \natural_s B^p) \#_{\frac{(p-t)s-(1-t)}{(p-t)s+r}} A^{-r+t} \\ &= (A^t \natural_s B^p) \#_{\frac{(p-t)s-(1-t)}{(p-t)s}} \left\{ (A^t \natural_s B^p) \#_{\frac{(p-t)s}{(p-t)s+r}} A^{-r+t} \right\} \\ &\le (A^t \natural_s B^p) \#_{\frac{(p-t)s-(1-t)}{(p-t)s}} B^t \\ &= B^t \#_{\frac{1-t}{(p-t)s}} (A^t \natural_s B^p) \\ &\le B^t \#_{\frac{1-t}{p-t}} B^{(p-t)s+t} = B. \end{aligned}$$

Namely we have

$$A \ge B > 0 \implies A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) \le B \qquad (***)$$

for $1 \le s \le 2$, $r \ge t$ and $p \ge 1$.

Next we assume that $(***)$ holds for some $s \geq 1$. Then taking $r = t$, we have

$$B \geq (A^t \natural_s B^p)^{\frac{1}{(p-t)s+t}}.$$

Put $C = (A^t \natural_s B^p)^{\frac{1}{(p-t)s+t}}$, that is, $(A \geq) B \geq C$. By $(**)$ for $\frac{s+1}{s} \in [1,2]$, we obtain

$$\begin{aligned} C &\geq A^{-r+t} \#_{\frac{1-t+r}{((p-t)s+t-t)(\frac{s+1}{s})+r}} (A^t \natural_{\frac{s+1}{s}} C^{(p-t)s+t}) \\ &= A^{-r+t} \#_{\frac{1-t+r}{(p-t)(s+1)+r}} \left(A^t \natural_{\frac{s+1}{s}} (A^t \natural_s B^p)\right) \\ &= A^{-r+t} \#_{\frac{1-t+r}{(p-t)(s+1)+r}} (A^t \natural_{s+1} B^p). \end{aligned}$$

Hence we have

$$A^{-r+t} \#_{\frac{1-t+r}{(p-t)(s+1)+r}} (A^t \natural_{s+1} B^p) \leq C \leq B.$$

□

Remark 3.1 (GFI; $t = 1$) *implies a variant of (GFI) that*

$$\begin{aligned} & A \geq B > 0,\ t \in [0,1] \\ & \Longrightarrow A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) \leq A^t \#_{\frac{1-t}{p-t}} B^p \quad (r \geq t;\ p, s \geq 1) \end{aligned}$$

Here we note: (1) *The case $t = 0$ and $s = 1$ is just*

$$A \geq B > 0 \implies A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq B \quad (p \geq 1,\ r \geq 0). \tag{SF}$$

(2) *The case $t = 1$ and $r = s$ is the Ando-Hiai inequality:*

$$X \#_\alpha Y \leq I_H \implies X^r \#_\alpha Y^r \leq I_H \quad (r \geq 1). \tag{AH}$$

(Replace $X = A^{-1}$, $Y = A^{-\frac{1}{2}} B^p A^{-\frac{1}{2}}$ and $\alpha = \frac{1}{p}$.)

However, it easily follows from (SGF) because

$$A^{-r+t} \#_{\frac{1-t+r}{(p-t)s+r}} (A^t \natural_s B^p) \leq B = B^t \#_{\frac{1-t}{p-t}} B^p \leq A^t \#_{\frac{1-t}{p-t}} B^p$$

under the same condition as in the above.

3.4 The chaotic ordering

We first remark that $\log x$ is operator monotone, i.e. $A \geq B > 0$ implies $\log A \geq \log B$ by (LH) and $\frac{X^p - 1}{p} \to \log X$ for $X > 0$. By this fact, we can introduce the chaotic order as $\log A \geq \log B$ among positive invertible operators, which is weaker than the usual order $A \geq B$. In this section, we consider the Furuta inequality under the chaotic ordering.

Theorem 3.15 *The following assertions are mutually equivalent for $A,\ B > 0$:*

(i) $A \gg B$, *i.e.* $\log A \geq \log B$,

(ii) $A^p \geq (A^{\frac{p}{2}} B^p A^{\frac{p}{2}})^{\frac{1}{2}}$ *for* $p \geq 0$,

(iii) $A^r \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{r}{p+r}}$ *for* $p,\ r \geq 0$.

Proof.
(i) $\Longrightarrow$ (iii): First we note that $(I_H + \frac{\log X}{n})^n \ \to\ X$ for $X > 0$. Since

$$A_n = I_H + \frac{\log A}{n} \geq B_n = I_H + \frac{\log B}{n} > 0$$

for sufficiently large n, the Furuta inequality ensures that for given $p,\ r > 0$

$${A_n}^{1+nr} \geq ({A_n}^{\frac{nr}{2}} {B_n}^{np} {A_n}^{\frac{nr}{2}})^{\frac{1+nr}{n(p+r)}},$$

or equivalently

$${A_n}^{n(\frac{1}{n}+r)} \geq ({A_n}^{n\frac{r}{2}} {B_n}^{np} {A_n}^{n\frac{r}{2}})^{\frac{1}{n(p+r)}+\frac{r}{p+r}}.$$

Taking $n \to \infty$, we have the desired inequality (iii).
(iii) $\Longrightarrow$ (ii) is trivial by setting $r = p$.
(ii) $\Longrightarrow$ (i): Note that $\frac{X^p - I_H}{p} \to \log X$ for $X > 0$. The assumption (ii) implies that

$$\frac{A^p - I_H}{p} \geq \frac{(A^{\frac{p}{2}} B^p A^{\frac{p}{2}})^{\frac{1}{2}} - I_H}{p} = \frac{A^{\frac{p}{2}} B^p A^{\frac{p}{2}} - I_H}{p\left((A^{\frac{p}{2}} B^p A^{\frac{p}{2}})^{\frac{1}{2}} + I_H\right)} = \frac{A^{\frac{p}{2}} (B^p - 1) A^{\frac{p}{2}} + A^p - I_H}{p\left((A^{\frac{p}{2}} B^p A^{\frac{p}{2}})^{\frac{1}{2}} + I_H\right)}.$$

Taking $p \to +0$, we have

$$\log A \geq \frac{\log B + \log A}{2}, \text{ that is, } \log A \geq \log B.$$

So the proof is complete. □

Remark 3.2 *The order preserving operator inequality* (i) $\Longrightarrow$ (iii) *in above is called the chaotic Furuta inequality, simply (CFI). Here we note that* (iii) $\Longrightarrow$ (i) *is directly proved as follows:*

Take the logarithm on both side of (iii), *that is,*

$$r \log A \geq \frac{r}{p+r} \log A^{\frac{r}{2}} B^p A^{\frac{r}{2}}$$

for $p,\ r \geq 0$. *Therefore we have*

$$\log A \geq \frac{1}{p+r} \log A^{\frac{r}{2}} B^p A^{\frac{r}{2}}.$$

So we put $r = 0$ *in above. Namely it implies that*

$$\log A \geq \frac{1}{p} \log B^p = \log B.$$

As in the chaotic Furuta inequality, Theorem 3.13 has the following chaotic order version:

Theorem 3.16 *If* $\log A \geq \log B$ *for* $A, B > 0$, *then*

$$A^{\frac{(p+t)s+r}{q}} \geq \left(A^{\frac{r}{2}}(A^{\frac{t}{2}}B^pA^{\frac{t}{2}})^sA^{\frac{r}{2}}\right)^{\frac{1}{q}}$$

holds for all $p,\ t,\ s,\ r \geq 0$ *and* $q \geq 1$ *with* $(t+r)q \geq (p+t)s+r$.

Proof. As in the proof of the chaotic Furuta inequality (i) $\Longrightarrow$ (iii), we have

$$A_n = I_H + \frac{\log A}{n} \geq B_n = I_H + \frac{\log B}{n} > 0$$

for sufficiently large n. Thus Theorem 3.13 implies that

$$A_n^{\frac{(p_1+t_1)s+r_1}{q}} \geq \left(A_n^{\frac{r_1}{2}}(A_n^{\frac{t_1}{2}}B_n^{p_1}A_n^{\frac{t_1}{2}})^sA_n^{\frac{r_1}{2}}\right)^{\frac{1}{q}}$$

holds for all $p_1,\ t_1,\ s,\ r_1 \geq 0$ and $q \geq 1$ with $(t_1+r_1)q \geq (p_1+t_1)s+r_1$. Putting $p_1 = np$, $t_1 = nt$ and $r_1 = nr$, we have

$$A_n^{\frac{n((p+t)s+r}{q}} \geq \left(A_n^{\frac{nr}{2}}(A_n^{\frac{nt}{2}}B_n^{np}A_n^{\frac{nt}{2}})^sA_n^{\frac{nr}{2}}\right)^{\frac{1}{q}}$$

for all $p,\ t,\ s,\ r \geq 0$ and $q \geq 1$ with $(t+r)q \geq (p+t)s+r$. Finally, since $A_n{}^n \to A$ and $B_n{}^n \to B$, we have the desired inequality by tending $n \to \infty$. □

The chaotic Furuta inequality (CFI), Theorem 3.15 (iii), is expressed in terms of the weighted geometric mean as well as the Furuta inequality (FI) as follows:

$$A \geq B > 0 \implies A^{-r} \#_{\frac{r}{p+r}} B^p \leq I_H \tag{CFI}$$

holds for $p \geq 0$ and $r \geq 0$.

For the sake of convenience, we cite (AH): For $\alpha \in (0,1)$

$$A \#_\alpha B \leq I_H \implies A^r \#_\alpha B^r \leq I_H \tag{AH}$$

holds for $r \geq 1$.

Theorem 3.17 *The operator inequalities* (FI), (CFI) *and* (AH) *are mutually equivalent.*

Proof.
(CFI) $\Longrightarrow$ (FI): Suppose that (CFI) holds. Then we prove (FI), so we assume that $A \geq B > 0$. We have

$$\begin{aligned} A^{-r} \#_{\frac{1+r}{p+r}} B^p &= B^p \#_{\frac{p-1}{p+r}} A^{-r} = B^p \#_{\frac{p-1}{p}} (B^p \#_{\frac{p}{p+r}} A^{-r}) \\ &= B^p \#_{\frac{p-1}{p}} (A^{-r} \#_{\frac{r}{p+r}} B^p) \leq B^p \#_{\frac{p-1}{p}} I_H = B \leq A, \end{aligned}$$

which means that (FI) is shown.

(FI) $\Longrightarrow$ (AH): Suppose that (FI) holds. Then we prove (AH), so we assume that $A \#_\alpha B \le I_H$ and $r \ge 0$. Then, putting $C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$ and $p = \frac{1}{\alpha} > 1$, we have

$$B_1 = (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^\alpha = C^{\frac{1}{p}} \le A^{-1} = A_1.$$

Applying (FI) to $A_1 \ge B_1$, it follows that for $p \ge 1$,

$$A_1^{-r} \#_{\frac{1+r}{p+r}} B_1^p \le B_1 \le A_1.$$

Summing up the above discussion, for each $p > 1$,

$$A \#_{\frac{1}{p}} B \le I_H \Longrightarrow A^r \#_{\frac{1+r}{p+r}} A^{-\frac{1}{2}} B A^{-\frac{1}{2}} \le A^{-1}, \text{ or } A^{r+1} \#_{\frac{1+r}{p+r}} B \le I_H \text{ for } r \ge 0.$$

Note that

$$B \#_{\frac{p-1}{p+r}} A^{r+1} = A^{r+1} \#_{\frac{1+r}{p+r}} B \le I_H$$

holds. That is, we can assume this and so apply it for $q = \frac{p+r}{p-1} \ge 1$. Hence it implies that

$$I_H \ge B^{r+1} \#_{\frac{1+r}{q+r}} A^{r+1}.$$

Since $1 - \frac{1+r}{p_1+r} = \frac{1}{p}$,

$$I_H \ge B^{r+1} \#_{\frac{1+r}{q+r}} A^{r+1} = A^{r+1} \#_{\frac{1}{p}} B^{r+1}.$$

Namely we obtain (AH).

(AH) $\Longrightarrow$ (CFI): Suppose that (AH) holds. Then we prove (CFI), so we assume that $A \ge B > 0$ and $p, r > 1$ because it holds for $0 \le p, r \le 1$ by (LH). For given $p, r > 1$, we put $\alpha = \frac{r}{p+r}$ and $r_1 = \frac{r}{p}$. Then we have

$$A^{-r_1} \#_{\frac{r_1}{1+r_1}} B \le A^{-r_1} \#_{\frac{r_1}{1+r_1}} A = I_H.$$

Here we apply (AH) to this and so we have

$$I_H \ge A^{-r_1 p} \#_{\frac{r_1 p}{p+r_1 p}} B^p = A^{-r} \#_{\frac{r}{p+r}} B^p,$$

as desired. □

Here we present an interesting characterization of the chaotic ordering.

Theorem 3.18 *The following assertions are mutually equivalent for $A, B > 0$:*

(i) $\log A \ge \log B$,

(ii) *For each $\delta > 0$ there exists an $\alpha = \alpha_\delta > 0$ such that $(e^\delta A)^\alpha > B^\alpha$.*

The proof of Theorem 3.18 is not given here, but its essence is shown as follows:

Theorem 3.19 *If $\log A > \log B$ for A, $B > 0$, then there exists an $\alpha > 0$ such that $A^\alpha > B^\alpha$.*

Proof. Since $\log A - \log B \geq 2s > 0$ for some $s > 0$, there exists an $\alpha > 0$ such that

$$\left\| \frac{x^h - 1}{h} - \log x \right\|_J < s$$

for $0 < h \leq \alpha$, where J is a bounded interval including the spectra of A and B. Hence we have

$$0 \leq \frac{A^\alpha - I_H}{\alpha} - \log A \leq s, \;\; 0 \leq \frac{B^\alpha - I_H}{\alpha} - \log B \leq s,$$

so

$$\begin{aligned} \frac{A^\alpha - B^\alpha}{\alpha} &= \left(\frac{A^\alpha - I_H}{\alpha} - \log A \right) + \log A - \log B - \left(\frac{B^\alpha - I_H}{\alpha} - \log B \right) \\ &\geq \log A - \log B - \left(\frac{B^\alpha - I_H}{\alpha} - \log B \right) \\ &\geq \log A - \log B - \left\| \frac{B^\alpha - I_H}{\alpha} - \log B \right\|_J \\ &\geq 2s - s = s, \end{aligned}$$

that is $A^\alpha - B^\alpha \geq \alpha s > 0$ is shown. □

Related to this, there raises the problem: Does $\log A \geq \log B$ imply that there exists an $\alpha > 0$ such that $A^\alpha \geq B^\alpha$?

Example 3.1 *Take A and B as follows:*

$$A = U \begin{pmatrix} e^4 & 0 \\ 0 & e^{-1} \end{pmatrix} U, \quad U = \frac{1}{\sqrt{5}} \begin{pmatrix} \sqrt{3} & \sqrt{2} \\ \sqrt{2} & -\sqrt{3} \end{pmatrix} \quad \textit{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & e^{-2} \end{pmatrix}.$$

Then we have

$$\log A = \begin{pmatrix} \sqrt{2} & \sqrt{6} \\ \sqrt{6} & 1 \end{pmatrix} \quad \textit{and} \quad \log B = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix},$$

so that $\log A \geq \log B$ *is easily checked.*

On the other hand, putting $x = e^\alpha$ *for* $\alpha > 0$,

$$\det(A^\alpha - B^\alpha) = -x^{-3}(x+1)(x-1)^4(2x^2 + x + 2) < 0$$

for all $x > 1$. *Hence* $A^\alpha \geq B^\alpha$ *does not hold for any* $\alpha > 0$.

Concluding this section, we mention some operator inequalities related to (CFI).

Theorem 3.20 *Let A and B be positive invertible operators. Then the following statements are mutually equivalent:*

(1) $\log A \leq \log B$.

(2) $A^{-r} \#_{\frac{r}{p+r}} B^p \geq I_H$ *for* $p, \; r \geq 0$.

(3) $A^{-r} \#_{\frac{\delta+r}{p+r}} B^p \geq B^\delta$ *for* $p,\ r \geq 0$ *and* $0 \leq \delta \leq p$.

(4) *The operator function* $f(p) = A^{-r} \#_{\frac{r}{p+r}} B^p$ *is increasing on* p.

Proof.
(1) $\Longleftrightarrow$ (2): It follows from (i) $\Longleftrightarrow$ (iii) in Theorem 3.15.
(2) $\Longrightarrow$ (3): By using Lemma 3.2, we have

$$\begin{aligned} A^{-r} \#_{\frac{\delta+r}{p+r}} B^p = B^p \#_{\frac{p-\delta}{p+r}} A^{-r} = B^p \#_{\frac{p-\delta}{p}} (B^p \#_{\frac{p}{p+r}} A^{-r}) \\ = B^p \#_{\frac{p-\delta}{p}} (A^{-r} \#_{\frac{r}{p+r}} B^p) \leq B^p \#_{\frac{p-\delta}{p}} I_H = B^\delta. \end{aligned}$$

(3) $\Longrightarrow$ (2): It is trivial by putting $\delta = 0$.
(3) $\Longrightarrow$ (4): By using (iv) of Lemma 3.2, we have

$$\begin{aligned} f(p+\varepsilon) &= A^{-r} \#_{\frac{r}{p+\varepsilon+r}} B^{p+\varepsilon} \\ &= A^{-r} \#_{\frac{r}{p+r}} (A^{-r} \#_{\frac{p+r}{p+\varepsilon+r}} B^{p+\varepsilon}) \\ &\geq A^{-r} \#_{\frac{r}{p+r}} B^p = f(p). \end{aligned}$$

(4) $\Longrightarrow$ (2): It is obtained by $f(p) \geq f(0) = 1$. □

3.5 The chaotically geometric mean

We consider the monotonicity of the operator function for a fixed $\mu \in [0,1]$ and $A, B > 0$ defined by

$$F(s) = ((1-\mu)A^s + \mu B^s)^{\frac{1}{s}} \quad \text{for } s \in \mathbb{R}.$$

Lemma 3.5 *Let* $F(s)$ *be as in above for a fixed* $\mu \in [0,1]$ *and* $A, B > 0$. *Then*

(1) $F(s)$ *is monotone increasing on* $[1,\infty)$ *and not so on* $(0,1]$ *under the usual order.*

(2) $F(s)$ *is monotone increasing on* $\mathbb{R}$ *under the chaotic order. Consequently there exists* $F(0) = s - \lim_{h \to 0} F(h)$ *and*

$$F(0) = \mathrm{e}^{(1-\mu)\log A + \mu \log B}.$$

We call it the chaotically μ-geometric mean for $A,\ B > 0$ and denote it by $A\diamondsuit_\mu B$, so

$$A\diamondsuit_\mu B := \mathrm{e}^{(1-\mu)\log A+\mu\log B}.$$

Proof of Lemma 3.5. We first note that the function $x \mapsto x^r$ is operator concave for $r \in [0,1]$.

(1) If $t \geq s \geq 1$, then $r = \frac{s}{t} \in (0,1]$ and so

$$((1-\mu)A^t + \mu B^t)^{\frac{s}{t}} \geq (1-\mu)A^s + \mu B^s.$$

Hence (LH) for $\frac{1}{s}$ implies that $F(t) \geq F(s)$.

Next a counterexample to the latter for $\mu = \frac{1}{2}$ is given by

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^3, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}^3.$$

Then we have

$$F(1) = \frac{1}{2}(A+B) = \begin{pmatrix} 14 & 14 \\ 14 & 20 \end{pmatrix}$$

and

$$F\left(\frac{1}{3}\right) = \left(\frac{1}{2}(A^{\frac{1}{3}} + B^{\frac{1}{3}})\right)^3 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^3 = \begin{pmatrix} 14 & 13 \\ 13 & 14 \end{pmatrix},$$

so that

$$F(1) - F\left(\frac{1}{3}\right) = \begin{pmatrix} 0 & 1 \\ 1 & 6 \end{pmatrix} \not\geq 0.$$

(2) We show that $\log F(s) \leq \log F(t)$ for $s < t$ with $s,t \neq 0$. We first assume that $0 < s < t$. Since x^r is operator concave for $r \in [0,1]$ and $\log x$ is operator monotone on $(0,\infty)$, it follows that

$$\log((1-\mu)A^t + \mu B^t)^{\frac{s}{t}} \geq \log((1-\mu)A^s + \mu B^s),$$

so that

$$\log F(t) \geq \log F(s).$$

The case $s < t < 0$ is similar to the above.

We now prove the second assertion. It follows from the concavity of $\log x$ and the Krein inequality that

$$(1-\mu)\log A + \mu\log B = \frac{1}{t}((1-\mu)\log A^t + \mu\log B^t)$$

$$\leq \frac{1}{t}\log((1-\mu)A^t + \mu B^t) \leq \frac{1}{t}((1-\mu)A^t + \mu B^t)$$

$$= (1-\mu)\frac{A^t - I_H}{t} + \mu\frac{B^t - I_H}{t} \Longrightarrow (1-\mu)\log A + \mu\log B \quad (t \to +0).$$

Moreover it follows that for $t = -s < 0$

$$F_{A,B}(t) = F_{A^{-1},B^{-1}}(-s)^{-1} \Longrightarrow \left[e^{(1-\mu)\log A^{-1}+\mu\log B^{-1}}\right]^{-1} = e^{(1-\mu)\log A+\mu\log B}.$$

Therefore there exists the limit $s - \lim_{t\to 0} F(t)$ and it is $F(0) = e^{(1-\mu)\log A+\mu\log B}$.

Consequently we obtain that if $s < 0 < t$, then

$$\log F(s) \le \log F(0) \le \log F(t),$$

and that $F(s)$ is monotone increasing on $\mathbb{R}$ under the chaotic order. □

Remark 3.3 (1) *On the other hand, we note that x^r is operator convex for $r \in [1,2]$. So, if $0 < s \le t \le 2s$ and $t \ge 1$, then $F(s) \le F(t)$. For example, we have $F(s) \le F(1)$ for $\frac{1}{2} \le s \le 1$.*

(2) *It is proved that $F(s)$ converges to $A \diamondsuit_\mu B$ unifromly.*

We recall that μ-arithmetic mean and μ-harmonic mean are denoted by $A \nabla_\mu B = (1-\mu)A + \mu B$ and $A \,!_\mu\, B = ((1-\mu)A^{-1} + \mu B^{-1})^{-1}$, respectively.

Theorem 3.21 *Let $A, B > 0$ and $\mu \in [0,1]$. Then both $(A^t \nabla_\mu B^t)^{\frac{1}{t}}$ and $(A^t \,!_\mu\, B^t)^{\frac{1}{t}}$ converge to $A\diamondsuit_\mu B$ as $t \to +0$. Consequently*

$$s - \lim_{t\to+0} (A^t \#_\mu B^t)^{\frac{1}{t}} = A\diamondsuit_\mu B.$$

Proof. The first assertion follows from Lemma 3.5 and the second one does from the well-known fact that

$$A^t \,!_\mu\, B^t \le A^t \#_\mu B^t \le A^t \nabla_\mu B^t.$$

□

Remark 3.4 *Theorem 3.21 is closely related to the Golden-Thompson inequality*

$$\|e^{H+K}\| \le \|e^H e^K\| \quad \textit{for self-adjoint} \quad H,\ K$$

and its complementary inequality

$$\|(e^{pH} \#_\mu e^{pK})^{\frac{1}{p}}\| \le \|e^{(1-\mu)H+\mu K}\|$$

for self-adjoint H, K, $p > 0$ and $\mu \in [0,1]$.

As an application of the chaotically geometric mean, we have three operator version of the Furuta inequality.

Theorem 3.22 *Let $A, B, C > 0$ and $\mu \in [0,1]$. Then the following statements are mutually equivalent:*

(1) $\log A \le \log(B\diamondsuit_\mu C)$,

(2) $B^s \nabla_\mu C^s \le A^{-r} \#_{\frac{s+r}{t+r}} (B^t \nabla_\mu C^t)$ *for* $r \ge 0$ *and* $t \ge s \ge 0$.

(3) *For each* $r,\ s \ge 0$, $f(t) = A^{-r} \#_{\frac{s+r}{t+r}} (B^t \nabla_\mu C^t)$ *is an increasing function of* $t \ge s$.

Proof.
(1) $\Longrightarrow$ (2): We note that (1) is equivalent to $\log A \le \log(B^t \nabla C^t)^{\frac{1}{t}}$ for $t > 0$ by the preceding theorem. Therefore (2) follows from Theorem 3.20.
(2) $\Longrightarrow$ (3): Suppose that (2) holds. By Theorem 3.20 again, we have

$$A^{-r} \#_{\frac{t+r}{t+\varepsilon+r}} (B^{t+\varepsilon} \nabla_\mu C^{t+\varepsilon}) \ge (B^{t+\varepsilon} \nabla_\mu C^{t+\varepsilon})^{\frac{t}{t+\varepsilon}},$$

so that

$$\begin{aligned} f(t+\varepsilon) &= A^{-r} \#_{\frac{s+r}{t+r}} (A^{-r} \#_{\frac{t+r}{t+\varepsilon+r}} (B^{t+\varepsilon} \nabla_\mu C^{t+\varepsilon})) \\ &\ge A^{-r} \#_{\frac{s+r}{t+r}} (B^{t+\varepsilon} \nabla_\mu C^{t+\varepsilon})^{\frac{t}{t+\varepsilon}} \ge A^{-r} \#_{\frac{s+r}{t+r}} (B^t \nabla_\mu C^t) = f(t), \end{aligned}$$

where the second inequality is ensured by Jensen's inequality for the function $x^{\frac{t}{t+\varepsilon}}$.
(3) $\Longrightarrow$ (1): If (3) holds, then $f(s) \le f(t)$ for $s \le t$. It implies (1) by Theorem 3.20, too. □

The following theorem is a complement of the preceding theorem.

Theorem 3.23 *Let* $A, B, C > 0$ *and* $\mu \in [0,1]$. *Then the following statements are mutually equivalent:*

(1) $\log A \ge \log(B \diamondsuit_\mu C)$,

(2) $B^s\ !_\mu\ C^s \le A^{-r} \#_{\frac{s+r}{t+r}} (B^t\ !_\mu\ C^t)$ *for* $r \ge 0$ *and* $t \ge s \ge 0$.

(3) *For each* $r,\ s \ge 0$, $h(t) = A^{-r} \#_{\frac{s+r}{t+r}} (B^t\ !_\mu\ C^t)$ *is a decreasing function of* $t \ge s$.

Proof. We note that (1) is equivalent to $\log A^{-1} \le \log(B^{-1} \diamondsuit_\mu C^{-1})$ and $h(t) = h_{A,B,C}(t) = f_{A^{-1},B^{-1},C^{-1}}(t)^{-1}$. So we have the conclusion by the preceding theorem. □

3.6 Generalized the Bebiano-Lemos-Providência inequalities

It is known that the Löwner-Heinz inequality (LH) is equivalent to the Araki-Cordes inequality (AC):

$$\|A^{\frac{t}{2}} B^t A^{\frac{t}{2}}\| \quad \le \quad \|(A^{\frac{1}{2}} B A^{\frac{1}{2}})^t\| \qquad \text{(AC)}$$

for $0 \le t \le 1$,

As a matter of fact, it is easily proved as follows: Let $t \in (0,1)$ be fixed. Suppose that (AC) holds for t, and that $A \geq B > 0$. Since $A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq I$, we have

$$\|A^{-\frac{t}{2}}B^tA^{-\frac{t}{2}}\| \leq \|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|^t \leq 1,$$

so that $A^{-\frac{t}{2}}B^tA^{-\frac{t}{2}} \leq I$, or $B^t \leq A^t$. Conversely assume that (LH) holds for t, and put $\|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\| = b$. Then $A \geq \frac{B}{b}$ and so $A^t \geq (\frac{B}{b})^t$. Hence it follows that $b^t \geq A^{-\frac{t}{2}}B^tA^{-\frac{t}{2}}$ and

$$\|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|^t = b^t \geq \|A^{-\frac{t}{2}}B^tA^{-\frac{t}{2}}\|.$$

Recently, Bebiano, Lemos and Providência showed the following norm inequality, say the BLP inequality, which is an extension of the Araki-Cordes inequality (AC) in some sense.

Theorem 3.24 (BLP) *If $A,B \geq 0$, then*

$$\|A^{\frac{1+t}{2}}B^tA^{\frac{1+t}{2}}\| \quad \leq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^sA^{\frac{s}{2}})^{\frac{t}{s}}A^{\frac{1}{2}}\| \tag{BLP}$$

for all $s \geq t \geq 0$.

The following operator inequality is corresponding to (BLP):
For $A,B \geq 0$ and $t > 0$,

$$A^s\#_{\frac{t}{s}}B^s \ \leq \ A^{1+s} \text{ for some } s \geq t \quad \Longrightarrow \quad B^t \ \leq \ A^{1+t} \tag{3.5}$$

Here replacing B by $B^{\frac{1+t}{t}}$, and putting $p = \frac{s}{t}(\geq 1)$ in (3.5), it is rewritten as follows.

Theorem 3.25 *For $A,B \geq 0$*

$$A^s\#_{\frac{1}{p}}B^{p+s} \ \leq \ A^{1+s} \textit{ for some } p \geq 1 \textit{ and } s \geq 0 \quad \Longrightarrow \quad B^{1+\frac{s}{p}} \ \leq \ A^{1+\frac{s}{p}}. \tag{3.6}$$

As in (BLP), our base is the Furuta inequality. Nevertheless, (BLP) can be improved by reviewing as an operator inequality expression in Theorem 3.25:

Theorem 3.26 *Let A and B be positive operators and $s \geq 0$. Then*

$$A^s\#_{\frac{1}{p}}B^{p+s} \ \leq \ A^{1+s} \textit{ for some } p \geq 1 \quad \Longrightarrow \quad B^{1+s} \ \leq \ A^{1+s}. \tag{3.7}$$

Proof. We put

$$C = (A^{-\frac{s}{2}}B^{p+s}A^{-\frac{s}{2}})^{\frac{1}{p}}, \quad \text{or} \quad B^{p+s} = A^{\frac{s}{2}}C^pA^{\frac{s}{2}}.$$

Then the assumption says that $A \geq C \geq 0$, and so the Furuta inequality ensures that

$$B^{1+s} = (A^{\frac{s}{2}}C^pA^{\frac{s}{2}})^{\frac{1+s}{p+s}} \leq A^{1+s}.$$

That is, the desired inequality (3.7) is proved. □

Now we have a norm inequality equivalent to (3.7) in Theorem 3.26.

Corollary 3.1 *Let A and B be positive operators. Then*

$$\|A^{\frac{1+s}{2}}B^{1+s}A^{\frac{1+s}{2}}\|^{\frac{p+s}{p(1+s)}} \quad \leq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{p+s}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}}\|$$

for all $p \geq 1$ and $s \geq 0$.

In addition, Theorem 3.26 has the following expression by the Löwner-Heinz inequality.

Corollary 3.2 *Let A and B be positive operators. Then*

$$A^s\#_{\frac{1}{p}}B^{p+s} \leq A^{1+s} \text{ for some } p \geq 1 \text{ and } s \geq 0 \quad \Longrightarrow \quad B^{1+t} \leq A^{1+t}$$

for $t \in [0,s]$, or equivalently

$$\|A^{\frac{1+t}{2}}B^{1+t}A^{\frac{1+t}{2}}\|^{\frac{p+s}{p(1+t)}} \quad \leq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{p+s}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}}\| \tag{3.8}$$

for $p \geq 1$ and $s \geq t \geq 0$.

Remark 3.5 *Replacing B by $B^{\frac{t}{1+t}}$,* (3.8) *is expressed as follows: For $A,B \geq 0$*

$$\|A^{\frac{1+t}{2}}B^{t}A^{\frac{1+t}{2}}\|^{\frac{p+s}{p(1+t)}} \quad \leq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{\frac{t(p+s)}{1+t}}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}}\|$$

for $p \geq 1$ and $s \geq t \geq 0$. Thus if we take $p = \frac{s}{t}$ for $s \geq t \geq 0$, then we have the original BLP inequality (BLP) *because $\frac{p+s}{p(1+t)} = 1$ and $\frac{t(p+s)}{1+t} = s$.*

Next, we approach to (BLP) from the reverse direction. That is,

Theorem 3.27 *The Furuta inequality is equivalent to the following norm inequality:*

$$\|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{\frac{t(p+s)}{1+t}}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}}\| \quad \geq \quad \|A^{\frac{1+t}{2}}B^{t}A^{\frac{1+t}{2}}\|^{\frac{p+s}{p(1+t)}}$$

for $p \geq 1$ and $s \geq t \geq 0$.

Proof. First of all, the proposed norm inequality is rephrased by replacing A to A^{-1} as follows:

$$\|A^{-\frac{1}{2}}(A^{-\frac{s}{2}}B^{\frac{t(p+s)}{1+t}}A^{-\frac{s}{2}})^{\frac{1}{p}}A^{-\frac{1}{2}}\| \geq \|A^{-\frac{1+t}{2}}B^{t}A^{-\frac{1+t}{2}}\|^{\frac{p+s}{p(1+t)}}$$

for $p \geq 1$ and $s \geq t \geq 0$. Moreover, putting

$$C = (A^{-\frac{s}{2}}B^{\frac{t(p+s)}{1+t}}A^{-\frac{s}{2}})^{\frac{1}{p}}, \quad \text{or} \quad B^t = (A^{\frac{s}{2}}C^pA^{\frac{s}{2}})^{\frac{1+t}{p+s}},$$

it is also represented as

$$\|A^{-\frac{1}{2}}CA^{-\frac{1}{2}}\| \geq \|A^{-\frac{1+t}{2}}(A^{\frac{s}{2}}C^pA^{\frac{s}{2}})^{\frac{1+t}{p+s}}A^{-\frac{1+t}{2}}\|^{\frac{p+s}{p(1+t)}}$$

for $p \geq 1$ and $s \geq t \geq 0$.

Hence it suffices to show that it is equivalent to the Furuta inequality, which follows from rewriting the Furuta inequality by the help of the Löwner-Heinz inequality:

$$A \geq C > 0 \implies A^{1+t} \geq (A^{\frac{s}{2}} C^p A^{\frac{s}{2}})^{\frac{1+t}{p+s}} \quad \text{for} \quad p \geq 1 \text{ and } s \geq t \geq 0.$$

The way from Theorem 3.27 to Theorem 3.24 (the BLP inequality) is as follows:
We take $p = \frac{s}{t} \geq 1$ in Theorem 3.27. Then

$$\frac{1+t}{p+s} = \frac{t}{s} \quad \text{and} \quad \frac{p+s}{p(1+t)} = 1.$$

So we have (BLP)

$$\|A^{\frac{1}{2}}(A^{\frac{s}{2}} B^s A^{\frac{s}{2}})^{\frac{t}{s}} A^{\frac{1}{2}}\| \geq \|A^{\frac{1+t}{2}} B^t A^{\frac{1+t}{2}}\| \quad \text{for} \quad s \geq t > 0.$$

□

Now we return to (AC), which is the starting point of (BLP). It is easily seen that the Araki-Cordes inequality

$$\|A^{\frac{t}{2}} B^t A^{\frac{t}{2}}\| \leq \|(A^{\frac{1}{2}} B A^{\frac{1}{2}})^t\| \quad \text{for } 0 \leq t \leq 1$$

is equivalent to the following reverse inequality:

$$\|A^{\frac{t}{2}} B^t A^{\frac{t}{2}}\| \geq \|(A^{\frac{1}{2}} B A^{\frac{1}{2}})^t\| \quad \text{for } t \geq 1.$$

Inspired by this fact, we discuss appropriate conditions for which the reverse order of the BLP inequality holds.

Theorem 3.28 *For $A,\ B > 0$,*

$$\|A^{\frac{1+t}{2}} B^t A^{\frac{1+t}{2}}\| \quad \geq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}} B^s A^{\frac{s}{2}})^{\frac{t}{s}} A^{\frac{1}{2}}\| \tag{3.9}$$

holds for all $t \geq s \geq 1$.

More generally, the reverse inequality of the one in Theorem 3.27, the generalized BLP inequality, is given by the following way.

Theorem 3.29 *Let $A, B \geq 0$ and $0 < p \leq 1$. Then*

$$\|A^{\frac{1+s}{2}} B^{1+s} A^{\frac{1+s}{2}}\|^{\frac{p+s}{p(1+s)}} \quad \geq \quad \|A^{\frac{1}{2}}(A^{\frac{s}{2}} B^{p+s} A^{\frac{s}{2}})^{\frac{1}{p}} A^{\frac{1}{2}}\|$$

for all $s \geq 0$ with $s \geq 1 - 2p$.

To prove it, Kamei's theorem on complement of the Furuta inequality is available:

Theorem K. If $A \geq B > 0$, then for $0 < p \leq \frac{1}{2}$

$$A^t \natural_{\frac{2p-t}{p-t}} B^p \le A^{2p} \quad \text{for} \quad 0 \le t \le p$$

and for $\frac{1}{2} \le p \le 1$

$$A^t \natural_{\frac{1-t}{p-t}} B^p \le A \quad \text{for} \quad 0 \le t \le p.$$

Proof of Theorem 3.29. It suffices to show that

$$B^{1+s} \le A^{-(1+s)} \implies A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{p+s}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}} \le I_H \tag{3.10}$$

for $0 < p \le 1$ and $s \ge 0$ with $s \ge 1-2p$. So we put

$$A_1 = A^{-(1+s)},\ B_1 = B^{1+s}.$$

Then (3.10) is rephrased as

$$A_1 \ge B_1 > 0 \implies A_1^{\frac{s}{1+s}} \natural_{\frac{1}{p}} B_1^{\frac{p+s}{1+s}} \le A_1$$

for $0 < p \le 1$ and $s \ge 0$ with $s \ge 1-2p$. Moreover, if we replace

$$t_1 = \frac{s}{1+s},\ p_1 = \frac{p+s}{1+s},$$

then we have $\frac{1-t_1}{p_1-t_1} = \frac{1}{p}$, and $\frac{1}{2} \le p_1(\le 1)$ if and only if $1-2p \le s$, so that (3.10) has the following equivalent expression:

$$A_1 \ge B_1 > 0 \implies A_1^{t_1} \natural_{\frac{1-t_1}{p_1-t_1}} B_1^{p} \le A_1 \quad \text{for} \quad 0 \le t_1 < p_1.$$

Since $\frac{1}{2} \le p_1 \le 1$, this is ensured by Theorem K due to Kamei. □

Next we show that Theorem 3.28 is obtained as a corollary of Theorem 3.29.

Proof of Theorem 3.28. We put $p = \frac{s}{t}$ for $t \ge s \ge 1$. Then we have $1-2p \le s$ if and only if $\frac{t}{t+2} \le s$. Since $s \ge 1$ is assumed, $\frac{t}{t+2} \le s$ holds for arbitrary $t > 0$, so that Theorem 3.29 is applicable.

Now we take $B = B_1^{\frac{t}{1+t}}$ for a given arbitrary $B_1 \ge 0$, i.e. $B_1 = B^{\frac{1+t}{t}}$. Then the Araki-Cordes inequality and Theorem 3.29 imply that

$$\begin{aligned}\|A^{\frac{1+t}{2}}B_1^tA^{\frac{1+t}{2}}\| &\ge \|A^{\frac{1+s}{2}}B_1^{\frac{t(1+s)}{1+t}}A^{\frac{1+s}{2}}\|^{\frac{1+t}{1+s}} = \|A^{\frac{1+s}{2}}B^{1+s}A^{\frac{1+s}{2}}\|^{\frac{p+s}{p(1+s)}} \\ &\ge \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B^{p+s}A^{\frac{s}{2}})^{\frac{1}{p}}A^{\frac{1}{2}}\| = \|A^{\frac{1}{2}}(A^{\frac{s}{2}}B_1^sA^{\frac{s}{2}})^{\frac{t}{s}}A^{\frac{1}{2}}\|,\end{aligned}$$

which proves (3.9). □

Theorem 3.28 is slightly generalized as follows:

If $A,\ B > 0$ and $r \ge 0$, then

$$\|A^{\frac{r+t}{2}}B^tA^{\frac{r+t}{2}}\| \ge \|A^{\frac{r}{2}}(A^{\frac{s}{2}}B^sA^{\frac{s}{2}})^{\frac{t}{s}}A^{\frac{r}{2}}\| \tag{3.11}$$

for all $t \ge s \ge r$.

It is proved by applying Theorem 3.28 to $A_1 = A^r,\ B_1 = B^r$ and $t_1 = \frac{t}{r},\ s_1 = \frac{s}{r}$.

Finally we consider a converse inequality of the generalized BLP inequality which corresponds to another Kamei's complement: If $A \geq B > 0$, then for $0 < p \leq \frac{1}{2}$

$$A^t \natural_{\frac{2p-t}{p-t}} B^p \leq A^{2p} \quad \text{for} \quad 0 \leq t < p.$$

Theorem 3.30 *Let $A, B \geq 0$ and $0 < p \leq \frac{1}{2}$. Then*

$$\|A^{\frac{1+s}{2}} B^{1+s} A^{\frac{1+s}{2}}\|^{\frac{(2p+s)(p+s)}{p(1+s)}} \geq \|A^{p+\frac{s}{2}} (A^{\frac{s}{2}} B^{p+s} A^{\frac{s}{2}})^{\frac{2p+s}{p}} A^{p+\frac{s}{2}}\| \tag{3.12}$$

for all $0 \leq s \leq 1-2p$.

Proof. The proof is quite similar to that of Theorem 3.29. We put

$$A_1 = A^{-(1+s)},\ B_1 = B^{1+s};\ t_1 = \frac{s}{1+s},\ p_1 = \frac{p+s}{1+s}.$$

Then Theorem K gives

$$A_1 \geq B_1 > 0 \implies A_1^{t_1} \natural_{\frac{2p_1-t_1}{p_1-t_1}} B_1^{p_1} \leq A_1^{2p_1},$$

for $0 \leq t_1 < p_1 \leq \frac{1}{2}$, so that

$$A^{-(1+s)} \geq B^{1+s} \implies A^{-s} \natural_{\frac{2p+s}{p}} B^{p+s} \leq A^{-2(p+s)}$$

for $0 \leq s \leq 1-2p$. Obviously, it implies the desired norm inequality (3.12). □

Remark 3.6 *In Theorem 3.30, if we take $s = 0$, then we obtain the Araki-Cordes inequality*

$$\|A^{\frac{1}{2}} B A^{\frac{1}{2}}\|^{2p} \geq \|A^p B^{2p} A^p\|$$

for $0 \leq p \leq \frac{1}{2}$. Also it appears in (3.11) by taking $r = 0$. Actually we have

$$\|A^{\frac{t}{2}} B^t A^{\frac{t}{2}}\| \geq \|(A^{\frac{s}{2}} B^s A^{\frac{s}{2}})^{\frac{t}{s}}\| = \|(A^{\frac{s}{2}} B^s A^{\frac{s}{2}})\|^{\frac{t}{s}}$$

for $t \geq s > 0$.

3.7 Riccati's equation

The following equation is said to be the algebraic Riccati equation:

$$X^* B^{-1} X - T^* X - X^* T = C \tag{3.13}$$

for positive definite matrices B, C and arbitrary T. The simple case $T = 0$ in (3.13)

$$X^*B^{-1}X = C \tag{3.14}$$

is called Riccati's equation by several authors. It is known that the geometric mean $B \# C$ is the unique positive definite solution of (3.14). We recall Ando's definition of it in terms of operator matrix: for positive operators B, C on a Hilbert space,

$$B \# C = \max\left\{X \geq 0; \begin{pmatrix} B & X \\ X & C \end{pmatrix} \geq 0\right\}. \tag{3.15}$$

If B is invertible, it is expressed by

$$B \# C = B^{\frac{1}{2}}(B^{-\frac{1}{2}}CB^{-\frac{1}{2}})^{\frac{1}{2}}B^{\frac{1}{2}}.$$

We first discuss a relation between solutions of Riccati's equations (3.13) and (3.14), by which solutions of (3.13) can be given. The following lemma says that (3.14) is substantial in a mathematical sense.

Lemma 3.6 *Let B be positive invertible, C positive and T arbitrary operators on a Hilbert space. Then W is a solution of Riccati's equation*

$$W^*B^{-1}W = C + T^*BT$$

if and only if $X = W + BT$ is a solution of the algebraic Riccati equation

$$X^*B^{-1}X - T^*X - X^*T = C.$$

Proof. Put $X = W + BT$. Since

$$X^*B^{-1}X - T^*X - X^*T = W^*B^{-1}W - T^*BT$$

we have the conclusion immediately. □

Next we determine solutions of Riccati's equation (3.7.2):

Lemma 3.7 *Let B be positive invertible and A positive. Then W is a solution of Riccati's equation*

$$W^*B^{-1}W = A$$

if and only if W is in the form of $W = B^{\frac{1}{2}}UA^{\frac{1}{2}}$ for some partial isometry U whose initial space contains $\mathrm{ranA}^{\frac{1}{2}}$.

Proof. If W is a solution, then $\|B^{-\frac{1}{2}}Wx\| = \|A^{\frac{1}{2}}x\|$ for all vectors x. It ensures the existence of a partial isometry U such that $B^{-\frac{1}{2}}W = UA^{\frac{1}{2}}W$, i.e. $W = B^{\frac{1}{2}}UA^{\frac{1}{2}}$. □

Consequently, we have solutions of the algebraic Riccati equation (3.13).

Theorem 3.31 *The solutions of the algebraic Riccati equation* (3.13)

$$X^*B^{-1}X - T^*X - X^*T = C.$$

is given by $X = B^{\frac{1}{2}}U(C+T^*BT)^{\frac{1}{2}} + BT$ *for some partial isometry* U *whose initial space contains* $\mathrm{ran}(C+T^*BT)^{\frac{1}{2}}$.

In addition, the following result due to Trapp [283] is obtained by Lemma 3.6.

Corollary 3.3 *Under the assumption that* BT *is self-adjoint, the self-adjoint solution of the algebraic Riccati equation* (3.13) *is in the form of*

$$X = (T^*BT + C)\ \#\ B + BT.$$

Proof. The uniqueness of solution follows from the fact that $A \# B$ is the unique positive solution of $XB^{-1}X = A$. □

Next we will generalize Riccati's equation. Actually it is realized as the positivity of an operator matrix $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \geq 0$ for given positive operators B and A. Roughly speaking, it is regarded as an operator inequality $W^*B^{-1}W \leq A$. As a matter of fact, it is correct if B is invertible.

Lemma 3.8 *Let* A *be a positive operator. Then*

$$\begin{pmatrix} I_H & X \\ X^* & A \end{pmatrix} \geq 0 \quad \textit{if and only if} \quad A \geq X^*X.$$

Proof. Since

$$\begin{pmatrix} I_H & 0 \\ 0 & A - X^*X \end{pmatrix} = \begin{pmatrix} I_H & 0 \\ -X^* & I_H \end{pmatrix} \begin{pmatrix} I_H & X \\ X^* & A \end{pmatrix} \begin{pmatrix} I_H & -X \\ 0 & I_H \end{pmatrix},$$

it follows that

$$\begin{pmatrix} I_H & X \\ X^* & A \end{pmatrix} \geq 0 \quad \text{if and only if} \quad A \geq X^*X.$$

□

The following majorization theorem is quite useful in the below. For convenience, we cite it.

Theorem 3.32 (Douglas' majorization theorem (DM)) *The following statements are mutually equivalent:*

(i) *ran* $X \subset$ *ran* Y.

(ii) *There exists a constant* $k > 0$ *such that* $XX^* \leq k^2YY^*$.

(iii) *There exists an operator* C *such that* $X = YC$ *(and* $\|C\| \leq k$ *if* (ii) *is assumed).*

Incidentally, the unicity of C in (iii) *is ensured by the conditions that*
(1) $\|C\| = \inf\{k > 0;\ XX^* \le k^2 YY^*\}$, (2) *ker X = ker C,* (3) *ran $C \subset$ ker $Y^{\perp}$.*

Proof. We show (i) $\Longrightarrow$ (iii) only. Since $Y_1|_{\ker Y^{\perp}}$ is a bijection onto ran Y, for each $x \in H$ there exists a vector $y \in \ker B^{\perp}$ with $Xx = Y_1 y$. In other words, we can define a linear operator C on H such that $Cx = y$, i.e. $X = YC$ and ran $C \subset \ker Y^{\perp}$. Finally the boundedness of C is shown by the closed graph theorem; if $\{(x_n, Cx_n)\} \subset G(C)$ satisfies $x_n \to x$ and $Cx_n \to y$ for some $x, y \in H$, then

$$Yy = \lim YCx_n = \lim Xx_n = Ax.$$

Since $Cx_n \in \ker B^{\perp}$ and so $y \in \ker B^{\perp}$, we have $Cx = y$. □

Lemma 3.9 *Let A and B be positive operators. Then*

$$\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \ge 0 \quad \textit{implies} \quad \mathrm{ranW} \subseteq \mathrm{ranB}^{\frac{1}{2}}.$$

and so $X = B^{-\frac{1}{2}} W$ is well-defined as a mapping.

Proof. Let $S = \begin{pmatrix} a & b \\ b^* & d \end{pmatrix}$ be the square root of $R = \begin{pmatrix} B & W \\ W^* & A \end{pmatrix}$. Then

$$R = S^2 = \begin{pmatrix} a^2 + bb^* & ab + bd \\ b^* a + db^* & b^* b + d^2 \end{pmatrix},$$

that is,

$$B = a^2 + bb^* \text{ and } W = ab + bd.$$

Since ran $B^{\frac{1}{2}}$ contains both ran a and ran b by (DM), it contains ran $a+$ ran b. Moreover ran W is contained in ran $a+$ ran b by $W = ab + bd$. □

Theorem 3.33 *Let A and B be positive operators on K and H respectively, and W be an operator from K to H. Then* $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \ge 0$ *if and only if $W = B^{\frac{1}{2}} X$ for some operator X from K to H and $A \ge X^* X$.*

Proof. Suppose that $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \ge 0$. Since $\mathrm{ranW} \subseteq \mathrm{ranB}^{\frac{1}{2}}$ by Lemma 3.9, (DM) says that $W = B^{\frac{1}{2}} X$ for some operator X. Moreover we restrict X by $P_B X = X$, where P_B is the range projection of B. Noting that $y \in \mathrm{ranB}$ if and only if $y = B^{\frac{1}{2}} x$ for some $x \in \mathrm{ranB}^{\frac{1}{2}}$, the assumption implies that

$$\left\langle \begin{pmatrix} P_B & X \\ X^* & A \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix}, \begin{pmatrix} x \\ z \end{pmatrix} \right\rangle \ge 0$$

for all $y \in \mathrm{ranB}$ and $z \in K$. This means that $\begin{pmatrix} P_B & X \\ X^* & A \end{pmatrix} \ge 0$, and so

$$\begin{pmatrix} P_B & 0 \\ 0 & A - X^* X \end{pmatrix} = \begin{pmatrix} I_H & 0 \\ -X^* & A \end{pmatrix} \begin{pmatrix} P_B & X \\ X^* & A \end{pmatrix} \begin{pmatrix} I_H & -X \\ 0 & I_H \end{pmatrix} \ge 0,$$

that is, $A \ge X^* X$, as required. The converse is easily checked. □

The following factorization theorem due to Ando is led by Theorem 3.33.

Theorem 3.34 *Let A and B be positive operators. Then* $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \geq 0$ *if and only if* $W = B^{\frac{1}{2}}VA^{\frac{1}{2}}$ *for some contraction* V.

Proof. Suppose that $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \geq 0$. Then it follows from Theorem 3.7.7 that $W = B^{\frac{1}{2}}X$ for some bounded X satisfying $A \geq X^*X$. Hence we can find a contraction V with $X = VA^{\frac{1}{2}}$ by (DM), so that $W = B^{\frac{1}{2}}VA^{\frac{1}{2}}$ is shown.

The converse is proved by Lemma 3.7.5 as follows:

$$\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} = \begin{pmatrix} B & B^{\frac{1}{2}}VA^{\frac{1}{2}} \\ A^{\frac{1}{2}}V^*B^{\frac{1}{2}} & A \end{pmatrix}$$
$$= \begin{pmatrix} B^{\frac{1}{2}} & 0 \\ 0 & A^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} I_H & V \\ V^* & I_H \end{pmatrix} \begin{pmatrix} B^{\frac{1}{2}} & 0 \\ 0 & A^{\frac{1}{2}} \end{pmatrix} \geq 0.$$

□

Finally we consider the geometric mean and the harmonic one, as an application of the preceding paragraph. The former is defined by (3.15).

If B is invertible, then Theorem 3.33 says that $\begin{pmatrix} B & X \\ X & C \end{pmatrix} \geq 0$ if and only if $C \geq X^*B^{-1}X$.

By the way, we can directly obtain the desired inequality $C \geq X^*B^{-1}X$ by the following identity:

$$\begin{pmatrix} I_H & 0 \\ -X^*B^{-1} & I_H \end{pmatrix} \begin{pmatrix} B & X \\ X^* & C \end{pmatrix} \begin{pmatrix} I_H & -B^{-1}X \\ 0 & I_H \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & C - X^*B^{-1}X \end{pmatrix}.$$

Anyway the maximum is given by

$$B \,\#\, C = B^{\frac{1}{2}}(B^{-\frac{1}{2}}CB^{-\frac{1}{2}})^{\frac{1}{2}}B^{\frac{1}{2}}.$$

Next we review a work of Pedersen and Takesaki [252]. They proved that if B and C are positive operators and B is nonsingular, then there exists a positive solution X of $XBX = C$ if and only if $(B^{\frac{1}{2}}CB^{\frac{1}{2}})^{\frac{1}{2}} \leq kB$ holds for some $k > 0$.

From the viewpoint of Riccati's inequality, we add another equivalent condition to the Pedersen-Takesaki theorem:

Theorem 3.35 *Let B and C be positive operators and B be nonsingular. Then the following statements are mutually equivalent:*

(1) *Riccati's equation* $XBX = C$ *has a positive solution.*

(2) $(B^{\frac{1}{2}}CB^{\frac{1}{2}})^{\frac{1}{2}} \leq kB$ *for some* $k > 0$.

(3) *There exists the minimum of* $\{X \geq 0; C \leq XBX\}$.

(3') *There exists the minimum of* $\left\{ X \geq 0; \begin{pmatrix} I_H & C^{\frac{1}{2}} \\ C^{\frac{1}{2}} & XBX \end{pmatrix} \geq 0 \right\}$.

Proof. We first note that (3) and (3') are equivalent by Lemma 3.8.

Now we suppose (1), i.e. $X_0BX_0 = C$ for some $X_0 \geq 0$. If $X \geq 0$ satisfies $C \leq XBX$, then

$$(B^{\frac{1}{2}}X_0B^{\frac{1}{2}})^2 = B^{\frac{1}{2}}CB^{\frac{1}{2}} \leq (B^{\frac{1}{2}}XB^{\frac{1}{2}})^2$$

and so

$$B^{\frac{1}{2}}X_0B^{\frac{1}{2}} \leq B^{\frac{1}{2}}XB^{\frac{1}{2}}.$$

Since B is nonsingular, we have $X_0 \leq X$, namely (3) is proved.

Next we suppose (3). Since $C \leq XBX$ for some X, we have

$$B^{\frac{1}{2}}CB^{\frac{1}{2}} \leq (B^{\frac{1}{2}}XB^{\frac{1}{2}})^2$$

and so

$$(B^{\frac{1}{2}}CB^{\frac{1}{2}})^{\frac{1}{2}} \leq B^{\frac{1}{2}}XB^{\frac{1}{2}} \leq \|X\|B,$$

which shows (2).

The implication (2) $\Longrightarrow$ (1) has been shown by Pedersen-Takesaki [252] and Nakamoto [233], but we sketch it for convenience. By Douglas' majorization theorem [45], we have

$$(B^{\frac{1}{2}}CB^{\frac{1}{2}})^{\frac{1}{4}} = ZB^{\frac{1}{2}}$$

for some Z, so that

$$(B^{\frac{1}{2}}CB^{\frac{1}{2}})^{\frac{1}{2}} = B^{\frac{1}{2}}Z^*ZB^{\frac{1}{2}} \quad \text{and} \quad B^{\frac{1}{2}}CB^{\frac{1}{2}} = B^{\frac{1}{2}}(Z^*ZBZ^*Z)B^{\frac{1}{2}}.$$

Since B is nonsingular, Z^*Z is a solution of $XBX = C$.

In addition, we consider an operator matrix $M_{B,C}(X) = \begin{pmatrix} I_H & B^{\frac{1}{2}}X \\ XB^{\frac{1}{2}} & C \end{pmatrix}$ for $B, C, X \geq 0$. We know that $M_{B,C}(X) \geq 0$ if and only if $C \geq XBX$ by Lemma 3.8. We remark that there exists the maximum of $\{X \geq 0; M_{B,C}(X) \geq 0\}$ if (1) in Theorem 3.35 holds. As a matter of fact, if $X_0BX_0 = C$ for some $X_0 \geq 0$, then it follows from Lemma 3.8 that

$$X_0BX_0 = C \geq XBX \quad \text{and so} \quad B^{\frac{1}{2}}X_0B^{\frac{1}{2}} \geq B^{\frac{1}{2}}XB^{\frac{1}{2}}$$

for $X \geq 0$ with $M_{B,C}(X) \geq 0$. Finally the nonsingularity of B implies $X_0 \geq X$, as desired. □

On the other hand, the harmonic mean is defined by

$$B \,!\, C = \max \left\{ X \geq 0; \begin{pmatrix} 2B & 0 \\ 0 & 2C \end{pmatrix} \geq \begin{pmatrix} X & X \\ X & X \end{pmatrix} \right\}.$$

To obtain the exact expression of the harmonic mean, we need the following lemma which is more explicit than Theorem 3.33.

Lemma 3.10 *If* $\begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \geq 0$, *then* $X = B^{-\frac{1}{2}}W$ *is bounded and* $A \geq X^*X$.

Proof. For a fixed vector x, we put $x_1 = B^{-\frac{1}{2}}Wx$. Since $B^{\frac{1}{2}}x_1 = Wx$, we may assume $x_1 \in (\ker B^{\frac{1}{2}})^{\perp}$. So it follows that

$$\begin{aligned} \|B^{-\frac{1}{2}}Wx\| &= \sup\{|(Wx,v)|; \|v\| = 1\} \\ &= \sup\{|(B^{-\frac{1}{2}}Wx, B^{\frac{1}{2}}u)|; \|B^{\frac{1}{2}}u\| = 1\} \\ &= \sup\{|(Wx,u)|; (Bu,u) = 1\}. \end{aligned}$$

On the other hand, since

$$\left\langle \begin{pmatrix} B & W \\ W^* & A \end{pmatrix} \begin{pmatrix} u \\ tx \end{pmatrix}, \begin{pmatrix} u \\ tx \end{pmatrix} \right\rangle = |t|^2\langle Ax,x\rangle + 2\mathrm{Re}\, t\langle Wx,u\rangle + \langle Bu,u\rangle \geq 0$$

for all scalars t, we have

$$|\langle Wx,u\rangle|^2 \leq \langle Ax,x\rangle\langle Bu,u\rangle.$$

Hence it follows that

$$\|B^{-\frac{1}{2}}Wx\|^2 = \sup\{|\langle Wx,u\rangle|; \langle Bu,u\rangle = 1\} \leq \langle Ax,x\rangle,$$

which implies that $X = B^{-\frac{1}{2}}W$ is bounded and $A \geq X^*X$. □

Theorem 3.36 *Let* B, C *be positive operators. Then*

$$B\ !\ C = 2(B - [(B+C)^{-\frac{1}{2}}B]^*[(B+C)^{-\frac{1}{2}}B]).$$

In particular, if $B+C$ *is invertible, then*

$$B\ !\ C = 2(B - B(B+C)^{-1}B) = 2B(B+C)^{-1}C.$$

Proof. First of all, the inequality $\begin{pmatrix} 2B & 0 \\ 0 & 2C \end{pmatrix} \geq \begin{pmatrix} X & X \\ X & X \end{pmatrix}$ is equivalent to

$$\begin{pmatrix} 2(B+C) & -2B \\ -2B & 2B-X \end{pmatrix} = \begin{pmatrix} -I_H & I_H \\ I_H & 0 \end{pmatrix} \begin{pmatrix} 2B-X & -X \\ -X & 2C-X \end{pmatrix} \begin{pmatrix} -I_H & I_H \\ I_H & 0 \end{pmatrix} \geq 0.$$

Then it follows from Lemma 3.8 that $D = [2(B+C)]^{-\frac{1}{2}}(-2B)$ is bounded and $D^*D \leq 2B - X$. Therefore we have the explicit expression of $B\ !\ C$ even if both B and C are non-inbertible:

$$B\ !\ C = \max\{X \geq 0; D^*D \leq 2B - X\} = 2B - D^*D.$$

In particular, if $B+C$ is invertible, then

$$B\ !\ C = 2B - D^*D = 2(B - B(B+C)^{-1}B) = 2B(B+C)^{-1}C.$$

□

Incidentally we consider the set

$$\mathscr{F}_E = \{X \in B(H); X^*EX \leq E\}$$

for a projection E, where $B(H)$ is the set of all bounded linear operators on H.

Lemma 3.11 *Let E be a projection. Then*

$$\mathscr{F}_E = \left\{ \begin{pmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{pmatrix} \text{ on } EH \oplus (I_H - E)H;\ \|X_{11}\| \leq 1 \right\}.$$

Proof. If $X \in \mathscr{F}_E$, then $EX^*EXE \leq E$ and so EXE is a contraction. On the other hand, since $(I_H - E)X^*EX(I_H - E) = 0$, we have $EX(I_H - E) = 0$.

Conversely suppose that

$$X = \begin{pmatrix} X_{11} & 0 \\ X_{21} & X_{22} \end{pmatrix} \text{ on } EH \oplus (I_H - E)H \text{ and } \|X_{11}\| \leq 1.$$

Then

$$X^*EX = \begin{pmatrix} X_{11} & 0 \\ 0 & 0 \end{pmatrix} \leq \begin{pmatrix} I_H & 0 \\ 0 & 0 \end{pmatrix} = E.$$

□

Consequently we have the following:

Theorem 3.37 *Let E be a projection. Then*

(1) *A positive operator X belongs to $\mathscr{F}_E$ if and only if $X = X_1 \oplus X_2$ on $EH \oplus (I_H - E)H$ and $X_1 \leq I_H$.*

(2) *A projection F belongs to $\mathscr{F}_E$ if and only if F commutes with E.*

(3) *A projection F satisfies $FEF = E$ if and only if $F \leq E$.*

Proof. (1) follows from the preceding lemma, and (2) from (1). For (3), first suppose that a projection F satisfies $FEF = E$. Then F commutes with E by (2), so that $FE = FEF = E$. The converse is clear. □

3.8 Hua's inequality

Classical Hua's inequality says that

$$\left| \delta - \sum_{k=1}^{n} a_k \right|^2 \geq \frac{\delta^2 \alpha}{\alpha + n} - \alpha \sum_{k=1}^{n} a_k^2$$

for every $\delta, \alpha > 0$ and $a_k \in R$. By putting $b_k = na_k/\delta$, $\beta = \alpha/n$, $\tau(X) = (\operatorname{Tr} X)/n$, the normalized trace, and $B = \operatorname{diag}(b_1, \cdots, b_n)$, it is expressed as the following brief form:

$$|\tau(I-B)|^2 \geq \frac{\beta}{\beta+1} - \beta\tau(B^2)$$

for $\beta > 0$.

On the other hand, Hua gave the determinant inequality as follows:

$$|\det(I-B^*A)|^2 = \det|1-B^*A|^2 \geq \det(I-A^*A)\det(I-B^*B)$$

for contractive matrices A and B.

In this section, we generalize them in a noncommutative field as a good use of the operator geometric mean. For this, we explain Schwarz's inequality for positive mapping between C^*-algebras. A (not necessarily linear) mapping Φ between C*-algebras is called *2-positive* if

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \geq 0 \quad \text{implies} \quad \begin{pmatrix} \Phi(A) & \Phi(B) \\ \Phi(C) & \Phi(D) \end{pmatrix} \geq 0$$

for all operators A, B, C and D in a C^*-algebra. The determinant on matrix algebras is a (non-linear) 2-positive mapping by Theorem 3.34. For a state ϕ, a normalized positive linear functioanl on a C^*-algebra, we have

$$\begin{pmatrix} \phi(A^*A) & \overline{\phi(B^*A)} \\ \phi(B^*A) & \phi(B^*B) \end{pmatrix} = \begin{pmatrix} \phi(A^*A) & \phi(A^*B) \\ \phi(B^*A) & \phi(B^*B) \end{pmatrix} \geq 0 \qquad \text{by} \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B \end{pmatrix} \geq 0.$$

Thus the 2-positivity of arbitrary states is supported by Schwarz's inequality, i.e.

$$|\phi(B^*A)|^2 \leq \phi(A^*A)\phi(B^*B)$$

for operators A and B in a C*-algebra.

Now we mention Schwarz's operator inequality:

Lemma 3.12 *Let Φ be a 2-positive mapping and $\Phi(B^*A) = U|\Phi(B^*A)|$ the polar decomposition. Then,*

$$|\Phi(B^*A)| \leq \Phi(A^*A) \,\#\, U^*\Phi(B^*B)U.$$

Proof. Since

$$\begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B \end{pmatrix} = \begin{pmatrix} A^* & 0 \\ B^* & 0 \end{pmatrix} \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix} \geq 0,$$

the 2-positivity of Φ implies that $\begin{pmatrix} \Phi(A^*A) & \Phi(A^*B) \\ \Phi(B^*A) & \phi(B^*B) \end{pmatrix} \geq 0.$

So we show that if $\begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} \geq 0$ and $X > 0$, then

$$X \,\#\, U^*ZU \geq |Y^*|,$$

where $Y^* = U|Y^*|$ is the polar decomposition of Y^*. Now it follows from the assumption that

$$\begin{pmatrix} I & Y_1 \\ Y_1^* & Z_1 \end{pmatrix} \geq 0, \text{ where } Y_1 = X^{-\frac{1}{2}} Y X^{-\frac{1}{2}} \text{ and } Z_1 = X^{-\frac{1}{2}} Z X^{-\frac{1}{2}}.$$

Therefore Lemma 3.8 ensures that $Z_1 \geq Y_1^* Y_1$, or $Z \geq Y^* X^{-1} Y$. Since $U^* Y^* = |Y^*|$, we have

$$U^* Z U \geq |Y^*| X^{-1} |Y^*|,$$

so that

$$\begin{aligned} X \# U^* Z U &\geq X \# |Y^*| X^{-1} |Y^*| \\ &= X^{\frac{1}{2}} \left(X^{-\frac{1}{2}} |Y^*| X^{-1} |Y^*| X^{-\frac{1}{2}} \right)^{\frac{1}{2}} X^{\frac{1}{2}} \\ &= X^{\frac{1}{2}} (X^{-\frac{1}{2}} |Y^*| X^{-\frac{1}{2}}) X^{\frac{1}{2}} = |Y^*|. \end{aligned}$$

□

The above lemma leads us to an operator inequality for the modulus of operators:

Corollary 3.4 *If $\Phi(X) = U|\Phi(X)|$ is the polar decomposition of $\Phi(X)$ for a 2-positive mapping Φ, then*

$$|\Phi(X)| \leq \Phi(|X|) \# U^* \Phi(|X^*|) U.$$

In particular, if $\Phi = \phi$ is a state, then $|\phi(X)| \leq \sqrt{\phi(|X|)\phi(|X^*|)}$.

Proof. Let $X = V|X|$ be the polar decomposition of X. Since $V|X|V^* = |X^*|$, we have

$$|\Phi(X)| = |\Phi(V|X|^{\frac{1}{2}} |X|^{\frac{1}{2}})| \leq \Phi(|X|) \# U^* \Phi(V|X|V^*) U = \Phi(|X|) \# U^* \Phi(|X^*|) U.$$

□

Theorem 3.38 *Let A and B be operators on a Hilbert space and Φ a contractive 2-positive mapping for a C^*-algebra including A, B and the identity operator. If $\Phi(B^*A) = U|\Phi(B^*A)|$ is the polar decomposition of a normal operator $\Phi(B^*A)$, then*

$$|I - \Phi(B^*A)| \geq I - |\Phi(B^*A)| \geq I - \Phi(A^*A) \# U^* \Phi(B^*B) U.$$

In addition, if A and B are contractions and Φ is linear, then

$$I - \Phi(A^*A) \# U^* \Phi(B^*B) U \geq \Phi(I - A^*A) \# U^* \Phi(I - B^*B) U.$$

Proof. The first inequality follows from the normality of $X = \Phi(B^*A)$, i.e. $|I - X| \geq I - |X|$ and the second from Lemma 3.12. The last inequality does from the subadditivity and the monotonicity of the geometric mean:

$$\begin{aligned} &\Phi(A^*A) \# U^* \Phi(B^*B) U + \Phi(I - A^*A) \# U^* \Phi(I - B^*B) U \\ &\leq \Phi(A^*A + I - A^*A) \# U^* \Phi(B^*B + I - B^*B) U \\ &= \Phi(I) \# U^* \Phi(I) U \leq I \# I = I, \end{aligned}$$

because $\Phi(I) \leq I$ and $U^*U \leq I$. □

Corollary 3.5 *If A and B are contractions and ϕ is a state, then*

$$|\phi(I-B^*A)|^2 \geq (I-|\phi(B^*A)|)^2 \geq \left(I-\sqrt{\phi(A^*A)\phi(B^*B)}\right)^2$$
$$\geq \phi(I-A^*A)\phi(I-B^*B).$$

Proof. We have only to check the last inequality, which is shown by

$$2\sqrt{\phi(A^*A)\phi(B^*B)} \leq \phi(A^*A)+\phi(B^*B),$$

that is, the arithmetic-geometric mean inequality for positive numbers. □

In the remainder, we mention relations among them.

(1) We claim that Corollary 3.5 implies Hua's determinant inequality. To show this, we may assume that

$$\det|I-B^*A|^2 = \prod_k |\langle |I-B^*A|e_k, e_k\rangle|^2$$

for some complete orthonormal base $\{e_k\}$. Noting that $I-A^*A \geq 0$ and $I-B^*B \geq 0$, it follows from Corollary 3.5 that for each e_k

$$|\langle |I-B^*A|e_k, e_k\rangle|^2 \geq \langle (I-A^*A)e_k, e_k\rangle\langle (I-B^*B)e_k, e_k\rangle,$$

so that

$$\det|I-B^*A|^2 = \prod_k |\langle |I-B^*A|e_k, e_k\rangle|^2 \geq \prod_k \langle (I-A^*A)e_k, e_k\rangle\langle (I-B^*B)e_k, e_k\rangle.$$

Since each $\langle He_k, e_k\rangle$ is a diagonal entry of H with respect to the base $\{e_k\}$, we have

$$\prod_k \langle (I-A^*A)e_k, e_k\rangle\langle (I-B^*B)e_k, e_k\rangle \geq \det(I-A^*A)\det(I-B^*B)$$

by the Hadamard theorem, which obtains the determinant inequality.

(2) Hua's inequality follows from Schwarz's inequality for states. As a matter of fact, it is proved by the use of a simpler inequality; $\phi(A)^2 \leq \phi(A^2)$ for $A \geq 0$. We have to show that

$$|\tau(I-B)|^2 \geq \frac{\beta}{\beta+1}I - \beta\tau(B^2).$$

Instead of showing it, we easily checked that

$$|\tau(I-B)|^2 \geq \frac{\beta}{\beta+1}I - \beta\tau(B)^2 \geq \frac{\beta}{\beta+1}I - \beta\tau(B^2).$$

3.9 The Heinz inequality

In this section, we investigate several norm inequalities equivalent to the Heinz inequality.

Theorem 3.39 (HEINZ INEQUALITY (HI)) *Let A and B be positive operators and $t \in [0,1]$. Then*

$$\|AQ + QB\| \geq \|A^tQB^{1-t} + A^{1-t}QB^t\|$$

for arbitrary operators Q.

The case $t = \frac{1}{2}$ in above is expressed by

$$\|P^*PQ + QRR^*\| \geq 2\|PQR\|$$

for arbitrary operators P and Q. Furthermore it is reduced to the following:

$$\|\mathrm{Re}\, QP\| \geq \|PQ\| \quad \text{if } PQ \text{ is self-adjoint.}$$

Adding to some other inequalities, we have the equivalence among them:

Theorem 3.40 *The following inequalities hold and are mutually equivalent:*

(1) (HI)

(2) $\|P^*PQ + QRR^*\| \geq 2\|PQR\|$ *for arbitrary operators P and Q.*

(3) $\|STR^{-1} + S^{-1}TR\| \geq 2\|T\|$ *for invertible self-adjoint S, R and arbitrary T.*

(4) $\|STS^{-1} + S^{-1}TS\| \geq 2\|T\|$ *for invertible self-adjoint S and arbitrary T.*

(5) $\|STS^{-1} + S^{-1}TS\| \geq 2\|T\|$ *for invertible self-adjoint S and self-adjoint T.*

(6) $\|S^{2m+n}TR^{-n} + S^{-n}TR^{2m+n}\| \geq 2\|S^{2m}T + TR^{2m}\|$ *for invertible self-adjoint S, R, arbitrary T and nonnegative integers m, n.*

(7) $\|\mathrm{Re}\, A^2Q\| \geq \|AQA\|$ *for $A \geq 0$ and self-adjoint Q.*

(8) $\|\mathrm{Re}\, QP\| \geq \|PQ\|$ *for arbitrary P, Q whose product PQ is self-adjoint.*

Proof. We prove it by the following implication:

$$(1) \Longrightarrow (6) \Longrightarrow (5) \Longrightarrow (4) \Longrightarrow (3) \Longrightarrow (2) \Longrightarrow (1) \text{ and } (5) \Longrightarrow (8) \Longrightarrow (7) \Longrightarrow (2).$$

(1) $\Longrightarrow$ (6): In (1), we replace A and B by A^{2m+2n} and B^{2m+2n} respectively, and take $t = (2m+n)(2m+2n)^{-1}$. Then we obtain (6).

(6) $\Longrightarrow$ (5): It is trivial by taking $m = 0$ and $n = 1$ in (6).

(5) $\Longrightarrow$ (4): Since $\begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix}$ is self-adjoint, (5) implies that

$$\begin{aligned}\| \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}^{-1} - \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}^{-1} \begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix} \| \\ \geq 2\| \begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix} \| = 2\|T\|.\end{aligned}$$

Since the left hand side in above equals to $\|STS^{-1} + S^{-1}TS\|$, (4) is obtained.
(4) $\Longrightarrow$ (3): We use Berberian's magic. That is,

$$\begin{aligned}\| \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} 0 & T \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix}^{-1} - \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix}^{-1} \begin{pmatrix} 0 & T \\ 0 & 0 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & R \end{pmatrix} \| \\ \geq 2\| \begin{pmatrix} 0 & T \\ T^* & 0 \end{pmatrix} \| = 2\|T\|.\end{aligned}$$

Since the right hand side in above is $\|STR^{-1} + S^{-1}TR\|$, (3) is obtained.
(3) $\Longrightarrow$ (2): We may assume that both P^*P and RR^* are invertible. Then we have

$$\begin{aligned}\|P^*PQ + QRR^*\| &= \||P|(|P|Q|R^*|)|R^*|^{-1} + |P|^{-1}(|P|Q|R^*|)|R^*|\| \\ &\geq 2\||P|Q|R^*|\| = 2\|PQR\|.\end{aligned}$$

(2) $\Longrightarrow$ (1): The proof is an analogy to Pedersen's one for (LH), stated in Section 3.1. We define the operator function on $[0,1]$ by

$$f(t) = \|A^t QB^{1-t} + A^{1-t}QB^t\| \quad \text{for} \quad t \in [0,1].$$

Thus we prove that $I = \{t \in [0,1];\ f(t) \leq f(1)\} = [0,1]$. Since $0, 1 \in I$ and $f(t)$ is norm continuous, it suffices to show that it is a convex function. For given $\alpha, \gamma \in I$ with $\alpha < \gamma$, we put $\beta = (\alpha+\gamma)/2$, i.e. $\alpha = \beta - \varepsilon$ and $\gamma = \beta + \varepsilon$ for $\varepsilon = (\alpha - \gamma)/2$. Then we have

$$\begin{aligned}2f(\beta) &= 2\|A^\beta QB^{1-\beta} + A^{1-\beta}QB^\beta\| \\ &= 2\|A^\varepsilon(A^\alpha QB^{1-\gamma} + A^{1-\gamma}QB^\alpha)B^\varepsilon\| \\ &\leq \|A^{2\varepsilon}(A^\alpha QB^{1-\gamma} + A^{1-\gamma}QB^\alpha) + (A^\alpha QB^{1-\gamma} + A^{1-\gamma}QB^\alpha)B^{2\varepsilon}\| \quad \text{by (2)} \\ &= \|A^\gamma QB^{1-\gamma} + A^{1-\alpha}QB^\alpha + A^\alpha QB^{1-\alpha} + A^{1-\gamma}QB^\gamma\| \\ &\leq f(\alpha) + f(\gamma),\end{aligned}$$

as desired.

Next we show the second: (5) $\Longrightarrow$ (8) $\Longrightarrow$ (7) $\Longrightarrow$ (2).
(5) $\Longrightarrow$ (8): Let $Q = UH$ be the polar decomposition. We may assume that U is unitary (by extending the space) and H is invertible. Then we have

$$2\|\mathrm{Re}\ QP\| = \|UHP + P^*HU^*\| = \|HPU + U^*P^*H\| = 2\|\mathrm{Re}\ HPU\|.$$

Here we apply (5) for $T = PQ$ and S=H;

$$\begin{aligned}2\|PQ\| &\le \|HPQH^{-1} + H^{-1}Q^*P^*H\| = \|HPU + U^*P^*H\| \\ &= \|QP + P^*Q^*\| = 2\|\text{Re } QP\|,\end{aligned}$$

so that we have (8).

(8) $\Longrightarrow$ (7): It is trivial by putting $Q = A$ and $P = AT$.

(7) $\Longrightarrow$ (2): We put

$$T = \begin{pmatrix} P & 0 \\ 0 & R^* \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & Q \\ Q^* & 0 \end{pmatrix},$$

and apply (7) for $A = |T|$ and $Q = S$. Then we have

$$\|T^*TS + ST^*T\| \ge 2\| \, |T|S|T| \, \|.$$

Moreover, since $\| \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix} \| = \|X\|$, we have

$$\|T^*TS + ST^*T\| = \|P^*PQ + QRR^*\|$$

and

$$\| \, |T|S|T| \, \| = \|TST^*\| = \|PQR\|,$$

which imply (2).

Finally we prove (8) on behalf of them.

If PQ is self-adjoint, then the spectrum $\sigma(PQ)$ lies in the real axis and so does $\sigma(QP)$. Since the closed numerical range $W(QP)^-$ contains $\sigma(QP)$, $\sigma(PQ)$ is contained in Re $W(QP)^- = W(\text{Re } QP)^-$, so that the spectral radius $r(QP)$ is not greater than the numerical radius $w(\text{Re } QP)$. Hence we have

$$\|PQ\| = r(PQ) = r(QP) \le w(\text{Re } QP) = \|\text{Re } QP\|.$$

This completes the proof. □

Next we mention several inequalities equivalent to (LH). Among them, the Heinz-Kato inequality is important from the historical view.

Theorem 3.41 (HEINZ-KATO INEQUALITY (HK)) *Let A and B positive operators on H. Then*

$$T^*T \le A^2, \; TT^* \le B^2 \implies |(Tx,y)| \le \|A^s x\| \|B^{1-s}\| \textit{ for } s \in [0,1], \; x,y \in H.$$

Afterwards, it was extended by Furuta:

Theorem 3.42 (HEINZ-KATO-FURUTA INEQUALITY (HKF)) *Let A and B positive operators on H. Then*

$$T^*T \le A^2, \; TT^* \le B^2 \implies |(U|T|^{s+t}x,y)| \le \|A^s x\| \|B^t\|$$

holds for $s,t \in [0,1]$, $x,y \in H$, where $T = U|T|$ is the polar decomposition of T.

Proof. Noting that $U|T|^tU^* = |T^*|^t$ for $t > 0$, it follows that

$$\begin{aligned}|(U|T|^{s+t}x,y)| &= |(|T|^s x, |T|^t U^* y)| \le \||T|^s x\|\||T|^t U^* y\| \\ &= \||T|^s x\|\|U|T|^t U^* y\| = \||T|^s x\|\||T^*|^t y\|.\end{aligned}$$

Since $\||T|^s x\| \le \|A^s x\|$ and $\||T^*|^t y\| \le \|B^t y\|$ by (LH) we have (HKF). □

Theorem 3.43 *The following inequalities are mutually equivalent:*

(1) (LH) *or* (AC),

(2) (HK) *or* (HKF),

(3) $\|A^sTB^{1-s}\| \le \|AT\|^s\|TB\|^{1-s}$ *for* $s \in [0,1]$.

(4) $\|ABA\| \le \|A^2B\|$ *for* $A,B \ge 0$.

(5) $\|TS\| \ge \|ST\|$ *if* ST *is self-adjoint.*

Proof. First of all, we note that a proof of (LH) is written in the below of Theorem 3.1, in which the equivalence (1) $\Longleftrightarrow$ (2) is implicitly explained.
(1) $\Longrightarrow$ (2) is done in the proof of (HKF) in above.
(2) $\Longrightarrow$ (3): (2) says that

$$\|TA\| \le 1,\ \|T^*B\| \le 1 \implies |(Tx,y)| \le \|A^{-s}x\|\|B^{s-1}y\| \quad \text{for } x,y \in H.$$

If we replace x and y by A^sx and $B^{1-s}y$ respectively, then we have

$$|(B^{1-s}TA^sx,y)| \le \|x\|\|y\|,$$

that is, we obtain that if $\|TA\| \le 1$ and $\|T^*B\| \le 1$, then $\|B^{1-s}TA^s\| \le 1$. By the use of Berberian's operator matrix magic, we may assume that T is self-adjoint in (3). Hence we have

$$\|A^sTB^{1-s}\| = \|B^{1-s}TA^s\| \le \|AT\|^s\|TB^{1-s}\|.$$

(3) $\Longrightarrow$ (4): Put $T = B = A$ and $s = \frac{1}{2}$ in (3).
(4) $\Longrightarrow$ (5): Let $T^* = UH$ be the polar decomposition of T^*. Then

$$\begin{aligned}\|ST\|^2 &= \|UHS^*SHU^*\| \le \|HS^*SH\| \\ &\le \|H^2S^*S\| \quad \text{by (4)} \\ &= \|TT^*S^*S\| = \|T(ST)^*S\| = \|TSTS\| \le \|TS\|^2.\end{aligned}$$

(5) $\Longrightarrow$ (1): We show (5) $\Longrightarrow$ (AC). For convenience, (AC; t) holds for $t \in [0,1]$ means that $\|A^tB^t\| \le \|AB\|^t$ holds for all $A,B \ge 0$. We first prove that (5) implies (AC; $\frac{1}{2}$) holds. For this, we put $S = B^{\frac{1}{2}}$ and $T = AB^{\frac{1}{2}}$. Then $ST \ge 0$, it follows from (5) that

$$\|AB\| = \|TS\| \ge \|ST\| = \|B^{\frac{1}{2}}AB^{\frac{1}{2}}\| = \|A^{\frac{1}{2}}B^{\frac{1}{2}}\|^2.$$

Next we show (AC; $\frac{3}{4}$) by (AC; $\frac{1}{2}$).

$$\begin{aligned}\|A^{\frac{3}{4}}B^{\frac{3}{4}}\|^2 &= \|B^{\frac{3}{4}}A^{\frac{3}{2}}B^{\frac{3}{4}}\| = \|B^{\frac{1}{4}}(B^{\frac{1}{2}}A^{\frac{3}{2}}B^{\frac{3}{4}})\| \\ &\le \|B^{\frac{1}{2}}A^{\frac{3}{2}}B\| \le \|B^{\frac{1}{2}}A^{\frac{1}{2}}\|\|AB\| \\ &\le \|AB\|^{\frac{3}{2}}.\end{aligned}$$

In general, if (AC; s) and (AC; t) holds, then so does (AC; $\frac{s+t}{2}$). Assume $s<t$, and put $r=\frac{s+t}{2}$ and $d=\frac{t-s}{2}$. Then

$$\begin{aligned}\|A^rB^r\|^2 &= \|B^rA^{s+t}B^r\| = \|B^d(B^sA^{s+t}B^r)\| \\ &\le \|B^sA^{s+t}B^{r+d}\| \le \|B^sA^s\|\|A^tB^t\| \\ &\le \|AB\|^{s+t} = \|AB\|^{2r}.\end{aligned}$$

Since $\{m/2^n; n=1,2,\dots,m=1,2,\dots,2^n\}$ is dense in $[0,1]$, (AC) is proved under the assumption (5). □

Remark 3.7 *We now mention an interesting relation between (HI) and (LH): We compare Theorem 3.9.1 (8) and Theorem 3.9.2 (5). We pick out them.*

(HI) $\iff$ $\|\mathrm{Re}\,TS\| \ge \|ST\|$ *if* ST *is self-adjoint.*

(LH) $\iff$ $\|TS\| \ge \|ST\|$ *if* ST *is self-adjoint.*

From this, it is obvious that (HI) is stronger than (LH).

Finally we discuss a norm inequality considered in the Corach-Porta-Recht geometry:

Theorem 3.44 (CORACH-PORTA-RECHT INEQUALITY (CPR)) *Let* $A,B,C,D \ge 0$. *Then*

$$\|(A \#_t B)^{\frac{1}{2}}(C \#_t D)^{\frac{1}{2}}\| \le \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{1-t}\|B^{\frac{1}{2}}D^{\frac{1}{2}}\|^t \quad \textit{for } t\in[0,1].$$

Theorem 3.45 *The inequalities* (CPR), (LH) *and Jensen's inequality* (JI)*;*

$$(X^*AX)^t \ge X^*A^tX \quad \textit{for contractions } X, A\ge 0,\ t\in[0,1]$$

are mutually equivalent.

Proof.
(LH) $\Longrightarrow$ (JI): By virtue of the polar decomposition, it suffices to show that

$$(CAC)^t \ge CA^tC \quad \text{for invertible positive contractions } C,$$

or equivalently

$$A^t \le C^{-1}(CAC)^tC^{-1} = C^{-2} \#_t A.$$

Since $C^{-1}\ge I_H$ and so $C^{-2}\ge I_H$, it is ensured by the monotonicity of $\#_t$, namely (LH).
(JI) $\Longrightarrow$ (CPR): First of all, (JI) is explicitly expressed as

$$(X^*AX)^t \le \|X\|^{2-2t}X^*A^tX \quad \text{for arbitrary } X,A\ge 0,\ t\in[0,1].$$

Thus it follows that for $C > 0$

$$\begin{aligned} A \#_t B &= C^{-\frac{1}{2}}(C^{\frac{1}{2}}A^{\frac{1}{2}})(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t(A^{\frac{1}{2}}C^{\frac{1}{2}})C^{-\frac{1}{2}} \\ &\le C^{-\frac{1}{2}}\|C^{\frac{1}{2}}A^{\frac{1}{2}}\|^{2-2t}((C^{\frac{1}{2}}A^{\frac{1}{2}})(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})(A^{\frac{1}{2}}C^{\frac{1}{2}}))^tC^{-\frac{1}{2}} \\ &= \|C^{\frac{1}{2}}A^{\frac{1}{2}}\|^{2-2t}C^{-\frac{1}{2}}(C^{\frac{1}{2}}BC^{\frac{1}{2}})^tC^{-\frac{1}{2}} \\ &= \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{2-2t}C^{-1} \#_t B, \end{aligned}$$

and that for $B > 0$

$$C \#_t D = D \#_{1-t} C \le \|B^{\frac{1}{2}}D^{\frac{1}{2}}\|^{2-2(1-t)}(B^{-1} \#_{1-t} C) = \|B^{\frac{1}{2}}D^{\frac{1}{2}}\|^{2t}(C \#_t B^{-1}).$$

Therefore we have

$$\begin{aligned} \|(A \#_t B)^{\frac{1}{2}}(C \#_t D)^{\frac{1}{2}}\|^2 &= \|(C \#_t D)^{\frac{1}{2}}(A \#_t B)(C \#_t D)^{\frac{1}{2}}\| \\ &\le \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{2-2t}\|(C \#_t D)^{\frac{1}{2}}(C^{-1} \#_t B)\|\|(C \#_t D)^{\frac{1}{2}}\| \\ &= \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{2-2t}\|(C^{-1} \#_t B)^{\frac{1}{2}}(C \#_t D)(C^{-1} \#_t B)^{\frac{1}{2}}\| \\ &\le \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{2-2t}\|B^{\frac{1}{2}}D^{\frac{1}{2}}\|^{2t}\|(C^{-1} \#_t B)^{\frac{1}{2}}(C \#_t B^{-1})(C^{-1} \#_t B)^{\frac{1}{2}}\| \\ &= \|A^{\frac{1}{2}}C^{\frac{1}{2}}\|^{2-2t}\|B^{\frac{1}{2}}D^{\frac{1}{2}}\|^{2t}, \end{aligned}$$

because $C \#_t B^{-1} = (C^{-1} \#_t B)^{-1}$. So we obtain (CPR).

(CPR) $\Longrightarrow$ (LH): Put $A = C = I_H$ in (CPR) and $X = B^{\frac{1}{2}}, Y = D^{\frac{1}{2}}$. Then we have (AC), which is equivalent to (LH). □

3.10 Notes

(LH) was considered in general setting by Löwner [171] and explicitly proved by Heinz [140]. Another proof is given by Kato [157], and interesting proof is presented by Pedersen [251]. A step of the way from (LH) to (FI) was set up by Chan-Kwong [31]. (FI) was established by Furuta [106] in 1987. A simple proof was given by himself [107], and mean theoretic approach was done by [151] and [79]. Among others, Tanahashi [277] considered the best possibility of the exponent in (FI).

(AH) is an essential part of the proof of a majorization inequality in [12]. Its 2 variable version (GAH) was given by [90], and (GAH) is equivalent to (GFI; $t = 1$) by [91].

(GFI) was established in order to discuss (AH) in the flame of (FI). As similar to (FI), the best possibility of (GFI) is obtained in [278] cf. [289, 95], see also [124, Chapter 7].

The chaotic order was introduced in [89]. The Furuta inequality for the chaotic order was essentially initiated by Ando [8]: For self-adjoint operators A and B,

$$A \geq B \iff e^{pA} \geq (e^{\frac{pA}{2}} e^{pB} e^{\frac{pA}{2}})^{\frac{1}{2}} \quad \text{for } p \geq 0,$$

which appears in (i) and (ii) of Theorem 3.4.1, and (iii) is posed in [80], see also [108].

The Furuta inequality induces another geometric mean, so called the chaotically geometric mean $A \diamondsuit_\mu B = e^{(1-\mu)\log A + \mu \log B}$, [98]. It is closely related to the Golden-Thompson inequality:

$$\|e^{H+K}\| \leq \|e^H e^K\| \quad \text{for self-adjoint } H, K.$$

Theorem 3.5.3 was obtained by Hiai-Petz [142].

BLP inequality [19] is a generalization of the Araki-Cordes inequality in some sense. It is discussed from the viewpoint of the difference from (FI). Consequently (BLP) is generalized in [100, 180].

Section 3.7 is written by depending on [59] mainly. The study of Riccati's equation was initiated by Pedersen-Takesaki [252]. In particular, the geometric mean $A \# B$ is the unique self-adjoint solution of $XA^{-1}X = B$ for given $A, B > 0$. The definition of the geometric mean by using operator matrix was introduced by Ando [6]. The algebraic Riccati equation is solved by Trapp [283] under some additional assumption.

Chapter 4

Kantorovich-Furuta Type Inequalities

In this chapter, we study order preserving operator inequalities in another direction which differs from the Furuta inequality. We investigate the Kantorovich-Furuta type inequalities related to the operator ordering and the chaotic one.

4.1 Introduction

Let A and B be positive operators on a Hilbert space H. The Löwner-Heinz theorem asserts that $A \geq B \geq 0$ ensures $A^p \geq B^p$ for all $p \in [0,1]$. However $A \geq B$ does not always ensure $A^p \geq B^p$ for each $p > 1$ in general. In order to study operator inequality, the Löwner-Heinz theorem is very useful, but the fact above is inconvenient, because the condition "$p \in [0,1]$" is too restrictive. Thus, excluding the limit of p, it is the Furuta inequality that devises methods to preserve the order for $p > 1$. Namely, by considering the magic box

$$f(\square) = \left(B^{\frac{r}{2}} \square B^{\frac{r}{2}}\right)^{\frac{1}{q}},$$

then the Furuta inequality asserts that $A \geq B \geq 0$ ensures

$$f(A^p) \geq f(B^p)$$

holds for all $p > 1$ and additional conditions of q and r.

We study order preserving operator inequalities in another direction which differ from the Furuta inequality. First of all, to explain it, we present the following simple example. By virtue of the Kantorovich inequality, a function $f(t) = t^2$ is order preserving in the following sense.

Theorem 4.1 *Let A and B be positive operators. Then*

$$A \geq B \geq 0 \quad \textit{and} \quad MI_H \geq B \geq mI_H > 0 \quad \textit{imply} \quad \frac{(M+m)^2}{4Mm}A^2 \geq B^2.$$

Proof. Refer to [124, Theorem 8.1] for the proof. □

Theorem 4.1 is a new view of operator inequality which differ from the Furuta inequality. Namely, $f(t) = t^2$ preserves the order in terms of the spectrum of given positive operators by virtue of the Kantorovich inequality. Thus, we call it the Kantorovich-Furuta type operator inequality.

By using a generalization of the Kantorovich inequality, we get the following Kantorovich-Furuta type operator inequality.

Theorem 4.2 *Let A and B be positive operators such that $MI_H \geq B \geq mI_H$ for some scalars $M > m > 0$. If $A \geq B$, then*

$$\left(\frac{M}{m}\right)^{p-1} A^p \geq K(m,M,p)A^p \geq B^p \qquad \textit{for all } p \geq 1,$$

where the generalized Kantorovich constant $K(m,M,p)$ is defined by (2.29)*.*

Proof. Refer to [124, Theorem 8.3] for the proof. □

Theorem 4.3 *Let A and B be positive operators such that $M_1 I \geq A \geq m_1 I$ and $M_2 I_H \geq B \geq m_2 I_H$ for some scalars $M_j > m_j > 0$ $(j = 1,2)$. If $A \geq B$, then the following inequalities hold:*

(i) $K(m_j,M_j,p)A^p \geq B^p$ *for all $p > 1$ and $j = 1,2$,*

(ii) $K(m_j,M_j,p)B^p \geq A^p$ *for all $p < -1$ and $j = 1,2$.*

Proof. Refer to [124, p.220,232,250] for the proof. □

For positive invertible operators A and B, the order $A \gg B$ defined by $\log A \geq \log B$ is called the chaotic order. Since $\log t$ is an operator monotone function, the chaotic order is weaker than the operator order $A \geq B$.

The following theorem is a Kantorovich-Furuta type operator inequality related to the chaotic order which is parallel to Theorem 4.2.

Theorem 4.4 *Let A and B be positive invertible operators such that $MI_H \geq B \geq mI_H$ for some scalars $M > m > 0$. If $\log A \geq \log B$, then*

$$\left(\frac{M}{m}\right)^{p} A^p \geq K(m,M,p+1)A^p \geq B^p \qquad \textit{for all } p \geq 0.$$

Proof. Refer to [124, Theorem 8.4] for the proof. □

Remark 4.1 *In fact, the chaotic order* $\log A \geq \log B$ *does not always ensure the operator order* $A \geq B$ *in general. However, by Theorem 4.4, it follows that*

$$\log A \geq \log B \quad \textit{and} \quad MI \geq B \geq mI > 0 \quad \textit{imply} \quad \frac{(M+m)^2}{4Mm} A \geq B.$$

In terms of the Kantorovich constant, we show Kantorovich type operator inequalities related to the operator ordering and the chaotic one:

Theorem 4.5 *Let A and B be positive operators such that* $MI_H \geq B \geq mI_H$ *for some scalars* $0 < m \leq M$. *If* $A \geq B$, *then*

$$\frac{(M^{p-1}+m^{p-1})^2}{4M^{p-1}m^{p-1}} A^p \geq B^p \qquad \textit{for all } p \geq 2.$$

Proof. For each $p \geq 2$, put $r = p-2$ and $q = \frac{p+r}{1+r}$ in the Furuta inequality (FI). Then the Furuta inequality ensures

$$\left(B^{\frac{p-2}{2}} A^p B^{\frac{p-2}{2}}\right)^{\frac{1}{2}} \geq B^{p-1}. \tag{4.1}$$

Square both sides of (4.1), it follows from $M^{p-1}I_H \geq B \geq m^{p-1}I_H$ and Theorem 4.1 that

$$\frac{(M^{p-1}+m^{p-1})^2}{4M^{p-1}m^{p-1}} B^{\frac{p-2}{2}} A^p B^{\frac{p-2}{2}} \geq B^{2(p-1)},$$

and hence

$$\frac{(M^{p-1}+m^{p-1})^2}{4M^{p-1}m^{p-1}} A^p \geq B^p \qquad \text{for all } p \geq 2.$$

□

Theorem 4.6 *Let A and B be positive invertible operators such that* $MI_H \geq B \geq mI_H$ *for some scalars* $M > m > 0$. *Then the following assertions are mutually equivalent:*

(i) $\log A \geq \log B$.

(ii) $\frac{(M^p+m^p)^2}{4M^pm^p} A^p \geq B^p$ *for all* $p \geq 0$.

Proof. Refer to [124, Theorem 8.5] for the proof. □

The exponential function $t \mapsto \exp(t)$ is not operator monotone. By virtue of the Mond-Pečarić method, the exponential function preserves the operator order in the following sense.

Theorem 4.7 *Let A and B be self-adjoint operators such that* $MI_H \geq B \geq mI_H$ *for some scalars* $m \leq M$. *If* $A \geq B$, *then*

$$S(e^{M-m}) \exp A \geq \exp B,$$

where the Specht ratio $S(h)$ *is defined by* (2.35).

Proof. By Theorem 2.15, we have $\langle \exp Bx, x\rangle \le S(e^{M-m}) \exp\langle Bx, x\rangle$ for every unit vector $x \in H$. Hence it follows that

$$\begin{aligned}\langle \exp Bx, x\rangle &\le S(e^{M-m}) \exp\langle Bx, x\rangle \\ &\le S(e^{M-m}) \exp\langle Ax, x\rangle \qquad \text{by the assumption } A \ge B \\ &\le S(e^{M-m}) \langle \exp A\, x, x\rangle \qquad \text{by Jensen's inequality}\end{aligned}$$

for every unit vector $x \in H$, so that we have $S(e^{M-m}) \exp A \ge \exp B$. □

Furthermore, by the property of the Specht ratio, we have the following characterization of the operator ordering.

Theorem 4.8 *Let A and B be positive operators such that $MI_H \ge B \ge mI_H$ for some scalars $0 < m \le M$. Then the following assertions are mutually equivalent:*

(i) $A \ge B$.

(ii) $S(e^{p(M-m)}) \exp(pA) \ge \exp(pB)$ *for all* $p \ge 0$,

where the Specht ratio $S(h)$ is defined by (2.35).

Proof.
Suppose (i): Since $pA \ge pB$ and $pMI_H \ge pB \ge pmI_H$ for all $p > 0$, we have (i) $\Longrightarrow$ (ii) by Theorem 4.7.
Conversely, suppose (ii): Taking the logarithm of both sides of (ii), we have

$$\log S(e^{p(M-m)})^{\frac{1}{p}} + A \ge B.$$

Since $S(e^{p(M-m)})^{\frac{1}{p}} \to 1$ as $p \to 0$ by (iv) of Theorem 2.16, we have $A \ge B$. □

The following theorem is a more precise characterization of the chaotic ordering.

Theorem 4.9 *Let A and B be positive invertible operators such that $MI_H \ge B \ge mI_H$ for some scalars $M > m > 0$. Put $h = \frac{M}{m} (\ge 1)$. Then the following assertions are mutually equivalent:*

(i) $\log A \ge \log B$.

(ii) $S(h^p)A^p \ge B^p$ *for all* $p \ge 0$.

Proof. Refer to [124, Theorem 8.7] for the proof. □

4.2 Difference version

In this section, we show new order preserving operator inequality on the operator order and the chaotic order by estimating the upper bound of the difference.

First of all, we present that the function $t \mapsto t^2$ preserves the operator order in the following sense associated with the difference.

Theorem 4.10 *If A and B are positive operators such that $MI_H \geq B \geq mI_H$ for some scalars $0 < m < M$, then*

$$A \geq B \qquad \textit{implies} \qquad A^2 + \frac{(M-m)^2}{4} I_H \geq B^2. \tag{4.2}$$

Proof. By a difference type of the Kantorovich inequality (Theorem 2.20), we have

$$\begin{aligned}
\langle B^2 x, x\rangle &\leq \langle Bx, x\rangle^2 + \frac{(M-m)^2}{4} && \text{by } MI_H \geq B \geq mI_H \\
&\leq \langle Ax, x\rangle^2 + \frac{(M-m)^2}{4} && \text{by } A \geq B \\
&\leq \langle A^2 x, x\rangle + \frac{(M-m)^2}{4} && \text{by the Hölder-McCarthy inequality}
\end{aligned}$$

for every unit vector $x \in H$. Hence we have (4.2). □

Moreover, we have the following order preserving operator inequality associated with the difference, which is a parallel result to the Kantorovich type inequality in Theorem 4.1.

Theorem 4.11 *If A and B are positive operators such that $MI_H \geq B \geq mI_H$ for some scalars $0, m < M$, then*

$$A^p + M(M^{p-1} - m^{p-1})I_H \geq A^p + C(m, M, p)I_H \geq B^p \quad \textit{for all } p > 1,$$

where the Kantorovich constant for the difference $C(m, M, p)$ is defined by (2.38).

Proof. The former inequality follows from (ii) of Theorem 2.24. The latter follows from

$$\begin{aligned}
\langle B^p x, x\rangle &\leq \langle Bx, x\rangle^p + C(m, M, p) && \text{by Theorem 2.23} \\
&\leq \langle Ax, x\rangle^p + C(m, M, p) && \text{by } A \geq B \\
&\leq \langle A^p x, x\rangle + C(m, M, p) && \text{by the Hölder-McCarthy inequality}
\end{aligned}$$

for every unit vector $x \in H$. □

Theorem 4.12 *Let A and B be positive invertible operators such that $MI_H \geq B \geq mI_H$ for some scalars $0 < m < M$. Put $h = \frac{M}{m} > 1$. Then the following assertions are mutually equivalent:*

(i) $\log A \geq \log B$.

(ii) $A^p + L(m^p, M^p)\log S(h^p)I_H \geq B^p$ *for all* $p > 0$,
where the Specht ratio $S(h)$ *is defined by* (2.35) *and the logarithmic mean* $L(m,M)$ *is defined by* (2.41).

Proof.
(i) $\Longrightarrow$(ii): By using Theorem 2.27 and $M^p I_H \geq B^p \geq m^p I_H$, it follows that

$$\begin{aligned}\langle B^p x, x\rangle &\leq \exp\langle \log B^p x, x\rangle + D(m^p, M^p) \\ &\leq \exp\langle \log A^p x, x\rangle + D(m^p, M^p) \qquad \text{by } \log B \leq \log A \\ &\leq \langle A^p x, x\rangle + D(m^p, M^p) \qquad \text{by convexity of the exp function}\end{aligned}$$

holds for every unit vector $x \in H$. Hence it follows from Lemma 2.4 that

$$A^p + L(m^p, M^p)\log S(h^p)I \geq B^p \qquad \text{for all } p > 0.$$

(ii) $\Longrightarrow$(i): We have

$$\lim_{p\to 0}\frac{1}{p}D(m^p, M^p) = \lim_{p\to 0}\frac{1}{p}L(m^p, M^p)\log S(h^p) = 0$$

by (iv) of Theorem 2.16. Therefore, we have

$$\frac{A^p - I}{p} + \frac{1}{p}D(m^p, M^p)I \geq \frac{B^p - I}{p}$$

and hence $\log A \geq \log B$ as $p \to 0$. □

As an application of the Furuta inequality, we shall show order preserving operator inequality associated with the difference which is parametrized the operator order and the chaotic order.

Let A and B be positive invertible operators on a Hilbert space H. We consider an order $A^\delta \geq B^\delta$ for $\delta \in [0,1]$ which interpolates usual order $A \geq B$ and chaotic order $A \gg B$ continuously. We consider that the case $\delta = 0$ means the chaotic order since $\lim_{\delta\to 0}\frac{A^\delta - I}{\delta} = \log A$ for a positive invertible operator A.

The following lemma shows that the Furuta inequality interpolates the usual order and the chaotic one.

Lemma 4.1 *Let A and B be positive invertible operators. The following statements are mutually equivalent for each* $\delta \in [0,1]$*:*

(i) $A^\delta \geq B^\delta$, *where the case* $\delta = 0$ *means* $A \gg B$.

(ii) $\left(B^{\frac{p}{2}} A^{p+\delta} B^{\frac{p}{2}}\right)^{\frac{p+\delta}{2p+\delta}} \geq B^{p+\delta}$ *for all* $p \geq 0$.

(iii) $\left(B^{\frac{r}{2}} A^{p+\delta} B^{\frac{r}{2}}\right)^{\frac{r+\delta}{p+r+\delta}} \geq B^{r+\delta}$ *for all* $p \geq 0$ *and* $r \geq 0$.

Proof. The case of $0 < \delta \leq 1$ is ensured by the Furuta inequality and the case of $\delta = 0$ by the chaotic Furuta inequality (CFI). □

By virtue of Lemma 4.1, we shall obtain the following order preserving operator inequality associated with the difference.

Theorem 4.13 *Let A and B be positive operators on H satisfying $MI_H \geq B \geq mI_H > 0$. Then the following implication* (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iv) $\Longrightarrow$ (iii) *holds for some* $\delta \in [0,1]$*:*

(i) $A^{\delta} \geq B^{\delta}$, *where the case* $\delta = 0$ *means* $A \gg B$.

(ii) $A^{p+\delta} + \frac{1}{m^{1-\delta}} C(m,M,p+1) I_H \geq B^{p+\delta}$ *for all* $p > 0$.

(iii) $A^{p+\delta} + \frac{(M^p - m^p)^2}{4m^{p-\delta}} I_H \geq B^{p+\delta}$ *for all* $p > \delta$.

(iv) $A^{p+\delta} + \frac{1}{m^r} C(m^{r+\delta}, M^{r+\delta}, \frac{p+r+\delta}{r+\delta}) I_H \geq B^{p+\delta}$ *for all* $p > 0$ *and* $r > 0$,
where $C(m,M,p)$ *is defined by* (2.38).

Proof.
(i)$\Longrightarrow$(iv): It follows from Lemma 4.1 that $A^{\delta} \geq B^{\delta}$ is equivalent to the following inequality:

$$\left(B^{\frac{r}{2}} A^{p+\delta} B^{\frac{r}{2}}\right)^{\frac{r+\delta}{p+r+\delta}} \geq B^{r+\delta} \qquad \text{for all } p > 0 \text{ and } r > 0.$$

Put $A_1 = \left(B^{\frac{r}{2}} A^{p+\delta} B^{\frac{r}{2}}\right)^{\frac{r+\delta}{p+r+\delta}}$ and $B_1 = B^{r+\delta}$, then A_1 and B_1 satisfy $A_1 \geq B_1 > 0$ and $M^{r+\delta} I_H \geq B_1 \geq m^{r+\delta} I_H > 0$. Applying Theorem 4.11 to A_1 and B_1, we have

$$A_1^{\frac{p+r+\delta}{r+\delta}} + C\left(m^{r+\delta}, M^{r+\delta}, \frac{p+r+\delta}{r+\delta}\right) I_H \geq B_1^{\frac{p+r+\delta}{r+\delta}}.$$

Therefore we have

$$B^{\frac{r}{2}} A^{p+\delta} B^{\frac{r}{2}} + C\left(m^{r+\delta}, M^{r+\delta}, \frac{p+r+\delta}{r+\delta}\right) I_H \geq B^{p+r+\delta},$$

so that it follows that

$$A^{p+\delta} + C\left(m^{r+\delta}, M^{r+\delta}, \frac{p+r+\delta}{r+\delta}\right) B^{-r} \geq B^{p+\delta}.$$

(iv)$\Longrightarrow$(ii): Put $r = 1 - \delta > 0$ in (iv).
(iv)$\Longrightarrow$(iii): Put $r = p - \delta > 0$ in (iv). Then we have

$$\frac{1}{m^r} C\left(m^{r+\delta}, M^{r+\delta}, \frac{p+r+\delta}{r+\delta}\right) = \frac{1}{m^{p-\delta}} C(m^p, M^p, 2) = \frac{(M^p - m^p)^2}{4m^{p-\delta}}$$

and $p > \delta$.

(ii)$\Longrightarrow$(i): It follows from (ii) of Theorem 2.13 that

$$\left(\frac{M}{m}\right)^p \geq K(m,M,p+1) \geq 1 \qquad \text{for all } p>0$$

and hence

$$\frac{mM^p - Mm^p}{M-m}\left(\frac{M}{m}-1\right) \geq \frac{mM^p - Mm^p}{M-m}\left(K(m,M,p+1)^{\frac{1}{p}}-1\right) \geq 0.$$

Therefore we have $\lim_{p\to 0} C(m,M,p+1) = 0$.

Hence the proof of Theorem 4.13 is complete. □

Theorem 4.13 interpolates the following two theorems. As a matter of fact, if we put $\delta = 0$ in Theorem 4.13, then we have Theorem 4.14 which make a paraphrase of Theorem 4.12. Also, if we put $\delta = 1$ in Theorem 4.13, then we obtain order preserving operator inequality under the operator order associated with the difference.

Theorem 4.14 *Let A and B be positive invertible operators such that $MI \geq B \geq mI$ for some scalars $0<m<M$. Then the following implication* (i) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (iv) $\Longrightarrow$ (ii) *holds:*

(i) $\log A \geq \log B$.

(ii) $A^p + \frac{1}{m}C(m,M,p+1)I_H \geq B^p \qquad$ *for all* $p>0$.

(iii) $A^p + \frac{(M^p-m^p)^2}{4m^p}I_H \geq B^p \qquad$ *for all* $p>0$.

(iv) $A^p + \frac{1}{m^r}C\left(m^r,M^r,\frac{p+r}{r}\right)I_H \geq B^p \qquad$ *for all* $p>0$ *and* $r>0$,
where $C(m,M,p)$ *is defined by* (2.38).

Proof. If we put $\delta=0$ in Theorem 4.13, then we have the implication (i) $\Longrightarrow$ (iv) $\Longrightarrow$ (iii) and (iv) $\Longrightarrow$ (ii). For (iii) $\Longrightarrow$ (i), since

$$\frac{A^p - I_H}{p} + \frac{1}{p}\frac{(M^p-m^p)^2}{4m^p}I_H \geq \frac{B^p - I_H}{p},$$

we have (i) as $p \to 0$. □

Theorem 4.15 *Let A and B be positive operators such that $MI \geq B \geq mI$ for some scalars $0<m<M$. Then the following implication* (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iv) $\Longrightarrow$ (iii) *holds:*

(i) $A \geq B$.

(ii) $A^p + C(m,M,p)I_H \geq B^p \qquad$ *for all* $p \geq 1$.

(iii) $A^{p+1} + \frac{(M^p-m^p)^2}{4m^{p-1}}I_H \geq B^{p+1} \qquad$ *for all* $p \geq 1$.

(iv) $A^p + \frac{1}{m^{r-1}} C\big(m^r, M^r, \frac{p+r-1}{r}\big) I_H \geq B^p$ *for all* $p \geq 1$ *and* $r \geq 1$,
where $C(m,M,p)$ *is defined by* (2.38).

Remark 4.2 *Theorem 4.13 interpolates Theorem 4.11 and Theorem 4.12. Let A and B be positive invertible operators such that* $MI_H \geq B \geq mI_H$*. Then the following assertions hold:*

(i) $A \geq B$ *implies* $A^p + C(m,M,p)I_H \geq B^p$ *for all* $p \geq 1$.

(ii) $A^\delta \geq B^\delta$ *implies* $A^p + C(m^\delta, M^\delta, \frac{p}{\delta})I_H \geq B^p$ *for all* $p \geq \delta$.

(iii) $\log A \geq \log B$ *implies* $A^p + L(m^p, M^p)\log S(h^p) I_H \geq B^p$ *for all* $p > 0$,
where the Specht ratio $S(h)$ *is defined by* (2.35) *and the logarithmic mean* $L(m,M)$ *is defined by* (2.41).

It follows that the constant of (ii) *interpolates the constant of* (i) *and* (iii) *continuously. In fact, if we put* $\delta = 1$ *in* (ii), *then we have* (i), *also if we put* $\delta \to 0$ *in* (ii), *then we have*

$$\begin{aligned} C(m^\delta, M^\delta, \frac{p}{\delta}) &= \frac{m^\delta M^p - M^\delta m^p}{M^\delta - m^\delta}\{K(m^\delta, M^\delta, \frac{p}{\delta})^{\frac{\delta}{p-\delta}} - 1\} \\ &= \frac{\delta}{h^\delta - 1} m^p (h^p - h^\delta) \frac{K(m^\delta, M^\delta, \frac{p}{\delta})^{\frac{\delta}{p-\delta}} - 1}{\delta} \\ &\to \frac{1}{\log h}(M^p - m^p)\log M_h(p)^{\frac{1}{p}} \quad (as\ \delta \to 0) \\ &= L(m^p, M^p)\log S(h^p), \end{aligned}$$

where $h = \frac{M}{m} > 1$.

4.3 Version with the Specht ratio

In this section, we see that the Specht ratio plays an important rule as characterizations of the chaotic order: Let A and B be positive invertible operators such that $MI_H \geq B \geq mI_H$ for some scalars $0 < m < M$ and $h = \frac{M}{m}$. Then

$$\log A \geq \log B \quad \Longleftrightarrow \quad S_h(p)A^p \geq B^p \qquad \text{for all } p > 0,$$

where the symbol S_h is defined by

$$S_h(p) = S(h^p) \tag{4.3}$$

and the Specht ratio $S(h)$ is defined by (2.35).

It is natural to ask what is characterizations of the operator order in terms of the Specht ratio. Thus, we compare Theorem 4.5 with Theorem 4.6: For $A,B>0$ with $MI_H \geq B \geq mI_H$

$$A \geq B \quad \Longrightarrow \quad \frac{(M^{p-1}+m^{p-1})^2}{4M^{p-1}m^{p-1}}A^p \geq B^p \qquad \text{for all } p \geq 2$$
$$\log A \geq \log B \quad \Longrightarrow \quad \frac{(M^p+m^p)^2}{4M^p m^p}A^p \geq B^p \qquad \text{for all } p > 0.$$

Therefore, we observe the difference between p and $p-1$ in the power of the constant. Hence one might expect that the following result holds under the operator order as a parallel result to Theorem 4.9: Let A and B be positive invertible operators such that $MI \geq B \geq mI$. Then

$$A \geq B \quad \text{implies} \quad S_h(p-1)A^p \geq B^p \quad \text{for all } p \geq 2, \text{ where } h = \frac{M}{m} \geq 1.$$

However, we have a counterexample to this conjecture. Put

$$A = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

then $A \geq B \geq 0$. And $m = \frac{1}{2}$ and $M = 2$, so that $h = \frac{M}{m} = 4$. Then we have $S(h) = 1.26374$ and $S(h^2) = 2.39434$. On the other hand, $\alpha A^2 \geq B^2$ holds if and only if $\alpha \geq 1.27389$, and $\beta A^3 \geq B^3$ holds if and only if $\beta \geq 2.396585$. Therefore $S(h)A^2 \ngeq B^2$ and $S(h^2)A^3 \ngeq B^3$.

Here, we present other characterizations of the chaotic ordering and the operator one associated with Kantorovich type inequalities via the Specht ratio:

Theorem 4.16 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. Then the following assertions are mutually equivalent:*

(i) $\log A \geq \log B$.

(ii) $S_k((p+t)s+r)^q A^{(p+t)s} \geq (A^{\frac{t}{2}}B^pA^{\frac{t}{2}})^s$
holds for $p \geq 0$, $t \geq 0$, $s \geq 0$, $r \geq 0$, $q \geq 1$ with $(t+r)q \geq (p+t)s+r$.

(iii) $S_k(2(p+t)s-2t)^2 A^{(p+t)s} \geq (A^{\frac{t}{2}}B^pA^{\frac{t}{2}})^s$
holds for $p \geq 0$, $t \geq 0$, $s \geq 0$ with $(p+t)s \geq 2t$.

(iv) $S_k(2ps)^{\frac{2}{s}}A^p \geq B^p$ *holds for $p \geq 0$ and $s \geq 1$,*
where $S_k(r)$ is defined by (4.3).

Theorem 4.17 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. Then the*

(i) $A \geq B$.

(ii) $S_k((p-t)s+r)^q A^{(p-t)s} \geq (A^{-\frac{t}{2}} B^p A^{-\frac{t}{2}})^s$
holds for $p \geq 1$, $t \in [0,1]$, $s \geq 1$, $q \geq 1$ *such that* $(1-t+r)q \geq (p-t)s+r$ *and* $r \geq t$.

(iii) $S_k(2(p-t)s-2(1-t))^2 A^{(p-t)s} \geq (A^{-\frac{t}{2}} B^p A^{-\frac{t}{2}})^s$
holds for $p \geq 1$, $t \in [0,1]$, $s \geq 1$ *such that* $(p-t)s \geq 2-t$.

(iv) $S_k(2(p-1)s)^{\frac{2}{s}} A^p \geq B^p$ *holds for* $p \geq 1$, $s \geq 1$ *such that* $p \geq \frac{1}{s}+1$.

(v) $k^{4(p-1)} A^p \geq B^p$ *holds for* $p \geq 1$,
where $S_k(r)$ *is defined by* (4.3).

The following corollary is easily obtained by Theorem 4.17.

Corollary 4.1 *Let A and B be positive invertible operators on a Hilbert space H such that* $kI_H \geq A \geq \frac{1}{k} I_H$ *for some* $k \geq 1$. *If* $A^\delta \geq B^\delta$ *for* $\delta \in (0,1]$, *then*

$$S_k(2(p-\delta)s)^{\frac{2}{s}} A^p \geq B^p$$

holds for $p \geq \delta$, $s \geq 1$ *such that* $p \geq (\frac{1}{s}+1)\delta$, *where* $S_k(r)$ *is defined by* (4.3).

Remark 4.3 *Corollary 4.1 interpolates* (iv) *of Theorem 4.16 and* (iv) *of Theorem 4.17 by means of the Specht ratio. Let A and B be positive invertible operators such that* $kI_H \geq A \geq \frac{1}{k} I_H$ *for some* $k \geq 1$. *Then the following assertions holds:*

(i) $A \geq B$ *implies* $S_k(2(p-1)s)^{\frac{2}{s}} A^p \geq B^p$ *for all* $p \geq \frac{1}{s}+1$ *and* $s \geq 1$,

(ii) $A^\delta \geq B^\delta$ *implies* $S_k(2(p-\delta)s)^{\frac{2}{s}} A^p \geq B^p$ *for all* $p \geq \delta$, $s \geq 1$ *such that* $p \geq (\frac{1}{s}+1)\delta$,

(iii) $\log A \geq \log B$ *implies* $S_k(2ps)^{\frac{2}{s}} A^p \geq B^p$ *for all* $p \geq 0$ *and* $s \geq 1$.

It follows that the Specht ratio of (ii) *interpolates the scalar of* (i) *and* (iii) *continuously. In fact, if we put* $\delta = 1$ *in* (ii), *then we have* (i). *Also, if we put* $\delta \to 0$ *in* (ii), *then we have* (iii).

Moreover, Corollary 4.1 interpolates the following result by means of the Specht ratio:

(i) $A \geq B$ *implies* $k^{4(p-1)} A^p \geq B^p$ *for all* $p \geq 1$,

(ii) $A^\delta \geq B^\delta$ *implies* $S_k(2(p-\delta)s)^{\frac{2}{s}} A^p \geq B^p$ *for all* $p \geq \delta$, $s \geq 1$ *such that* $p \geq (\frac{1}{s}+1)\delta$,

(iii) $\log A \geq \log B$ *implies* $k^{4p} A^p \geq B^p$ *for all* $p \geq 0$.

The Specht ratio of (ii) *interpolates the scalar of* (i) *and* (iii). *In fact, if we put* $\delta = 1$ *and* $s \to \infty$ *in* (ii), *then we have* (i). *Also, if we put* $\delta \to 0$ *and* $s \to \infty$ *in* (ii), *then we have* (iii).

To prove them, we need some preliminaries.

In the following lemma a complementary inequality to the Hölder-McCarthy inequality via the Specht ratio is given.

Lemma 4.2 *Let A be a positive operator such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. Then the following inequalities hold for every unit vector $x \in H$:*

(i) $S_k(1)\langle A^p x, x\rangle \geq \langle Ax, x\rangle^p \, (\geq \langle A^p x, x\rangle)$ *for all* $0 < p < 1$,

(ii) $S_k(p)^p \langle Ax, x\rangle^p \geq \langle A^p x, x\rangle \, (\geq \langle Ax, x\rangle^p)$ *for all* $p \geq 1$.

Proof.

(i): The following converse of Young's inequality is shown in [280]: For a given $a > 0$,

$$S_a(1)a^p \geq pa + (1-p)$$

holds for all $1 > p > 0$. If $k \geq a \geq \frac{1}{k} > 0$, then it follows from (ii) and (iii) of Theorem 2.16 that $S_k(1) = S_{k^{-1}}(1) \geq S_a(1)$. Therefore we have

$$S_k(1)A^p \geq pA + (1-p)I_H \quad \text{for all } 1 > p > 0. \tag{4.4}$$

By (4.4) and Young's inequality, it follows that

$$S_k(1)\langle A^p x, x\rangle \geq p\langle Ax, x\rangle + (1-p) \geq \langle Ax, x\rangle^p$$

holds for every unit vector $x \in H$.

(ii): Next, suppose the case of $p \geq 1$. Replacing p by $1/p$ and A by A^p in (i), then $k^p I_H \geq A^p \geq k^{-p} I_H$ and we have

$$S_{k^p}(1)\langle (A^p)^{1/p} x, x\rangle \geq \langle A^p x, x\rangle^{1/p}.$$

Taking the p-th power on both sides, we have

$$S_k(p)^p \langle Ax, x\rangle^p \geq \langle A^p x, x\rangle.$$

□

The following lemma is a Kantorovich-Furuta type operator inequality via the Specht ratio.

Lemma 4.3 *Let A and B be positive operators such that*

$$\text{(i)} \;\; kI_H \geq A \geq \frac{1}{k}I_H \quad \textit{or} \quad \text{(ii)} \;\; kI_H \geq B \geq \frac{1}{k}I_H$$

for some $k \geq 1$. Then

$$A \geq B \quad \textit{implies} \quad S_k(p)^p A^p \geq B^p \quad \textit{for all } p \geq 1,$$

where $S_k(p)$ is defined by (4.3).

Proof.
Suppose (ii). Then we have

$$\begin{aligned} S_k(p)^p\langle A^p x,x\rangle &\geq S_k(p)^p\langle Ax,x\rangle^p && \text{by Hölder-McCarthy inequality and } p\geq 1\\ &\geq S_k(p)^p\langle Bx,x\rangle^p && \text{by } A\geq B\\ &\geq \langle B^p x,x\rangle && \text{by Lemma 4.2 and } kI_H\geq B\geq (1/k)I_H. \end{aligned}$$

Next, suppose (i). Since $B^{-1}\geq A^{-1}$ and $kI_H\geq A^{-1}\geq \frac{1}{k}I_H$, then it follows from above discussion that $S_k(p)^p B^{-p}\geq A^{-p}$. Hence we have $S_k(p)^p A^p\geq B^p$. □

Proof of Theorem 4.16.
(i) $\Longrightarrow$ (ii): By Theorem 3.16, (i) ensures

$$A^{\frac{(p+t)s+r}{q}}\geq \{A^{\frac{r}{2}}(A^{\frac{t}{2}}B^pA^{\frac{t}{2}})^sA^{\frac{r}{2}}\}^{\frac{1}{q}} \tag{4.5}$$

holds for $p,t,s,r\geq 0$ and $q\geq 1$ with

$$(t+r)q\geq (p+t)s+r. \tag{4.6}$$

Put $A_1=A^{\frac{(p+t)s+r}{q}}$ and $B_1=\{A^{\frac{r}{2}}(A^{\frac{t}{2}}B^pA^{\frac{t}{2}})^sA^{\frac{r}{2}}\}^{\frac{1}{q}}$, then $A_1\geq B_1$ by (4.5) and $kI_H\geq A\geq \frac{1}{k}I_H>0$ assures $k^{\frac{(p+t)s+r}{q}}I_H\geq A_1\geq k^{-\frac{(p+t)s+r}{q}}I_H$. By applying Lemma 4.3 to A_1 and B_1, we have

$$S_{k^{\frac{(p+t)s+r}{q}}}(q)^qA_1^q=S_k((p+t)s+r)^q\,A_1^q\geq B_1^q.$$

Multiplying $A^{-\frac{r}{2}}$ on both sides, we have (ii).
(ii) $\Longrightarrow$ (iii): Put $r=(p+t)s-2t\geq 0$ and $q=2$ in (ii). Then the condition (4.6) is satisfied and $(p+t)s\geq 2t$, so we have (iii).
(iii) $\Longrightarrow$ (iv): If we put $t=0$ in (iii), then we have (iv) by the Löwner-Heinz theorem.
(iv) $\Longrightarrow$ (i): If we put $s=1$ and take logarithm of both sides of (iv), we have

$$\log(S_{k^2}(p)A^p)\geq \log B^p \qquad \text{for all } p>0$$

and hence

$$\log S_{k^2}(p)^{1/p}+\log A\geq \log B \qquad \text{for all } p>0.$$

Then letting $p\to +0$, we have $\log A\geq \log B$ by (iv) of Theorem 2.16. □

Proof of Theorem 4.17.
(i) $\Longrightarrow$ (ii): By the grand Furuta inequality, (i) ensures

$$A^{\frac{(p-t)s+r}{q}}\geq \{A^{\frac{r}{2}}(A^{-\frac{t}{2}}B^pA^{-\frac{t}{2}})^sA^{\frac{r}{2}}\}^{\frac{1}{q}} \tag{4.7}$$

holds for $p\geq 1$, $t\in[0,1]$, $s\geq 1$, $q\geq 1$ and

$$r\geq t, \tag{4.8}$$

$$(1-t+r)q\geq (p-t)s+r. \tag{4.9}$$

Put $A_1 = A^{\frac{(p-t)s+r}{q}}$ and $B_1 = \{A^{\frac{r}{2}}(A^{-\frac{t}{2}}B^pA^{-\frac{t}{2}})^sA^{\frac{r}{2}}\}^{\frac{1}{q}}$, then $A_1 \geq B_1$ by (4.7) and $kI_H \geq A \geq \frac{1}{k}I_H > 0$ assures $k^{\frac{(p-t)s+r}{q}}I_H \geq A_1 \geq k^{-\frac{(p-t)s+r}{q}}I_H$. By applying Lemma 4.3 to A_1 and B_1, we have

$$S_{k^{\frac{(p-t)s+r}{q}}}(q)^q\, A_1^q = S_k((p-t)s+r)^q\, A_1^q \geq B_1^q.$$

Multiplying $A^{-\frac{r}{2}}$ on both sides, we have (ii).
(ii) $\Longrightarrow$ (iii): Put $r = (p-t)s - 2(1-t)$ and $q = 2$ in (ii), then the condition (4.9) is satisfied and the condition (4.8) is equivalent to $(p-t)s \geq 2-t$, so that we have (iii).
(iii) $\Longrightarrow$ (iv): Put $t = 1$ in (iii), then by taking the $\frac{1}{s}$-power of both sides, we have (iv).
(iv) $\Longrightarrow$ (v): It follows from (v) of Theorem 2.16 that

$$S_k(2(p-1)s)^{\frac{2}{s}} = \left(S_{k^{2(p-1)}}(s)^{\frac{1}{s}}\right)^2 \to k^{4(p-1)} \quad \text{as } s \to \infty,$$

so that we have (v).
(v) $\Longrightarrow$ (i): We have only to put $p = 1$ in (v). □

Proof of Corollary 4.1. Put $A_1 = A^{\delta}$ and $B_1 = B^{\delta}$, then $A_1 \geq B_1 > 0$ and $k^{\delta}I_H \geq A_1 \geq \frac{1}{k^{\delta}}I_H$. By applying (iv) of Theorem 4.16 to A_1 and B_1, it follows that

$$S_{k^{\delta}}(2(p_1-1)s)^{\frac{2}{s}}A_1^{p_1} \geq B_1^{p_1}$$

holds for $p_1 \geq 1$, $s \geq 1$ such that $p_1 \geq \frac{1}{s}+1$. Put $p_1 = \frac{p}{\delta} \geq \frac{1}{s}+1$, then we have

$$S_k(2(p-\delta)s)^{\frac{2}{s}}A^p \geq B^p$$

holds for $p \geq \delta$, $s \geq 1$ such that $p \geq (\frac{1}{s}+1)\delta$. □

Remark 4.4 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. By using Uchiyama's method,* (iv) *of Theorem 4.16 can be derived from* (iv) *of Theorem 4.17 directly. In fact, the hypothesis $\log A \geq \log B$ ensures $A_n = I_H + \frac{1}{n}\log A \geq I_H + \frac{1}{n}\log B = B_n > 0$ and $M_nI_H = (1+\frac{1}{n}\log k)I_H \geq I_H + \frac{1}{n}\log A \geq (1+\frac{1}{n}\log\frac{1}{k})I_H = m_nI_H$ for sufficiently large natural number n. By Theorem 4.17, we have*

$$\max\{S_{M_n}(2(p-1)s)^{\frac{2}{s}}, S_{m_n^{-1}}(2(p-1)s)^{\frac{2}{s}}\}A_n^p \geq B_n^p$$

for $p,s \geq 1$ with $p \geq 1+\frac{1}{s}$. By substituting np to p, we have

$$\max\{S_{M_n}(2(np-1)s)^{\frac{2}{s}}, S_{m_n^{-1}}(2(np-1)s)^{\frac{2}{s}}\}A_n^{np} \geq B_n^{np}$$

for $np, s \geq 1$ with $np \geq 1+\frac{1}{s}$. Since

$$\lim_{n\to\infty}(I+\frac{1}{n}\log X)^n = X \quad \textit{for any } X > 0,$$

we obtain

$$S_k(2ps)^{\frac{2}{s}}A^p \geq B^p \quad \textit{for } s \geq 1, p \geq 0.$$

Therefore, we have (iv) *of Theorem 4.16.*

We place an emphasis on the coherence of characterizations of the chaotic order and the operator one via the Specht ratio, though our estimates via the Specht ratio are not better than the ones in Theorem 4.2 and Theorem 4.9. We observe a connection between their constants just to make sure. First of all, we start with the following lemma.

Lemma 4.4 *Let* $h \geq 1$. *Then*

$$F(s) = S_h(s)^{\frac{1}{s}} = \left(\frac{(h^s-1)h^{\frac{s}{h^s-1}}}{es\log h} \right)^{\frac{1}{s}}$$

is an increasing function for $s \geq 1$ *and a decreasing function for* $0 < s \leq 1$.

Proof. Since

$$(\log F)'(s) = \frac{F'(s)}{F(s)} = \frac{1}{s^2}(-\log(h^s-1) - \frac{s}{h^s-1}\log h + \log s + \log(\log h) + \frac{h^s}{h^s-1}\log h^s + \frac{(h^s-1)-sh^s\log h}{(h^s-1)^2}\log h^s),$$

if we put $x = h^s\,(>1)$, then we have

$$\begin{aligned}
&(\log F)'(s) \\
&= \frac{1}{s^2}\left(-\log(x-1) - \frac{\log x}{x-1} + \log(\log x) + \frac{x\log x}{x-1} + \frac{(x-1)-x\log x}{(x-1)^2}\log x\right) \\
&= \frac{1}{s^2}\left(\log\left(\frac{\log x}{x-1}\right) + x\frac{\log x}{x-1} - x\left(\frac{\log x}{x-1}\right)^2\right).
\end{aligned}$$

Klein's inequality $1 - 1/x \leq \log x \leq x-1$ and $x = h^s > 1$ imply

$$1 \geq \frac{\log x}{x-1} \geq \frac{1}{x}.$$

Then, since $L(t) = \frac{\log t}{1-t}$ is negative and increasing for $t > 0$, we have

$$\begin{aligned}
s^2(\log F)'(s) &= \log\left(\frac{\log x}{x-1}\right) + x\frac{\log x}{x-1}\left(1 - \frac{\log x}{x-1}\right) \\
&= \left(1 - \frac{\log x}{x-1}\right)\left(L\left(\frac{\log x}{x-1}\right) + \frac{x\log x}{x-1}\right) \\
&\geq \left(1 - \frac{\log x}{x-1}\right)\left(L\left(\frac{1}{x}\right) + \frac{x\log x}{x-1}\right) \\
&= \left(1 - \frac{\log x}{x-1}\right)\left(-\frac{x\log x}{x-1} + \frac{x\log x}{x-1}\right) = 0.
\end{aligned}$$

Thus we have $(\log F)'(s) \geq 0$. By $F(s) \geq 0$, F itself is increasing for $s \geq 1$. □

Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. We have the following two characterizations of the chaotic order via the Specht ratio:

(i) $\log A \geq \log B \Longleftrightarrow S_{k^2}(p)A^p \geq B^p$ for all $p > 0$.

(ii) $\log A \geq \log B \Longleftrightarrow S_{k^2}(ps)^{\frac{2}{s}}A^p \geq B^p$ for all $p > 0$ and $s \geq 1$.

We have the following relation between the constants (i) and (ii):

Lemma 4.5 *For a given $p > 0$, the constant $S_{k^2}(ps)^{2/s}$ is not smaller than the constant $S_{k^2}(p)$ for all $s \geq 1$:*

$$S_{k^2}(ps)^{2/s} \geq S_{k^2}(p) \quad \textit{for all } s \geq 1.$$

Proof. By definition, we have $S_{k^2}(p) = S_{k^{2p}}(1)$ and $S_{k^2}(ps)^{2/s} = S_{k^{2p}}(s)^{2/s}$. If we put $s = 1$, then it obviously follows that $S_{k^{2p}}(1)^2 \geq S_{k^{2p}}(1)$. Therefore by Lemma 4.4 we have

$$S_{k^2}(ps)^{2/s} = S_{k^{2p}}(s)^{2/s} \geq S_{k^{2p}}(1)^2 \geq S_{k^{2p}}(1) = S_{k^2}(p).$$

□

Next, let A and B be positive operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some $k \geq 1$. We have the following two characterizations of the operator order via the Specht ratio and the Kantorovich constant:

(iii) $A \geq B \Longleftrightarrow K\left(\frac{1}{k},k,p\right)A^p \geq B^p$ for all $p \geq 1$.

(iv) $A \geq B \Longleftrightarrow S_{k^2}((p-1)s)^{\frac{2}{s}}A^p \geq B^p$ for all $p,s \geq 1$ with $p \geq 1+\frac{1}{s}$.

Here, we investigate a relation between the constants (iii) and (iv) in the case of $p = 2$. If we put $s = 1$, then it follows that

$$S_k(2)^2 \geq K(\frac{1}{k},k,2).$$

In fact, since an inequality $x \geq e\log x$ for $x > 0$ implies

$$h^{\frac{h+1}{2(h-1)}} \geq e\log h^{\frac{h+1}{2(h-1)}}$$

where $h = k^2$, it follows that

$$\frac{(h-1)h^{\frac{1}{h-1}}}{e\log h} \geq \frac{h+1}{2\sqrt{h}}$$

or equivalently

$$\left(\frac{(h-1)h^{\frac{1}{h-1}}}{e\log h}\right)^2 \geq \frac{(h+1)^2}{4h}.$$

Therefore, it follows from Lemma 4.4 that the constant $S_k(2)^2$ is not smaller than the constant $K\left(\frac{1}{k},k,2\right)$ for all $s \geq 1$:

$$S_{k^2}(s)^{\frac{2}{s}} \geq S_{k^2}(1)^2 \geq K\left(\frac{1}{k},k,2\right) \quad \text{for all } s \geq 1.$$

4.4 The Furuta inequality version

In this section, we shall present Kantorovich type operator inequalities for the Furuta inequality related to the usual ordering and the chaotic one in terms of the generalized Kantorovich constant, a generalized condition number and the Specht ratio, in which we use variants of the grand Furuta inequality (Theorem 3.13).

Theorem 4.18 *Let A and B be positive operators such that $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. If $A \geq B$, then for each $r \geq 0$ and $\alpha > 1$*

$$K\left(m^{\frac{1}{\alpha-1}\left(\frac{p+r}{q}-(1+r)\right)}, M^{\frac{1}{\alpha-1}\left(\frac{p+r}{q}-(1+r)\right)}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.10}$$

holds for all $p \geq 1$, $q \geq 0$ such that $p \geq \alpha(1+r)q - r$, and

$$K\left(m^{\frac{p+r}{\alpha q}}, M^{\frac{p+r}{\alpha q}}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.11}$$

holds for all $p \geq 1$, $q \geq 0$ such that $\alpha(1+r)q - r \geq p \geq (1+r)q - r$, where $K(m,M,p)$ is defined by (2.29)*.*

In particular,

$$\frac{(m^{\frac{p+r}{q}-(1+r)} + M^{\frac{p+r}{q}-(1+r)})^2}{4m^{\frac{p+r}{q}-(1+r)} M^{\frac{p+r}{q}-(1+r)}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.12}$$

holds for all $p \geq 1$, $q \geq 0$ such that $p \geq 2(1+r)q - r$.

Proof. For each $r \geq 0$ and $\alpha > 1$, it follows from Theorem 3.13 that

$$A^{\frac{(p+r)s+t}{\alpha}} \geq \{A^{\frac{t}{2}} (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^s A^{\frac{t}{2}}\}^{\frac{1}{\alpha}} \tag{4.13}$$

holds for all $p \geq 1$ and $t, s \geq 0$ with

$$(1+t+r)\alpha \geq (p+r)s + t. \tag{4.14}$$

Put $A_1 = A^{\frac{(p+r)s+t}{\alpha}}$ and $B_1 = \{A^{\frac{t}{2}} (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^s A^{\frac{t}{2}}\}^{\frac{1}{\alpha}}$, then $A_1 \geq B_1$ by (4.13) and $MI_H \geq A \geq mI_H$ assures $M^{\frac{(p+r)s+t}{\alpha}} I_H \geq A_1 \geq m^{\frac{(p+r)s+t}{\alpha}} I_H$. By applying Theorem 4.2 to A_1 and B_1, we have

$$K\left(m^{\frac{(p+r)s+t}{\alpha}}, M^{\frac{(p+r)s+t}{\alpha}}, \alpha\right) A_1^{\alpha} \geq B_1^{\alpha}.$$

Multiplying $A^{-\frac{t}{2}}$ on both sides, we have

$$K\left(m^{\frac{(p+r)s+t}{\alpha}}, M^{\frac{(p+r)s+t}{\alpha}}, \alpha\right) A^{(p+r)s} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^s.$$

Put $t = \frac{(p+r)s-(1+r)\alpha}{\alpha-1}$ and $s = \frac{1}{q}$. Since $p \geq \alpha(1+r)q - r$ and $q > 0$, then it follows that $t \geq 0$, $s \geq 0$ and the condition (4.14) is satisfied. Therefore, we have

$$K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(1+r))}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(1+r))}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}}$$

for all $p \geq 1$, $q \geq 0$ such that $p \geq \alpha(1+r)q - r$, so that we have the desired inequality (4.10).

Also, putting $t = 0$ and $s = \frac{1}{q}$ in (4.13) and (4.14), we have (4.11) by the same discussion above.

For (4.12), we have only to put $\alpha = 2$ in (4.10).

Hence the proof of Theorem 4.18 is complete. □

By Theorem 4.2 and Theorem 4.18, we have the following corollary.

Corollary 4.2 *Let A and B be positive operators satisfying $A \geq B$ and $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. Then for each $r \geq 0$*

$$\left(\frac{M}{m}\right)^{\frac{p+r}{q}-(1+r)} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.15}$$

holds for all $p \geq 1$, $q \geq 0$ such that $p \geq (1+r)q - r$.

Proof. By using Theorem 4.2 and Theorem 4.18, for each $r \geq 0$ and $\alpha \geq 1$

$$\left(\frac{M}{m}\right)^{\frac{p+r}{q}-(1+r)} A^{\frac{p+r}{q}} = \left(\frac{M}{m}\right)^{\left(\frac{1}{\alpha-1}(\frac{p+r}{q}-(1+r))\right)(\alpha-1)} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 1$, $q \geq 0$ such that $p \geq \alpha(1+r)q - r$. If we put $\alpha = 1$, then we have Corollary 4.2. □

Remark 4.5 *Putting $r = 0$, $q = 1$ and $p = \alpha \geq 1$ in (4.10) of Theorem 4.18 and $r = 0$, $q = 1$ in (4.15) of Corollary 4.2, we have Theorem 4.2. Hence Theorem 4.18 and Corollary 4.2 is an extension of Theorem 4.2.*

Next, we present Kantorovich type operator inequalities for the Furuta inequality related to the operator ordering in terms of the Specht ratio.

Theorem 4.19 *Let A and B be positive operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some scalar $k > 1$. If $A \geq B$, then for each $r \geq 0$ and $\alpha > 1$*

$$S\left((k^{\frac{p+r}{q}-(1+r)})^{2s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.16}$$

holds for all $p \geq 1$, $q \geq 0$, $s \geq 1$ such that $p \geq \alpha(1+r)q - r$ and $\alpha - 1 \geq \frac{1}{s}$, and

$$S\left((k^{\frac{\alpha-1}{\alpha}\frac{p+r}{q}})^{2s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.17}$$

holds for all $p \geq 1$, $q \geq 0$, $s \geq 1$ such that $\alpha - 1 \geq \frac{1}{s}$ and $\alpha(1+r)q - r \geq p \geq (1+r)q - r$, where the Specht ratio $S(h)$ is defined by (2.35).

Proof. For each $r \geq 0$ and $\alpha > 1$, it follows from Theorem 3.13 that

$$A^{\frac{(p+r)u+t}{\alpha}} \geq \{A^{\frac{t}{2}}(A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^uA^{\frac{t}{2}}\}^{\frac{1}{\alpha}} \tag{4.18}$$

holds for all $p \geq 1$ and $t,u \geq 0$ with

$$(1+t+r)\alpha \geq (p+r)u+t \tag{4.19}$$

Put $A_1 = A^{\frac{(p+r)u+t}{\alpha}}$ and $B_1 = \{A^{\frac{t}{2}}(A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^uA^{\frac{t}{2}}\}^{\frac{1}{\alpha}}$, then $A_1 \geq B_1 > 0$ by (4.18) and $kI_H \geq A \geq \frac{1}{k}I_H > 0$ assures $k^{\frac{(p+r)u+t}{\alpha}}I_H \geq A_1 \geq k^{-\frac{(p+r)u+t}{\alpha}}I_H > 0$. By applying (iv) of Theorem 4.17 to A_1 and B_1, we have

$$S\left((k^{\frac{(p+r)s+t}{\alpha}})^{2(\alpha-1)s}\right)^{\frac{2}{s}} A_1^{\alpha} \geq B_1^{\alpha}.$$

Multiplying $A^{-\frac{t}{2}}$ on both sides, we have

$$S\left((k^{\frac{(p+r)u+t}{\alpha}})^{2(\alpha-1)s}\right)^{\frac{2}{s}} A^{(p+r)u} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^u$$

holds for all $p \geq 1$, $u,t \geq 0$ and $s \geq 1$ such that $\alpha - 1 \geq \frac{1}{s}$ and the condition (4.19).

Put $t = \frac{(p+r)u-(1+r)\alpha}{\alpha-1}$ and $u = \frac{1}{q}$. Since $p \geq \alpha(1+r)q - r$ and $q > 0$, then it follows that $t \geq 0$, $u \geq 0$ and the condition (4.19) is satisfied. Therefore, we have

$$S\left((k^{\frac{(p+r)u+t}{\alpha}})^{2(\alpha-1)s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}}$$

for all $p \geq 1$, $q \geq 0$ and $s \geq 1$ such that $p \geq \alpha(1+r)q - r$ and $\alpha - 1 \geq \frac{1}{s}$, so that we have the desired inequality (4.16).

Also, putting $t = 0$ and $u = \frac{1}{q}$ in (4.18) and (4.19), we have (4.17) by the same discussion above.

Hence the proof of Theorem 4.19 is complete. □

Remark 4.6 *Putting $r = 0$, $q = 1$ and $p = \alpha > 1$ in (4.16) of Theorem 4.19, we have* (iv) *of Theorem 4.17. Hence Theorem 4.19 is an extension of* (iv) *in Theorem 4.17.*

Corollary 4.3 *Let A and B be positive operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some scalar $k > 1$. If $A \geq B$, then for each $r \geq 0$*

$$(k^4)^{\frac{p+r}{q}-(1+r)} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 1$, $q \geq 0$ such that $p \geq (1+r)q - r$.

Proof. Since it follows from (v) of Theorem 2.16 that

$$\lim_{s\to\infty} S(k^s)^{\frac{1}{s}} = k,$$

we have this corollary by using Theorem 4.19. □

Next, we present Kantorovich type operator inequalities for the chaotic Furuta inequality related to the chaotic ordering in terms of the generalized Kantorovich constant, a generalized condition number and the Specht ratio.

Theorem 4.20 *Let A and B be positive invertible operators such that $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. If $\log A \geq \log B$, then for each $r \geq 0$ and $\alpha > 1$*

$$K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-r)}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-r)}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.20}$$

holds for all $p \geq 0$, $q \geq 0$ such that $p \geq \alpha rq - r$, and

$$K\left(m^{\frac{p+r}{\alpha q}}, M^{\frac{p+r}{\alpha q}}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 0$, $q \geq 0$ such that $\alpha rq - r \geq p \geq rq - r$, where $K(m,M,p)$ is defined by (2.29).

In particular,

$$\frac{(m^{\frac{p+r}{q}-r} + M^{\frac{p+r}{q}-r})^2}{4m^{\frac{p+r}{q}-r} M^{\frac{p+r}{q}-r}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 0$, $q \geq 0$ such that $p \geq 2rq - r$.

Proof. We can prove this theorem by a similar method as Theorem 4.18 by using Theorem 3.16 instead of Theorem 3.13. □

By Theorem 4.20 and Theorem 4.4, we have the following corollary.

Corollary 4.4 *Let A and B be positive invertible operators satisfying $\log A \geq \log B$ and $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. Then for each $r \geq 0$*

$$\left(\frac{M}{m}\right)^{\frac{p+r}{q}-r} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.21}$$

holds for all $p \geq 0$, $q \geq 0$ such that $p \geq rq - r$.

Remark 4.7 *Putting $r = 0$, $q = 1$ and $p = \alpha - 1 > 0$ in (4.20) of Theorem 4.20 and $r = 0$, $q = 1$ in (4.21) of Corollary 4.4, we have Theorem 4.4. Hence Theorem 4.20 and Corollary 4.4 can be considered as an extension of Theorem 4.4.*

Similarly, we have the following result which is considered as an extension of (ii) of Theorem 4.16.

Theorem 4.21 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some scalar $k > 1$. If $\log A \geq \log B$, then for each $r \geq 0$ and $\alpha > 1$*

$$S\left((k^{\frac{p+r}{q}-r})^{2s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 0$, $q \geq 0$, $s \geq 1$ such that $p \geq \alpha rq - r$ and $\alpha - 1 \geq \frac{1}{s}$, and

$$S\left((k^{\frac{\alpha-1}{\alpha}\frac{p+r}{q}})^{2s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 0$, $q \geq 0$, $s \geq 1$ such that $\alpha - 1 \geq \frac{1}{s}$ and $\alpha rq - r \geq p \geq rq - r$, where the Specht ratio $S(h)$ is defined by (2.35).

Proof. We can prove this theorem by a similar method as Theorem 4.21 by using (ii) of Theorem 4.16 and Theorem 3.13 instead of Theorem 3.16. □

Corollary 4.5 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some scalar $k > 1$. If $\log A \geq \log B$, then for each $r \geq 0$*

$$(k^4)^{\frac{p+r}{q}-r} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq 0$ and $q \geq 0$ such that $p \geq rq - r$.

The following corollaries are easily obtained by Theorem 4.18 and Theorem 4.19, respectively.

Corollary 4.6 *Let A and B be positive invertible operators such that $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. If $A^\delta \geq B^\delta$ for $\delta \in (0,1]$, then for each $r \geq 0$ and $\alpha > 1$*

$$K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(\delta+r))}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(\delta+r))}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \tag{4.22}$$

holds for all $p \geq \delta$, $q \geq 0$ such that $p \geq \alpha(\delta + r)q - r$, where $K(m,M,p)$ is defined by (2.29).

Corollary 4.7 *Let A and B be positive invertible operators such that $kI_H \geq A \geq \frac{1}{k}I_H$ for some scalar $k > 1$. If $A^\delta \geq B^\delta$ for $\delta \in (0,1]$, then for each $r \geq 0$ and $\alpha > 1$*

$$S\left((k^{\frac{p+r}{q}-(\delta+r)})^{2s}\right)^{\frac{2}{s}} A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$$

holds for all $p \geq \delta$, $q \geq 0$ such that $p \geq \alpha(\delta + r)q - r$, where the Specht ratio $S(h)$ is defined by (2.35)*.*

Remark 4.8 (4.22) *in Corollary 4.6 interpolates (4.10) in Theorem 4.18 and (4.20) in Theorem 4.20 by means of the generalized Kantorovich constant. Let A and B be positive invertible operators such that $MI_H \geq A \geq mI_H$ for some scalars $M > m > 0$. Then the following assertions hold:*

(i) $A \geq B$ *implies* $K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(1+r))}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(1+r))}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$ *for all $p \geq 1$, $q \geq 0$ such that $p \geq \alpha(1+r)q - r$.*

(ii) $A^\delta \geq B^\delta$ *implies* $K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(\delta+r))}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-(\delta+r))}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$ *for all $p \geq \delta$, $q \geq 0$ such that $p \geq \alpha(\delta + r)q - r$.*

(iii) $\log A \geq \log B$ *implies* $K\left(m^{\frac{1}{\alpha-1}(\frac{p+r}{q}-r)}, M^{\frac{1}{\alpha-1}(\frac{p+r}{q}-r)}, \alpha\right) A^{\frac{p+r}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}}$ *for all $p \geq 0$, $q \geq 0$ with $p > \alpha rq - r$.*

It follows that the generalized Kantorovich constant of (ii) *interpolates the scalar of* (i) *and* (iii) *continuously. In fact, if we put $\delta = 1$ in* (ii)*, then we have* (i)*. Also, if we put $\delta \to 0$ in* (ii)*, then we have* (iii)*.*

4.5 Notes

Theorem 4.1 is due to M. Fujii, Izumino, Nakamoto and Seo [85]. Theorem 4.2 is due to Furuta [113]. Theorem 4.4, Theorem 4.6 and Theorem 4.9 are due to Yamzaki and Yanagida [293]. Theorem 4.5 is due to Seo [260].

The results in Section 4.3 are due to [104] and in Section 4.4 due to [262].

Chapter 5

Operator Norm

As applications of the Mond-Pečarić method for convex functions, we shall discuss inequalities involving the operator norm. Among others, we show a converse of the Araki-Cordes inequality, the norm inequality of several geometric means and a complement of the Ando-Hiai inequality. Also, we discuss Hölder's inequality and its converses in connection with the operator geometric mean.

5.1 Operator norm and spectral radius

Let A be a (bounded linear) operator on a Hilbert space H. By Theorem 1.6, we have the following relation between the operator norm $\|\cdot\|$ and the spectral radius $r(\cdot)$:

$$r(A) \leq \|A\|. \tag{5.1}$$

In this section, we shall discuss a converse of (5.1). To consider it, we use another interpretation of the Kantorovich inequality. By Schwarz's inequality (Theorem 1.1), it follows that

$$\langle Zh, h\rangle \leq \|Zh\|\|h\| \tag{5.2}$$

for a positive operator Z and a vector $h \in H$. We first show a converse of Schwarz's inequality (5.2):

Theorem 5.1 *Let Z be a positive operator such that $mI_H \le Z \le MI_H$ for some scalars $0 < m \le M$. Then*

$$\|Zh\|\|h\| \le \frac{M+m}{2\sqrt{Mm}}\langle Zh,h\rangle \tag{5.3}$$

for every vector $h \in H$.

Proof. Let $\mathscr{E}$ be any subspace of H and there exist M' and m' such that $0 < m'I_H \le Z_{\mathscr{E}} \le M'I_H$. Then $m \le m' \le M' \le M$ and setting $t = \sqrt{M/m}$ and $t' = \sqrt{M'/m'}$, we have $t \ge t' \ge 1$. Since $t \mapsto t + 1/t$ increases on $[1,\infty)$ and

$$\frac{M+m}{2\sqrt{Mm}} = \frac{1}{2}\left(t+\frac{1}{t}\right) \qquad \text{and} \qquad \frac{M'+m'}{2\sqrt{M'm'}} = \frac{1}{2}\left(t'+\frac{1}{t'}\right),$$

we infer

$$\frac{M+m}{2\sqrt{Mm}} \ge \frac{M'+m'}{2\sqrt{M'm'}}.$$

Therefore, for a unit vector $h \in H$, it suffices to prove the theorem for $Z_{\mathscr{E}}$ with $\mathscr{E} = \operatorname{span}\{h, Zh\}$. Hence we may assume $\dim H = 2, Z = Me_1 \otimes e_1 + me_2 \otimes e_2$ and $h = xe_1 + \sqrt{1-x^2}e_2$. Setting $x^2 = y$ we have

$$\frac{\|Zh\|}{\langle Zh,h\rangle} = \frac{\sqrt{M^2y+m^2(1-y)}}{My+m(1-y)}.$$

The right-hand side attains its maximum on $[0,1]$ at $y = m/(M+m)$, and then

$$\frac{\|Zh\|}{\langle Zh,h\rangle} = \frac{M+m}{2\sqrt{Mm}}.$$

Therefore, the proof is complete. □

Remark 5.1 (5.2) *in Theorem 5.1 is equivalent to the Kantorovich inequality* (2.24)*: $\langle Zx,x\rangle\langle Z^{-1}x,x\rangle \le (M+m)^2/4Mm$ for every unit vector $x \in H$. If we put $x = Z^{1/2}h/\|Z^{1/2}h\|$ for every vector $h \in H$, then*

$$\langle ZZ^{1/2}h, Z^{1/2}h\rangle\langle Z^{-1}Z^{1/2}h, Z^{1/2}h\rangle/\|Z^{1/2}h\|^4 \le \frac{(M+m)^2}{4Mm}$$

and hence $\|Zh\|^2\|h\|^2 \le \frac{(M+m)^2}{4Mm}\langle Zh,h\rangle^2$. Taking square roots of the inequality, we have Theorem 5.1. Conversely, suppose (5.3)*. If we replace h by $Z^{-\frac{1}{2}}x$ for every unit vector $x \in H$ in* (5.3)*, then we have*

$$\|ZZ^{-\frac{1}{2}}x\|\|Z^{-\frac{1}{2}}x\| \le \frac{M+m}{2\sqrt{Mm}}\langle ZZ^{-\frac{1}{2}}x, Z^{-\frac{1}{2}}x\rangle$$

and hence $\|Z^{\frac{1}{2}}x\|\|Z^{-\frac{1}{2}}x\| \le \frac{M+m}{2\sqrt{Mm}}$. Rasing it to the second powers, we have the Kantorovich inequality.

By virtue of Theorem 5.1, we show the following converse inequality of (5.1).

Theorem 5.2 *Let A and Z be positive operators such that $mI_H \leq Z \leq MI_H$ for some scalars $0 < m \leq M$. Then*

$$\|AZ\| \leq \frac{M+m}{2\sqrt{Mm}} r(AZ). \tag{5.4}$$

Proof. We may assume that there exists a unit vector f such that $\|ZA\| = \|ZAf\|$. Then $\|ZAf\|$ is expressed as follows:

$$\begin{aligned}\|ZAf\| &= \|ZA^{1/2}(A^{1/2}f)\|\|f\| = \|ZA^{1/2}(A^{1/2}f) \otimes f\| \\ &= \|(A^{1/2}f \otimes f)A^{1/2}Z\| = \|A^{1/2}f \otimes ZA^{1/2}f\| = \|A^{1/2}f\|\|ZA^{1/2}f\|.\end{aligned}$$

Hence we have

$$\begin{aligned}\|AZ\| = \|ZA\| \leq \|ZA^{1/2}f\|\|A^{1/2}f\| &\leq \frac{M+m}{2\sqrt{Mm}}\langle ZA^{1/2}f, A^{1/2}f\rangle \\ &= \frac{M+m}{2\sqrt{Mm}}\langle A^{1/2}ZA^{1/2}f, f\rangle \leq \frac{M+m}{2\sqrt{Mm}} r(A^{1/2}ZA^{1/2}) \\ &= \frac{M+m}{2\sqrt{Mm}} r(AZ)\end{aligned}$$

by Theorem 5.1. □

Remark 5.2 *Theorem 5.2 extends Theorem 5.1. Indeed, if we put $A = h \otimes h$ in Theorem 5.2, then we have Theorem 5.1.*

Let Z be a positive operator and A a contraction. Then

$$AZA \leq Z$$

does not always hold in general. As a matter of fact, if we put

$$Z = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

then we have $Z \geq 0$ and $0 \leq A \leq I_H$, but

$$Z - AZA = \begin{pmatrix} \frac{5}{4} & -\frac{3}{4} \\ -\frac{3}{4} & \frac{1}{4} \end{pmatrix} \ngeq 0.$$

By using Theorem 5.2, we have the following operator inequality.

Theorem 5.3 *Let A be a contraction and Z a positive operator such that $mI_H \leq Z \leq MI_H$ for some scalars $0 < m \leq M$. Then*

$$AZA \leq \frac{(M+m)^2}{4Mm} Z.$$

Proof. It follows from Theorem 5.2 that

$$\|Z^{-1/2}AZ^{1/2}\| = \|Z^{-1/2}AZ^{-1/2}Z\| \leq \frac{M+m}{2\sqrt{Mm}} r(Z^{-1/2}AZ^{-1/2}Z)$$
$$= \frac{M+m}{2\sqrt{Mm}} r(A) \leq \frac{M+m}{2\sqrt{Mm}}.$$

Hence we have

$$Z^{-1/2}AZAZ^{-1/2} \leq \|Z^{-1/2}AZAZ^{-1/2}\| I_H \leq \frac{(M+m)^2}{4Mm} I_H.$$

□

An important source of interesting inequalities in operator theory is the study of rearrangements in a product. The following rearrangement inequality is well known:

$$\|AB\| \leq \|BA\| \tag{5.5}$$

whenever AB is normal. In fact, since the spectral radii of AB and BA are equal and normality of AB implies $\|AB\| = r(AB)$, we have

$$\|AB\| = r(AB) = r(BA) \leq \|BA\|.$$

Thus, when $AB \geq 0$ the following theorem is a generalization of (5.5).

Theorem 5.4 *Let A, B be operators such that $AB \geq 0$ and let Z be a positive operator such that $mI_H \leq Z \leq MI_H$ for some scalars $0 < m \leq M$. Then*

$$\|ZAB\| \leq \frac{M+m}{2\sqrt{Mm}} \|BZA\|.$$

Proof. By Theorem 5.2, we have

$$\|ZAB\| \leq \frac{M+m}{2\sqrt{Mm}} r(ZAB) = \frac{M+m}{2\sqrt{Mm}} r(BZA) \leq \frac{M+m}{2\sqrt{Mm}} \|BZA\|.$$

□

We shall extend Theorem 5.2 by applying the Mond-Pečarić method for convex functions. For that purpose, we need some preliminaries.

Let A be a positive operator on a Hilbert space H and x a unit vector in H. By the Hölder-McCarthy inequality (Theorem 2.11), we have the relation between the continuous power mean and the continuous arithmetic one:

$$\langle Ax, x\rangle \leq \langle A^p x, x\rangle^{\frac{1}{p}} \qquad \text{for all } p > 1. \tag{5.6}$$

By using the Mond-Pečarić method, we have the following converse inequality of (5.6).

Lemma 5.1 *If A is a positive operator on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$, then for each $\alpha > 0$*

$$\langle A^p x, x\rangle^{\frac{1}{p}} \le \alpha\langle Ax, x\rangle + \beta(m, M, p, \alpha) \qquad \textit{for all } p > 1$$

holds for every unit vector $x \in H$, where

$$\beta(m,M,p,\alpha) = \begin{cases} \frac{p-1}{p}\left(\frac{a_p}{\alpha p}\right)^{\frac{1}{p-1}} + \alpha b_p & \textit{if} \quad \frac{a_p}{pM^{p-1}} \le \alpha \le \frac{a_p}{pm^{p-1}}, \\ (1-\alpha)M & \textit{if} \quad 0 < \alpha \le \frac{a_p}{pM^{p-1}}, \\ (1-\alpha)m & \textit{if} \quad \alpha \ge \frac{a_p}{pm^{p-1}} \end{cases} \tag{5.7}$$

and $\quad a_p := \dfrac{M^p - m^p}{M - m}, \quad b_p := \dfrac{Mm^p - mM^p}{M - m}.$

Proof. For the sake of reader's convenience, we give a proof. Put $f(t) = (a_p t + b_p)^{\frac{1}{p}} - \alpha t$ and $\beta = \beta(m, M, p, \alpha) = \max\{f(t) : m \le t \le M\}$. Then it follows that

$$f'(t) = \frac{a_p}{p}(a_p t + b_p)^{\frac{1}{p}-1} - \alpha$$

and the equation $f'(t) = 0$ has exactly one solution

$$t_0 = \frac{1}{a_p}\left(\frac{\alpha p}{a_p}\right)^{\frac{p}{1-p}} - \frac{b_p}{a_p}.$$

If $m \le t_0 \le M$, then we have $\beta = \max_{m \le t \le M} f(t) = f(t_0)$ since

$$f''(t) = \frac{a_p^2(1-p)}{p^2}(a_p t + b_p)^{\frac{1}{p}-2} < 0$$

and the condition $m \le t_0 \le M$ is equivalent to the condition

$$\frac{a_p}{pM^{p-1}} \le \alpha \le \frac{a_p}{pm^{p-1}}.$$

If $M \le t_0$, then $f(t)$ is increasing on $[m, M]$ and hence we have $\beta = f(t_0) = (1-\alpha)M$ for $t_0 = M$. Similarly, we have $\beta = f(t_0) = (1-\alpha)m$ for $t_0 = m$ if $t_0 \le m$. Hence it follows that

$$(a_p t + b_p)^{\frac{1}{p}} - \alpha t \le \beta \qquad \text{for all } t \in [m, M].$$

Since t^p is convex for $p > 1$, it follows that $t^p \le a_p t + b_p$ for $t \in [m, M]$. By the spectral theorem, we have $A^p \le a_p A + b_p I_H$ and hence $\langle A^p x, x\rangle \le a_p\langle Ax, x\rangle + b_p$ for every unit vector $x \in H$. Therefore we have

$$\begin{aligned} \langle A^p x, x\rangle^{\frac{1}{p}} - \alpha\langle Ax, x\rangle &\le (a_p\langle Ax, x\rangle + b_p)^{\frac{1}{p}} - \alpha\langle Ax, x\rangle \\ &\le \max_{m \le t \le M} f(t) = \beta(m, M, p, \alpha) \end{aligned}$$

as desired. □

As a complementary result, we state the following lemma.

Lemma 5.2 *If A is a positive operator on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$, then for each $\alpha > 0$*

$$\langle A^p x, x\rangle^{\frac{1}{p}} \ge \alpha\langle Ax, x\rangle + \overline{\beta}(m,M,p,\alpha) \qquad \text{for all } 0 < p < 1$$

holds for every unit vector $x \in H$, where

$$\overline{\beta}(m,M,p,\alpha) = \begin{cases} \frac{p-1}{p}\left(\frac{a_p}{\alpha p}\right)^{\frac{1}{p-1}} + \alpha b_p & \text{if} \quad \frac{a_p}{pm^{p-1}} \le \alpha \le \frac{a_p}{pM^{p-1}}, \\ (1-\alpha)M & \text{if} \quad \alpha \ge \frac{a_p}{pM^{p-1}}, \\ (1-\alpha)m & \text{if} \quad 0 < \alpha \le \frac{a_p}{pm^{p-1}}, \end{cases}$$

$$\text{and} \quad a_p := \frac{M^p - m^p}{M - m}, \quad b_p := \frac{Mm^p - mM^p}{M - m}.$$

By Lemma 5.1 and 5.2, we have the following estimates of both the difference and the ratio in the inequality (5.6).

Lemma 5.3 *If A is a positive operator on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$, then*

$$\langle A^p x, x\rangle^{\frac{1}{p}} \le K(m,M,p)^{\frac{1}{p}}\langle Ax, x\rangle \qquad \text{for all } p > 1 \tag{5.8}$$

and

$$K(m,M,p)^{\frac{1}{p}}\langle Ax, x\rangle \le \langle A^p x, x\rangle^{\frac{1}{p}} \qquad \text{for all } 0 < p < 1 \tag{5.9}$$

hold for every unit vector $x \in H$, where the generalized Kantorovich constant $K(m,M,p)$ is defined by (2.29)*.*

Proof. For $p > 1$, if we put $\beta(m,M,p,\alpha) = 0$ in Lemma 5.1, then it follows that

$$\frac{p-1}{p}\left(\frac{M^p - m^p}{p(M-m)}\right)^{\frac{1}{p-1}} + \alpha^{\frac{p}{p-1}}\frac{(Mm^p - mM^p)}{M^p - m^p} = 0$$

and hence

$$\alpha^{\frac{p}{p-1}} = -\frac{p-1}{p}\left(\frac{M^p - m^p}{p(M-m)}\right)^{\frac{1}{p-1}}\frac{M^p - m^p}{Mm^p - mM^p}.$$

Therefore, we have

$$\begin{aligned} \alpha^p &= \frac{M^p - m^p}{p(M-m)}\left(\frac{p-1}{p}\frac{M^p - m^p}{mM^p - Mm^p}\right)^{p-1} \\ &= K(m,M,p) \end{aligned}$$

and we obtain the desired inequality (5.8). For $0 < p < 1$, we similarly have the inequality (5.9) by Lemma 5.2. □

Lemma 5.4 *If A is a positive operator on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m < M$, then*

$$\langle A^p x, x\rangle^{\frac{1}{p}} - \langle Ax, x\rangle \le -C\Big(m^p, M^p, \frac{1}{p}\Big) \qquad \text{for all } p > 1 \tag{5.10}$$

and

$$-C\Big(m^p, M^p, \frac{1}{p}\Big) \le \langle A^p x, x\rangle^{\frac{1}{p}} - \langle Ax, x\rangle \qquad \text{for all } 0 < p < 1 \tag{5.11}$$

hold for every unit vector $x \in H$, where the constant $C(m,M,p)$ is defined by (2.38)*.*

Proof. For $p > 1$, if we put $\alpha = 1$ in Lemma 5.1, then it follows that

$$\begin{aligned} -C\Big(m^p, M^p, \frac{1}{p}\Big) &= (1-\frac{1}{p})\left(\frac{M-m}{\frac{1}{p}(M^p - m^p)}\right)^{\frac{\frac{1}{p}}{\frac{1}{p}-1}} - \frac{M^p m - m^p M}{M-m} \\ &= \beta(m,M,p,1) \end{aligned}$$

and we obtain the desired inequality (5.10). For $0 < p < 1$, we similarly have the inequality (5.11) by Lemma 5.2. □

The following theorem is a generalization of Theorem 5.2.

Theorem 5.5 *If A and Z are positive operators on H such that $mI_H \le Z \le MI_H$ for some scalars $0 < m \le M$, then for each $\alpha > 0$*

$$\|(AZ^pA)^{\frac{1}{p}}\| \le \alpha\, r(ZA^{\frac{2}{p}}) + \beta(m,M,p,\alpha)\|A\|^{\frac{2}{p}} \qquad \text{for all } p > 1,$$

where $\beta(m,M,p,\alpha)$ is defined by (5.7)*.*

Proof. For every unit vector $x \in H$, it follows from $0 < \frac{1}{p} < 1$ that

$$\begin{aligned} &\langle (AZ^pA)^{\frac{1}{p}}x, x\rangle \le \langle AZ^pAx, x\rangle^{\frac{1}{p}} \qquad \text{by the Hölder-McCarthy inequality} \\ &= \left\langle Z^p \frac{Ax}{\|Ax\|}, \frac{Ax}{\|Ax\|}\right\rangle^{\frac{1}{p}} \|Ax\|^{\frac{2}{p}} \\ &\le \left(\alpha\left\langle Z\frac{Ax}{\|Ax\|}, \frac{Ax}{\|Ax\|}\right\rangle + \beta(m,M,p,\alpha)\right)\|Ax\|^{\frac{2}{p}} \qquad \text{by Lemma 5.1} \\ &= \alpha\langle ZAx, Ax\rangle\|Ax\|^{\frac{2}{p}-2} + \beta(m,M,p,\alpha)\|Ax\|^{\frac{2}{p}} \\ &= \alpha\left\langle A^{\frac{1}{p}}ZA^{\frac{1}{p}}\frac{A^{1-\frac{1}{p}}x}{\|A^{1-\frac{1}{p}}x\|}, \frac{A^{1-\frac{1}{p}}x}{\|A^{1-\frac{1}{p}}x\|}\right\rangle \|Ax\|^{\frac{2}{p}-2}\|A^{1-\frac{1}{p}}x\|^2 + \beta(m,M,p,\alpha)\|Ax\|^{\frac{2}{p}} \end{aligned}$$

and

$$\begin{aligned} \|Ax\|^{\frac{2}{p}-2}\|A^{1-\frac{1}{p}}x\|^2 &= \langle A^2x, x\rangle^{\frac{1}{p}-1}\langle A^{2-\frac{2}{p}}x, x\rangle \\ &\le \langle A^2x, x\rangle^{\frac{1}{p}-1}\langle A^2x, x\rangle^{1-\frac{1}{p}} = 1 \qquad \text{by } 0 < 1-\tfrac{1}{p} < 1. \end{aligned}$$

By combining two inequalities above, we have

$$\begin{aligned}\langle (AZ^pA)^{\frac{1}{p}}x,x\rangle &\le \alpha\, \|A^{\frac{1}{p}}ZA^{\frac{1}{p}}\| + \beta(m,M,p,\alpha)\|Ax\|^{\frac{2}{p}} \\ &= \alpha\, r(A^{\frac{1}{p}}ZA^{\frac{1}{p}}) + \beta(m,M,p,\alpha)\|Ax\|^{\frac{2}{p}} \\ &\le \alpha\, r(ZA^{\frac{2}{p}}) + \beta(m,M,p,\alpha)\|A\|^{\frac{2}{p}}\end{aligned}$$

for every unit vector $x \in H$ and hence we have the desired inequality. □

Remark 5.3 *If A and Z are positive operators, then it follows that*

$$r(ZA^{\frac{2}{p}}) \le \|(AZ^pA)^{\frac{1}{p}}\| \qquad \text{for all } p > 1. \tag{5.12}$$

As a matter of fact, by the Araki-Cordes inequality (Theorem 5.9), we have

$$r(ZA^{\frac{2}{p}}) = r(A^{\frac{1}{p}}ZA^{\frac{1}{p}}) = \|A^{\frac{1}{p}}ZA^{\frac{1}{p}}\| \le \|(AZ^pA)^{\frac{1}{p}}\|$$

for all $p > 1$. Therefore, the inequality in Theorem 5.5 can be considered as a converse inequality of (5.12)*.*

The following theorem is a variant of Theorem 5.5 with 2-variables.

Theorem 5.6 *If A and Z are positive operators on H such that $mI_H \le Z \le MI_H$ for some scalars $0 < m \le M$, then for each $\alpha > 0$*

$$\|(AZ^pA)^{\frac{1}{q}}\| \le \alpha\, r(Z^{\frac{p}{q}}A^{\frac{2}{q}}) + \beta(m^{\frac{p}{q}}, M^{\frac{p}{q}}, q, \alpha)\|A\|^{\frac{2}{q}} \qquad \text{for all } p > 1 \text{ and } q > 1,$$

where $\beta(m,M,p,\alpha)$ is defined by (5.7)*.*

Proof. For every unit vector $x \in H$, we have

$$\begin{aligned}\langle (AZ^pA)^{\frac{1}{q}}x,x\rangle &\le \langle AZ^pAx,x\rangle^{\frac{1}{q}} \qquad \text{by } 0 < \tfrac{1}{q} < 1 \\ &= \left\langle (Z^{\frac{p}{q}})^q \frac{Ax}{\|Ax\|}, \frac{Ax}{\|Ax\|}\right\rangle^{\frac{1}{q}} \|Ax\|^{\frac{2}{q}} \\ &\le \left(\alpha \left\langle Z^{\frac{p}{q}} \frac{Ax}{\|Ax\|}, \frac{Ax}{\|Ax\|}\right\rangle + \beta(m^{\frac{p}{q}}, M^{\frac{p}{q}}, q, \alpha)\right) \|Ax\|^{\frac{2}{q}}.\end{aligned}$$

The rest of the proof is proved in a similar way as the proof of Theorem 5.5. □

Theorem 5.7 *Let A and Z be positive operators on H such that $mI_H \le Z \le MI_H$ for some scalars $0 < m \le M$. Then for each $p > 1$*

$$\|(AZ^pA)^{\frac{1}{p}}\| \le K(m,M,p)^{\frac{1}{p}}\; r(ZA^{\frac{2}{p}}). \tag{5.13}$$

In addition, (5.13) *is equivalent to* (5.8) *in Lemma 5.3.*

Proof. By using (5.8) of Lemma 5.3 instead of Lemma 5.1 in the proof of Theorem 5.5, we obtain (5.13). Conversely, for every unit vector $x \in H$, if we put $A = x \otimes x$ in (5.13), then

$$\|(x \otimes x)Z^p(x \otimes x)\|^{1/p} = \|\langle x, Z^p x\rangle\langle x,x\rangle\|^{1/p} = \langle Z^p x, x\rangle^{1/p}$$

and

$$r(ZA^{2/p}) = r(AZA) \le \|AZA\| = \|\langle Zx,x\rangle\langle x,x\rangle\| = \langle Zx,x\rangle.$$

Hence we have (5.8) of Lemma 5.3. □

Remark 5.4 *We have Theorem 5.2 as a special case of Theorem 5.7. As a matter of fact, if we put $p = 2$ in Theorem 5.7, then we have*

$$\|(AZ^2A)^{\frac{1}{2}}\| \le K(m,M,2)^{\frac{1}{2}}\, r(ZA).$$

Since $\|(AZ^2A)^{\frac{1}{2}}\| = \|(ZA)^*(ZA)\|^{\frac{1}{2}} = \|ZA\| = \|AZ\|$ *and* $K(m,M,2)^{\frac{1}{2}} = \left(\frac{(M+m)^2}{4Mm}\right)^{\frac{1}{2}} = \frac{M+m}{2\sqrt{Mm}}$, *we have the desired inequality* (5.4) *in Theorem 5.2.*

Theorem 5.8 *Let A and Z be positive operators on H such that $mI_H \le Z \le MI_H$ for some scalars $0 < m < M$. Then for each $p > 1$*

$$\|(AZ^pA)^{\frac{1}{p}}\| \le r(ZA^{\frac{2}{p}}) - C\left(m^p, M^p, \frac{1}{p}\right)\|A\|^{\frac{2}{p}}. \tag{5.14}$$

In addition, (5.14) *is equivalent to* (5.10) *in Lemma 5.4.*

Proof. By using (5.10) of Lemma 5.4 instead of Lemma 5.1 in the proof of Theorem 5.5, we obtain (5.14). Conversely, for every unit vector $x \in H$, if we put $A = x \otimes x$ in (5.14), then we have (5.10) of Lemma 5.4. □

We have the following corollary as a special case of (5.14) in Theorem 5.8, which is a difference type converse inequality of (5.1).

Corollary 5.1 *If A and Z are positive operators on H such that $0 < mI_H \le Z \le MI_H$ for some scalars $0 < m < M$, then*

$$\|ZA\| - r(ZA) \le \frac{(M-m)^2}{4(M+m)}\|A\|. \tag{5.15}$$

Proof. If we put $p = 2$ in Theorem 5.8, then we have (5.15) since

$$C\left(m^2, M^2, \frac{1}{2}\right) = \frac{(M-m)^2}{4(M+m)}.$$

□

5.2 The Araki-Cordes inequality

First of all, we recall the Araki-Cordes inequality (AC) for the operator norm in §3.6:

Theorem 5.9 *Let A and B be positive operators. Then*

$$\|B^pA^pB^p\| \le \|(BAB)^p\| \qquad \text{for all } 0 < p \le 1 \tag{5.16}$$

or equivalently

$$\|(BAB)^p\| \le \|B^pA^pB^p\| \qquad \text{for all } p > 1.$$

The Cordes inequality for the operator norm is as follows:

Theorem 5.10 *Let A and B be positive operators. Then*

$$\|A^pB^p\| \le \|AB\|^p \qquad \text{for all } 0 < p \le 1$$

or equivalently

$$\|AB\|^p \le \|A^pB^p\| \qquad \text{for all } p > 1.$$

Proof. By using the Araki-Cordes inequality, we have

$$\|A^pB^p\|^2 = \|B^pA^{2p}B^p\| \le \|(BA^2B)^p\| = \|BA^2B\|^p = \|AB\|^{2p}$$

for all $0 < p \le 1$. □

In this section, we show converse inequalities to these inequalities and investigate the equivalence among converse inequalities of Araki, Cordes and Löwner-Heinz inequalities.

First of all, we show the following ratio type converse inequality of the Araki-Cordes inequality.

Theorem 5.11 *If A and B are positive operators on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m \le M$, then*

$$K(m,M,p)\|BAB\|^p \le \|B^pA^pB^p\| \qquad \text{for all } 0 < p \le 1 \tag{5.17}$$

or equivalently

$$\|B^pA^pB^p\| \le K(m,M,p)\ \|BAB\|^p \qquad \text{for all } p > 1, \tag{5.18}$$

where $K(m,M,p)$ is defined by (2.29)*.*

In particular,

$$\|B^2A^2B^2\| \le \frac{(M+m)^2}{4Mm}\|BAB\|^2$$

and

$$\frac{2\sqrt[4]{Mm}}{\sqrt{M}+\sqrt{m}}\|BAB\|^{\frac{1}{2}} \le \|B^{\frac{1}{2}}A^{\frac{1}{2}}B^{\frac{1}{2}}\|.$$

Proof. Suppose that $0 < p \le 1$. For every unit vector $x \in H$, it follows that

$$\begin{aligned}
&\langle (BAB)^p x, x\rangle \\
&\le \langle BABx, x\rangle^p \qquad \text{by the Hölder-McCarthy inequality and } 0 < p \le 1 \\
&= \left\langle (A^p)^{\frac{1}{p}} \frac{Bx}{\|Bx\|}, \frac{Bx}{\|Bx\|} \right\rangle^p \|Bx\|^{2p} \\
&\le K\left(m^p, M^p, \frac{1}{p}\right)^p \left\langle A^p \frac{Bx}{\|Bx\|}, \frac{Bx}{\|Bx\|} \right\rangle \|Bx\|^{2p} \qquad \text{by Lemma 5.3 and } \tfrac{1}{p} > 1 \\
&= K\left(m^p, M^p, \frac{1}{p}\right)^p \langle A^p Bx, Bx\rangle \|Bx\|^{2p-2} \\
&= K\left(m^p, M^p, \frac{1}{p}\right)^p \left\langle B^p A^p B^p \frac{B^{1-p}x}{\|B^{1-p}x\|}, \frac{B^{1-p}x}{\|B^{1-p}x\|} \right\rangle \|Bx\|^{2p-2} \|B^{1-p}x\|^2
\end{aligned}$$

and

$$\begin{aligned}
\|Bx\|^{2p-2}\|B^{1-p}x\|^2 &= \langle B^2x, x\rangle^{p-1} \langle B^{2-2p}x, x\rangle \\
&\le \langle B^2x, x\rangle^{p-1} \langle B^2x, x\rangle^{1-p} = 1 \qquad \text{by } 0 < 1-p < 1.
\end{aligned}$$

By combining two inequalities above, we have

$$\begin{aligned}
\|BAB\|^p &= \|(BAB)^p\| \\
&\le K\left(m^p, M^p, \frac{1}{p}\right)^p \|B^pA^pB^p\| = K(m, M, p)^{-1} \|B^pA^pB^p\|
\end{aligned}$$

because $K(m,M,p)^{1/p} = K(m^p, M^p, 1/p)^{-1}$ by the inversion formula in Theorem 2.13. Hence we have the desired inequality (5.17).

Next, we show (5.17) $\Longrightarrow$ (5.18). For $p > 1$, since $0 < \frac{1}{p} < 1$, it follows from (5.17) that

$$K\left(m, M, \frac{1}{p}\right) \|BAB\|^{\frac{1}{p}} \le \|B^{\frac{1}{p}} A^{\frac{1}{p}} B^{\frac{1}{p}}\|.$$

By replacing A and B by A^p and B^p respectively, in the inequality above we have

$$K\left(m^p, M^p, \frac{1}{p}\right) \|B^pA^pB^p\|^{\frac{1}{p}} \le \|BAB\|,$$

and so

$$K(m, M, p)^{-1} \|B^pA^pB^p\| \le \|BAB\|^p$$

by the inversion formula in Theorem 2.13. Similarly we can show (5.18) $\Longrightarrow$ (5.17). □

Remark 5.5 *Theorem 5.11 implies Theorem 5.7. In fact, for each $p > 1$,*

$$\begin{aligned}
\|(AZ^pA)^{\frac{1}{p}}\| &\le K\left(m^p, M^p, \frac{1}{p}\right)^{-1} \|A^{\frac{1}{p}} Z A^{\frac{1}{p}}\| \\
&= K(m, M, p)^{\frac{1}{p}} r(A^{\frac{1}{p}} Z A^{\frac{1}{p}}) \\
&= K(m, M, p)^{\frac{1}{p}} r(Z A^{\frac{2}{p}}).
\end{aligned}$$

Next, we show a difference type converse inequality of the Araki-Cordes one.

Theorem 5.12 *If A and B are positive operators on H such that $mI_H \le A \le MI_H$ for some scalars $0 < m \le M$, then*

$$\|BAB\|^p \le \|B^pA^pB^p\| - C(m,M,p)\|B\|^{2p} \qquad \text{for all } 0 < p \le 1,$$

or equivalently

$$\|BAB\|^p \ge \|B^pA^pB^p\| - C(m,M,p)\|B\|^{2p} \qquad \text{for all } p > 1,$$

where $C(m,M,p)$ is defined by (2.38).

In particular,

$$\|BAB\|^{\frac{1}{2}} \le \|B^{\frac{1}{2}}A^{\frac{1}{2}}B^{\frac{1}{2}}\| + \frac{(\sqrt{M}-\sqrt{m})^2}{4(\sqrt{M}+\sqrt{m})}\|B\|.$$

and

$$\|B^2A^2B^2\| \le \|BAB\|^2 + \frac{(M-m)^2}{4}\|B\|^4$$

Proof. For $0 < p \le 1$,

$$\begin{aligned}
&\langle (BAB)^p x, x\rangle \le \langle BABx, x\rangle^p \\
&= \left\langle (A^p)^{\frac{1}{p}} \frac{Bx}{\|Bx\|}, \frac{Bx}{\|Bx\|} \right\rangle^p \|Bx\|^{2p} \\
&\le \left(\left\langle A^p \frac{Bx}{\|Bx\|}, \frac{Bx}{\|Bx\|} \right\rangle - C(m,M,p) \right) \|Bx\|^{2p} \\
&= \left\langle B^pA^pB^p \frac{B^{1-p}x}{\|B^{1-p}x\|}, \frac{B^{1-p}x}{\|B^{1-p}x\|} \right\rangle \|Bx\|^{2p-2}\|B^{1-p}x\|^{2p} - C(m,M,p)\|Bx\|^{2p} \\
&\le \|B^pA^pB^p\| - C(m,M,p)\|B\|^{2p}.
\end{aligned}$$

The last inequality holds since

$$\begin{aligned}
\|Bx\|^{2p-2}\|B^{1-p}x\|^2 &= \langle B^2x, x\rangle^{p-1}\langle B^{2-2p}x, x\rangle \\
&\le \langle B^2x, x\rangle^{p-1}\langle B^2x, x\rangle^{1-p} = 1 \qquad \text{by } 0 < 1-p < 1.
\end{aligned}$$

Hence we have

$$\|BAB\|^p \le \|B^pA^pB^p\| - C(m,M,p)\|B\|^{2p}.$$

Next, suppose that $p > 1$. For every unit vector $x \in H$ we have

$$\begin{aligned}
\|Bx\|^{2p-2}\|B^{1-p}x\|^2 &= \langle B^2x, x\rangle^{p-1}\langle B^{2-2p}x, x\rangle \\
&\ge \langle B^2x, x\rangle^{p-1}\langle B^2x, x\rangle^{1-p} = 1 \qquad \text{by } 1-p < 0
\end{aligned}$$

and

$$\begin{aligned}
&\langle (BAB)^p x, x\rangle \geq \langle BABx, x\rangle^p \\
&= \left\langle (A^p)^{\frac{1}{p}} \frac{Bx}{\|Bx\|}, \frac{Bx}{\|Bx\|} \right\rangle^p \|Bx\|^{2p} \\
&\geq \langle A^p Bx, Bx\rangle \|Bx\|^{2p-2} - C(m,M,p)\|Bx\|^{2p} \\
&\geq \left\langle B^p A^p B^p \frac{B^{1-p}x}{\|B^{1-p}x\|}, \frac{B^{1-p}x}{\|B^{1-p}x\|} \right\rangle \|Bx\|^{2p-2}\|B^{1-p}x\|^2 - C(m,M,p)\|Bx\|^{2p} \\
&\geq \left\langle B^p A^p B^p \frac{B^{1-p}x}{\|B^{1-p}x\|}, \frac{B^{1-p}x}{\|B^{1-p}x\|} \right\rangle - C(m,M,p)\|Bx\|^{2p}.
\end{aligned}$$

By a suitable unit vector $x \in H$, it follows that

$$\langle (BAB)^p x, xa\rangle \geq \|B^p A^p B^p\| - C(m,M,p)\|Bx\|^{2p}.$$

Since $\|Bx\|^{2p} \leq \|B\|^{2p}$, we have $-C(m,M,p)\|Bx\|^{2p} \geq -C(m,M,p)\|B\|^{2p}$ and hence

$$\|BAB\|^p \geq \|B^p A^p B^p\| - C(m,M,p)\|B\|^{2p}.$$

□

Moreover, we obtain the following converse inequality of the Cordes inequality by Theorem 5.11.

Theorem 5.13 *If A and B are positive operators on H such that $mI_H \leq A \leq MI_H$ for some scalars $0 < m \leq M$, then*

$$\|A^p B^p\| \leq K(m^2, M^2, p)^{\frac{1}{2}} \, \|AB\|^p \qquad \textit{for all } p > 1$$

or equivalently

$$K(m^2, M^2, p)^{\frac{1}{2}} \|AB\|^p \leq \|A^p B^p\| \qquad \textit{for all } 0 < p < 1.$$

In particular,

$$\|A^2 B^2\| \leq \frac{M^2 + m^2}{2Mm} \|AB\|^2$$

and

$$\sqrt{\frac{2\sqrt{Mm}}{M+m}} \|AB\|^{\frac{1}{2}} \leq \|A^{\frac{1}{2}} B^{\frac{1}{2}}\|.$$

Proof. For a given $p > 1$, it follows from Theorem 5.11 that

$$\|B^p A^p B^p\| \leq K(m,M,p)\|BAB\|^p$$

and hence

$$\|A^{\frac{p}{2}} B^p\|^2 \leq K(m,M,p)\|A^{\frac{1}{2}} B\|^{2p}.$$

If we replace A by A^2, then we have

$$\|A^pB^p\|^2 \le K(m^2,M^2,p)\|AB\|^{2p}$$

as desired. □

The equivalence among the converse inequalities of Araki, Cordes and Löwner-Heinz inequalities is now given as follows.

Theorem 5.14 *Let A,B be positive operators such that $mI_H \le A \le MI_H$ for some scalars $0 < m \le M$. Then for a given $p > 1$, the following are mutually equivalent:*

(a) $A \ge B \ge 0$ *implies* $K(m,M,p)A^p \ge B^p$.

(b) $\|A^pB^p\| \le K(m^2,M^2,p)^{1/2}\|AB\|^p$.

(c) $\|B^pA^pB^p\| \le K(m,M,p)\|BAB\|^p$.

(b') $K(m^2,M^2,1/p)^{1/2}\|AB\|^p \le \|A^pB^p\|$.

(c') $K(m,M,1/p)\|BAB\|^p \le \|B^pA^pB^p\|$.

Proof. The proof is divided into three parts, namely the equivalence $(a) \Longrightarrow (b) \Longrightarrow (c) \Longrightarrow (a)$, $(b) \Longleftrightarrow$ (b') and $(c) \Longleftrightarrow$ (c').

(a) $\Longrightarrow$ (b): It follows that

$$\begin{aligned}(a) \Longleftrightarrow\ & \left(\|A^{-\frac{1}{2}}B^{\frac{1}{2}}\| \le 1 \text{ implies } \|A^{-\frac{p}{2}}B^{\frac{p}{2}}\|^2 \le K(m,M,p)\right)\\ \Longleftrightarrow\ & \left(\|A^{\frac{1}{2}}B^{\frac{1}{2}}\| \le 1 \text{ implies } \|A^{\frac{p}{2}}B^{\frac{p}{2}}\|^2 \le K(M^{-1},m^{-1},p) = K(m,M,p)\right)\\ \Longleftrightarrow\ & \left(\|AB\| \le 1 \text{ implies } \|A^pB^p\| \le K(m^2,M^2,p).\right)\end{aligned}$$

If we put $B_1 = B/\|AB\|$, then it follows from $\|AB_1\| = 1$ that

$$\|A^pB_1^p\| \le K(m^2,M^2,p)^{\frac{1}{2}} \Longleftrightarrow \|A^pB^p\| \le K(m^2,M^2,p)^{\frac{1}{2}}\|AB\|^p.$$

(b) $\Longrightarrow$ (c): If we replace A by $A^{\frac{1}{2}}$ in (B), then it follows that

$$\|A^{\frac{p}{2}}B^p\| \le K(m,M,p)^{\frac{1}{2}}\|A^{\frac{1}{2}}B\|^p.$$

Squaring both sides, we have

$$\|B^pA^pB^p\| \le K(m,M,p)\|BAB\|^p.$$

(c) $\Longrightarrow$ (a): If we replace B by $B^{\frac{1}{2}}$ and A by A^{-1} in (C), then it follows that

$$\|B^{\frac{p}{2}}A^{-p}B^{\frac{p}{2}}\| \le K(M^{-1},m^{-1},p)\|B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}\|^p.$$

By rearranging it, we have

$$\|A^{-\frac{p}{2}}B^pA^{-\frac{p}{2}}\| \le K(m,M,p)\|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|^p.$$

Since $A \ge B \ge 0$, it follows from $A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le I_H$ that

$$\|A^{-\frac{p}{2}}B^pA^{-\frac{p}{2}}\| \le K(m,M,p)$$

and hence

$$B^p \le K(m,M,p)A^p.$$

(b) $\Longleftrightarrow$ (b'): If we replace A and B by $A^{\frac{1}{p}}$ and $B^{\frac{1}{p}}$ respectively in (B), then it follows that

$$\begin{aligned}
(B) \Longleftrightarrow \quad & \|AB\| \le K\left(m^{\frac{2}{p}},M^{\frac{2}{p}},p\right)^{\frac{1}{2}}\|A^{\frac{1}{p}}B^{\frac{1}{p}}\|^p \\
\Longleftrightarrow \quad & \|AB\|^{\frac{1}{p}} \le K\left(m^{\frac{2}{p}},M^{\frac{2}{p}},p\right)^{\frac{1}{2p}}\|A^{\frac{1}{p}}B^{\frac{1}{p}}\| \\
\Longleftrightarrow \quad & K(m^2,M^2,p)^{\frac{1}{2}}\|AB\|^{\frac{1}{p}} \le \|A^{\frac{1}{p}}B^{\frac{1}{p}}\| \qquad \text{by Theorem 2.13} \\
\Longleftrightarrow \quad & (B')
\end{aligned}$$

Similarly we have (c) $\Longleftrightarrow$ (c') and so the proof is complete. □

5.3 Norm inequality for the geometric mean

Let A and B be two positive operators on a Hilbert space. The arithmetic-geometric mean inequality says that

$$(1-\alpha)A+\alpha B \ge A \,\#_\alpha\, B \qquad \text{for all } 0 \le \alpha \le 1, \tag{5.19}$$

where the α-geometric mean $A \,\#_\alpha\, B$ is defined by

$$A \,\#_\alpha\, B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha A^{\frac{1}{2}} \qquad \text{for all } 0 \le \alpha \le 1.$$

In fact, put $C = A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Since $\alpha(t-1)+1 \ge t^\alpha$ for $t > 0$, we have $(1-\alpha)I_H + \alpha C \ge C^\alpha$. Therefore, we have (5.19).

On the other hand, it is known the following matrix Young inequality: For positive semi-definite matrices A, B and $p,q > 1$ such that $\frac{1}{p}+\frac{1}{q} = 1$

$$\frac{1}{p}A^p + \frac{1}{q}B^q \ge U^*|AB|U \tag{5.20}$$

for some unitary matrix U. By (5.20), for positive semi-definite matrices A,B

$$\|(1-\alpha)A+\alpha B\| \geq \|A^{1-\alpha}B^{\alpha}\| \qquad \text{for all } 0 \leq \alpha \leq 1 \tag{5.21}$$

and by (5.19) we have

$$\|(1-\alpha)A+\alpha B\| \geq \|A \#_{\alpha} B\| \qquad \text{for } 0 \leq \alpha \leq 1 \text{ and } A,B \geq 0.$$

Here we remark that McIntosh [182] proved that (5.21) holds for $\alpha = 1/2$ and positive operators.

In this section, we show a norm inequality and its converse on the geometric mean. In other words, we estimate $\|A \#_{\alpha} B\|$ by $\|A^{1-\alpha}B^{\alpha}\|$. Moreover we discuss it for the case $\alpha > 1$. Our main tools are the Araki-Cordes inequality (Theorem 5.9) and its converse one (Theorem 5.11).

We show the following norm inequality for the geometric mean, in which we use the Araki-Cordes inequality twice.

Theorem 5.15 *Let A and B be positive operators. Then for each* $0 \leq \alpha \leq 1$

$$\|A \#_{\alpha} B\| \leq \|A^{1-\alpha}B^{\alpha}\|. \tag{5.22}$$

Proof. It follows from (5.16) in Theorem 5.9 that

$$\|A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^{\alpha}A^{\frac{1}{2}}\| \leq \|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^{\alpha} = \|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^{\alpha}$$

for $0 \leq \alpha \leq 1$.

Furtheremore, if $\alpha \geq 1/2$, then by (5.16) in Theorem 5.9 again

$$\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^{\alpha} \leq \|A^{1-\alpha}B^{2\alpha}A^{1-\alpha}\|^{\frac{1}{2}} = \|A^{1-\alpha}B^{\alpha}\|.$$

Hence, if $1/2 \leq \alpha \leq 1$, then we have the desired inequality (5.22).

If $\alpha < 1/2$, then by using $A \#_{\alpha} B = B \#_{1-\alpha} A$, it reduces the proof to the case $\alpha \geq 1/2$ and so the proof is complete. □

As in Chapter 3, we use the notation ♮ to distinguish from the operator mean #:

$$A \natural_{\alpha} B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^{\alpha}A^{\frac{1}{2}} \qquad \text{for all } \alpha \notin [0,1].$$

Theorem 5.16 *Let A and B be positive operators. If* $3/2 \leq \alpha \leq 2$*, then*

$$\|A \natural_{\alpha} B\| \leq \|A^{1-\alpha}B^{\alpha}\|. \tag{5.23}$$

Proof. Put $\alpha = 1+\beta$ and $1/2 \le \beta \le 1$. Then we have

$$\begin{aligned}
\|A \natural_\alpha B\| &= \|B \natural_{-\beta} A\| = \|B^{\frac{1}{2}} \left(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\right)^{-\beta} B^{\frac{1}{2}}\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\le \|B^{\frac{1}{2\beta}} B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.9 and } 1/2 \le \beta \le 1 \\
&= \|B^{\frac{1+\beta}{2\beta}}A^{-1}B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\le \|B^{1+\beta}A^{-2\beta}B^{1+\beta}\|^{\frac{1}{2}} \qquad \text{by Theorem 5.9 and } 0 < \tfrac{1}{2\beta} \le 1 \\
&= \|A^{-\beta}B^{1+\beta}\| = \|A^{1-\alpha}B^{\alpha}\|.
\end{aligned}$$

□

Remark 5.6 *In Theorem 5.16, the inequality $\|A \natural_\alpha B\| \le \|A^{1-\alpha}B^\alpha\|$ does not always hold for $1 < \alpha < 3/2$. In fact, let $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Then we have $\|A \natural_{\frac{4}{3}} B\| =$ $3.38526 > \|A^{-\frac{1}{3}}B^{\frac{4}{3}}\| = 3.3759$. Also, $\|A \natural_{\frac{7}{5}} B\| = 3.49615 < \|A^{-\frac{2}{5}}B^{\frac{7}{5}}\| = 3.50464$.*

We show the following converse inequality of (5.22) in Theorem 5.15.

Theorem 5.17 *If A and B are positive operators such that $mI_H \le A, B \le MI_H$ for some scalars $0 < m \le M$ and $h = \frac{M}{m}$, then for each $0 \le \alpha \le 1$*

$$K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\| \le \|A \#_\alpha B\|,$$

where the generalized Kantorovich constant $K(h,\alpha)$ is defined by (2.31).

Proof. Suppose that $0 \le \alpha \le \frac{1}{2}$. Since $\frac{m}{M}I_H \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{M}{m}I_H$, it follows that a generalized condition number of $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ is $\frac{M}{m}/\frac{m}{M} = h^2$ and we have

$$\begin{aligned}
\|A \#_\alpha B\| &= \|(A^{\frac{1}{2\alpha}})^\alpha \left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha (A^{\frac{1}{2\alpha}})^\alpha\| \\
&\ge K(h^2,\alpha)\|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^\alpha \qquad \text{by Theorem 5.11 and } 0 \le \alpha \le \tfrac{1}{2} \\
&= K(h^2,\alpha)\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^\alpha \\
&\ge K(h^2,\alpha)\|A^{1-\alpha}B^{2\alpha}A^{1-\alpha}\|^{\frac{1}{2}} \qquad \text{by Theorem 5.9 and } \tfrac{1}{2\alpha} \ge 1 \\
&= K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\|.
\end{aligned}$$

Suppose that $\frac{1}{2} \le \alpha \le 1$. Since $0 \le 1-\alpha \le \frac{1}{2}$, we have

$$\begin{aligned}
\|A \#_\alpha B\| &= \|B \#_{1-\alpha} A\| \\
&\ge K(h^2,1-\alpha)\|B^{1-(1-\alpha)}A^{1-\alpha}\| \\
&= K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\| \qquad \text{by (ii) of Theorem 2.12}
\end{aligned}$$

and so the proof is complete. □

We show the following converse inequality of (5.23) in Theorem 5.16.

Theorem 5.18 *If A and B are positive operators such that $mI_H \le A, B \le MI_H$ for some scalars $0 < m \le M$ and $h = \frac{M}{m}$, then for each $\frac{3}{2} \le \alpha \le 2$*

$$K(h^2, \alpha - 1)K(h, 2(\alpha - 1))^{-\frac{1}{2}} \|A^{1-\alpha}B^{\alpha}\| \le \|A \natural_\alpha B\|,$$

where $K(h, \alpha)$ is defined by (2.31).

Proof. Put $\alpha = 1 + \beta$ and $1/2 \le \beta \le 1$. Then we have

$$\begin{aligned}
&\|A \natural_\alpha B\| = \|B \natural_{-\beta} A\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\ge K(h^2, \beta) \|B^{\frac{1}{2\beta}} B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.11 and } 1/2 \le \beta \le 1 \\
&= K(h^2, \beta) \|B^{\frac{1+\beta}{2\beta}} A^{-\frac{2\beta}{2\beta}} B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\ge K(h^2, \beta) \left(K\left(h^{-2\beta}, \frac{1}{2\beta}\right) \|B^{1+\beta} A^{-2\beta} B^{1+\beta}\|^{\frac{1}{2\beta}} \right)^{\beta} \qquad \text{by Theorem 5.11 and } 0 < \tfrac{1}{2\beta} \le 1 \\
&= K(h^2, \beta) K\left(h^{-2\beta}, \frac{1}{2\beta}\right)^{\beta} \|A^{-\beta} B^{1+\beta}\| \\
&= K(h^2, \beta) K(h, 2\beta)^{-\frac{1}{2}} \|A^{1-\alpha} B^{\alpha}\|.
\end{aligned}$$

The last equality follows from

$$K\left(h^{-2\beta}, \frac{1}{2\beta}\right)^{\beta} = K\left(h^{2\beta}, \frac{1}{2\beta}\right)^{\beta} = K(h, 2\beta)^{-\frac{\beta}{2\beta}} = K(h, 2\beta)^{-\frac{1}{2}}$$

by (i) of Theorem 2.12 and (i) of Theorem 2.13. □

As mentioned in Remark 5.6, we have no relation between $\|A \natural_\alpha B\|$ and $\|A^{1-\alpha}B^{\alpha}\|$ for $1 \le \alpha \le \frac{3}{2}$. We have the following result.

Theorem 5.19 *If A and B are positive operators such that $mI_H \le A, B \le MI_H$ for some scalars $0 < m \le M$ and $h = \frac{M}{m}$, then for each $1 \le \alpha \le \frac{3}{2}$*

$$K(h^2, \alpha - 1)\|A^{1-\alpha}B^{\alpha}\| \le \|A \natural_\alpha B\| \le K(h, 2(\alpha - 1))^{-\frac{1}{2}} \|A^{1-\alpha}B^{\alpha}\|,$$

where $K(h, \alpha)$ is defined by (2.31).

Proof. Put $\alpha = 1 + \beta$ and $0 \le \beta \le \frac{1}{2}$. Since a generalized condition number of $A^{-2\beta}$ is

$h^{-2\beta}$, it follows that

$$\begin{aligned}
\|A \natural_\alpha B\| &= \|B \natural_{-\beta} A\| = \|B^{\frac{1}{2}} \left(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\right)^{-\beta} B^{\frac{1}{2}}\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\leq \|B^{\frac{1}{2\beta}} B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.9 and } 0 \leq \beta \leq 1 \\
&= \|B^{\frac{1+\beta}{2\beta}}A^{-1}B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\leq \left(K\left(h^{-2\beta}, \frac{1}{2\beta}\right)\|B^{1+\beta}A^{-2\beta}B^{1+\beta}\|^{\frac{1}{2\beta}}\right)^{\beta} \qquad \text{by Theorem 5.11 and } 1 \leq \frac{1}{2\beta} \\
&= K(h, 2(\alpha-1))^{-\frac{1}{2}}\|A^{1-\alpha}B^{\alpha}\| \qquad \text{by Theorem 2.12 and 2.13.}
\end{aligned}$$

Also, we have

$$\begin{aligned}
\|A \natural_\alpha B\| &= \|B \natural_{-\beta} A\| = \|B^{\frac{1}{2}} \left(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\right)^{-\beta} B^{\frac{1}{2}}\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\geq K(h^2, \beta)\|B^{\frac{1}{2\beta}} B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.11 and } 0 \leq \beta \leq 1 \\
&= K(h^2, \beta)\|B^{\frac{1+\beta}{2\beta}}A^{-1}B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\geq K(h^2, \beta)\|B^{1+\beta}A^{-2\beta}B^{1+\beta}\|^{\frac{1}{2}} \qquad \text{by Theorem 5.9 and } \frac{1}{2\beta} \geq 1 \\
&= K(h^2, \alpha-1)\|A^{1-\alpha}B^{\alpha}\|
\end{aligned}$$

and so the proof is complete. □

Finally, we consider the case $\alpha \geq 2$:

Theorem 5.20 *If A and B are positive operators such that $mI_H \leq A, B \leq MI_H$ for some scalars $0 < m \leq M$ and $h = \frac{M}{m}$, then for each $\alpha \geq 2$*

$$K(h, 2(\alpha-1))^{-\frac{1}{2}}\|A^{1-\alpha}B^{\alpha}\| \leq \|A \natural_\alpha B\| \leq K(h^2, \alpha-1)\|A^{1-\alpha}B^{\alpha}\|,$$

where $K(h, \alpha)$ is defined by (2.31).

Proof. Put $\alpha = 1 + \beta$ and $\beta \geq 1$. Then we have

$$\begin{aligned}
\|A \natural_\alpha B\| &= \|B \natural_{-\beta} A\| = \|B^{\frac{1}{2}} \left(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\right)^{-\beta} B^{\frac{1}{2}}\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\leq K(h^2, \beta)\|B^{\frac{1}{2\beta}} B^{\frac{1}{2}}A^{-1}B^{\frac{1}{2}}B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.11 and } \beta \geq 1 \\
&= K(h^2, \beta)\|B^{\frac{1+\beta}{2\beta}}A^{-1}B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\leq K(h^2, \alpha-1)\|A^{1-\alpha}B^{\alpha}\| \qquad \text{by Theorem 5.9 and } 0 < \frac{1}{2\beta} \leq 1.
\end{aligned}$$

Also, it follows that

$$
\begin{aligned}
&\|A \natural_\alpha B\| = \|B \natural_{-\beta} A\| = \|B^{\frac{1}{2}} \left(B^{-\frac{1}{2}} A B^{-\frac{1}{2}}\right)^{-\beta} B^{\frac{1}{2}}\| \\
&= \|B^{\frac{1}{2}} \left(B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}}\right)^{\beta} B^{\frac{1}{2}}\| \\
&\geq \|B^{\frac{1}{2\beta}} B^{\frac{1}{2}} A^{-1} B^{\frac{1}{2}} B^{\frac{1}{2\beta}}\|^{\beta} \qquad \text{by Theorem 5.9 and } \beta \geq 1 \\
&= \|B^{\frac{1+\beta}{2\beta}} A^{-1} B^{\frac{1+\beta}{2\beta}}\|^{\beta} \\
&\geq K\left(h^{-2\beta}, \frac{1}{2\beta}\right)^{\beta} \|B^{1+\beta} A^{-2\beta} B^{1+\beta}\|^{\frac{1}{2}} \qquad \text{by Theorem 5.11 and } 0 < \tfrac{1}{2\beta} \leq 1 \\
&= K(h, 2(\alpha-1))^{-\frac{1}{2}} \|A^{1-\alpha} B^{\alpha}\| \qquad \text{by Theorem 2.12 and Theorem 2.13.}
\end{aligned}
$$

□

5.4 Norm inequality for the chaotically geometric mean

Let A and B be two positive invertible operators on a Hilbert space H. We recall that the chaotically geometric mean $A \diamondsuit_\alpha B$ for all $\alpha \in \mathbb{R}$ is defined by

$$A \diamondsuit_\alpha B = \exp\left((1-\alpha)\log A + \alpha \log B\right).$$

If A and B commute, then $A \diamondsuit_\alpha B = A^{1-\alpha} B^{\alpha}$ for all $\alpha \in \mathbb{R}$.

First of all, we recall the following Ando-Hiai inequality (Theorem 3.4).

Theorem AH *If A and B are positive operators, then for each $\alpha \in [0,1]$*

$$\|A^r \#_\alpha B^r\| \leq \|A \#_\alpha B\|^r \qquad \textit{for all } r \geq 1 \tag{5.24}$$

or equivalently

$$A \#_\alpha B \leq I_H \quad \Longrightarrow \quad A^r \#_\alpha B^r \leq I_H \qquad \textit{for all } r \geq 1.$$

The following result is a geometric mean version of the Lie-Trotter formula

$$\exp(A+B) = \lim_{n\to\infty} \left(\exp\left(\frac{A}{n}\right)\exp\left(\frac{B}{n}\right)\right)^n \tag{5.25}$$

for self-adjoint operators A and B.

Lemma 5.5 *If A and B are self-adjoint operators, then*

$$\exp((1-\alpha)A+\alpha B) = \lim_{r\to+0} (\exp(rA) \natural_\alpha \exp(rB))^{\frac{1}{r}}$$

in the operator norm topology for all $\alpha \in \mathbb{R}$

Proof. For $0 < r < 1$ and $\alpha \in \mathbb{R}$, let $X(r) = \exp(rA) \natural_\alpha \exp(rB)$, $Y(r) = \exp(r[(1-\alpha)A+\alpha B])$, and $r^{-1} = m+s$, where $m \in \mathbb{N}$ and $s \in [0,1)$. It is enough to prove that $\|X(r)^m - Y(r)^m\| \to 0$. Since, with the convention $o(r)/r \to 0$ as $r \to 0$,

$$\begin{aligned}
X(r) &= \exp\Big(\frac{rA}{2}\Big) \left(\sum_{k=0}^{\infty} \frac{1}{k!}\Big(-\frac{rA}{2}\Big)^k \sum_{k=0}^{\infty} \frac{(rB)^k}{k!} \sum_{k=0}^{\infty} \frac{1}{k!}\Big(-\frac{rA}{2}\Big)^k\right)^{\alpha} \exp\Big(\frac{rA}{2}\Big) \\
&= \exp\Big(\frac{rA}{2}\Big) (I_H + r(B-A)+o(r))^{\alpha} \exp\Big(\frac{rA}{2}\Big) \\
&= \left(I_H + \frac{rA}{2} + o(r)\right) (I_H + r\alpha(B-A)+o(r)) \left(I_H + \frac{rA}{2} + o(r)\right) \\
&= I_H + r[(1-\alpha)A+\alpha B] + o(r),
\end{aligned}$$

we get $X(r) - Y(r) = o(r)$. Since

$$X(r)^m - Y(r)^m = \sum_{j=0}^{m-1} X(r)^{m-j-1}(X(r)-Y(r))Y(r)^j,$$

it follows that

$$\begin{aligned}
\|X(r)^m - Y(r)^m\| &\le m\|X(r)-Y(r)\| \max\{\|X(r)\|, \|Y(r)\|\}^{m-1} \\
&\le \frac{1}{r}\|X(r)-Y(r)\| \exp((1-\alpha)\|A\| + \alpha\|B\|) \quad \to 0 \qquad \text{as } r \to 0.
\end{aligned}$$

□

By Lemma 5.5, we have the following formula for the chaotically geometric mean, which is an extension of Theorem 3.21.

Theorem 5.21 *Let A and B be positive invertible operators. Then for each* $\alpha \in [0,1]$

$$A \diamondsuit_\alpha B = \lim_{r\to+0} (A^r \#_\alpha B^r)^{\frac{1}{r}} \tag{5.26}$$

in the operator norm topology. Moreover, for each $\alpha \notin [0,1]$

$$A \diamondsuit_\alpha B = \lim_{r\to+0} (A^r \natural_\alpha B^r)^{\frac{1}{r}}.$$

We show the following norm inequality for the geometric mean, in which we use the Ando-Hiai inequality.

Theorem 5.22 *Let A and B be positive invertible operators. Then for each* $\alpha \in [0,1]$

$$\|A \#_\alpha B\| \le \|A \diamondsuit_\alpha B\| \le \|A^{1-\alpha}B^\alpha\|.$$

Proof. It follows from (5.24) in Theorem AH that

$$\|A \#_\alpha B\| \le \|A^r \#_\alpha B^r\|^{\frac{1}{r}} \qquad \text{for all } 0 < r < 1.$$

As $r \to 0$, we have $\|A \#_\alpha B\| \le \|A \diamondsuit_\alpha B\|$ by Theorem 5.21.

By Lie-Trotter formula (5.25) and the Cordes inequality (Theorem 5.10), we have

$$\|\exp(H+K)\| = \lim_{n\to\infty} \|\exp\left(\frac{H}{n}\right)\exp\left(\frac{K}{n}\right)\|^n \le \|\exp H \exp K\|$$

for self-adjoint operators H and K. Hence it follows that

$$\begin{aligned}\|A \diamondsuit_\alpha B\| &= \|\exp((1-\alpha)\log A + \alpha \log B)\| \\ &\le \|\exp\log A^{1-\alpha} \exp\log B^\alpha\| \\ &= \|A^{1-\alpha}B^\alpha\|.\end{aligned}$$

□

Remark 5.7 *By the proof above, we have*

$$\|A \diamondsuit_\alpha B\| \le \|A^{1-\alpha}B^\alpha\| \qquad \textit{for all } \alpha \in \mathbb{R}.$$

We show the following converse inequality for Theorem 5.22.

Theorem 5.23 *If A and B are positive operators such that* $mI_H \le A, B \le MI_H$ *for some scalars* $0 < m < M$ *and* $h = \frac{M}{m}$, *then for each* $0 \le \alpha \le 1$

$$K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\| \le \|A \#_\alpha B\|, \tag{5.27}$$

where $K(h,\alpha)$ *is defined by* (2.31).

Proof. Suppose that $0 \le \alpha \le \frac{1}{2}$. Since $\frac{m}{M}I_H \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{M}{m}I_H$, it follows that a generalized condition number of $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ is $\frac{M}{m}/\frac{m}{M} = h^2$ and by Theorem 5.9 we have $\|B^pA^pB^p\| \le \|(BAB)^p\|$ for all $p \in [0,1]$ and the opposite inequality holds for all $p > 1$. Hence it follows that

$$\begin{aligned}\|A \#_\alpha B\| &= \|(A^{\frac{1}{2\alpha}})^\alpha \left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha (A^{\frac{1}{2\alpha}})^\alpha\| \\ &\ge K(h^2,\alpha)\|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^\alpha \qquad \text{by } 0 \le \alpha \le \tfrac{1}{2} \\ &= K(h^2,\alpha)\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^\alpha \\ &\ge K(h^2,\alpha)\|A^{1-\alpha}B^{2\alpha}A^{1-\alpha}\|^{\frac{1}{2}} \qquad \text{by } \tfrac{1}{2\alpha} \ge 1 \\ &= K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\|.\end{aligned}$$

Suppose that $\frac{1}{2} \le \alpha \le 1$. Since $0 \le 1-\alpha \le \frac{1}{2}$ and $K(h,1-\alpha) = K(h,\alpha)$ by (i) of Theorem 2.13, we have

$$\begin{aligned}\|A \#_\alpha B\| &= \|B \#_{1-\alpha} A\| \\ &\ge K(h^2,1-\alpha)\|B^{1-(1-\alpha)}A^{1-\alpha}\| \\ &= K(h^2,\alpha)\|A^{1-\alpha}B^\alpha\|\end{aligned}$$

and so the proof is complete. □

We show the following complement of the Ando-Hiai inequality.

Theorem 5.24 *Let A and B be positive operators on H such that $mI_H \le A, B \le MI_H$ for some scalars $0 < m \le M$, $h = \frac{M}{m}$ and $0 \le \alpha \le 1$. Then*

$$\|A^r \#_\alpha B^r\| \le K(h^2,\alpha)^{-r}\|A \#_\alpha B\|^r \qquad \text{for all } 0 < r < 1 \tag{5.28}$$

or equivalently

$$A \#_\alpha B \le I_H \implies A^r \#_\alpha B^r \le K(h^2,\alpha)^{-r} \qquad \text{for all } 0 < r < 1, \tag{5.29}$$

where $K(h,\alpha)$ is defined by (2.31).

Proof. We firstly show (5.28). Since a generalized condition number of $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ is $h^2 = \left(\frac{M}{m}\right)^2$, it follows from Theorem 5.9 that for each $0 \le \alpha \le 1$

$$\begin{aligned}\|A^r \#_\alpha B^r\| &= \|A^{\frac{r}{2}}\left(A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}\right)^\alpha A^{\frac{r}{2}}\| \\ &\le \|A^{\frac{r}{2\alpha}}A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}A^{\frac{r}{2\alpha}}\|^\alpha \qquad \text{by } 0 \le \alpha \le 1 \\ &= \|A^{\frac{r-r\alpha}{2\alpha}}B^rA^{\frac{r-r\alpha}{2\alpha}}\|^\alpha \\ &\le \|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^{r\alpha} \qquad \text{by } 0 < r < 1 \\ &= \|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^{r\alpha} \\ &\le \left(K(h^2,\alpha)^{-1}\|A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha A^{\frac{1}{2}}\|\right)^r \qquad \text{by } 0 \le \alpha \le 1 \\ &= K(h^2,\alpha)^{-r}\|A \#_\alpha B\|^r\end{aligned}$$

for all $0 < r < 1$ and hence we have the desired inequality (5.28).
(5.28) $\Longrightarrow$ (5.29): is obvious.
(5.29) $\Longrightarrow$ (5.28): Since $A \#_\alpha B \le \|A \#_\alpha B\|$, it follows from the homogeneity of the geometric mean that

$$\frac{A}{\|A \#_\alpha B\|} \#_\alpha \frac{B}{\|A \#_\alpha B\|} \le I_H.$$

By (5.29), we have

$$\frac{A^r}{\|A \#_\alpha B\|^r} \#_\alpha \frac{B^r}{\|A \#_\alpha B\|^r} \le K(h^2,\alpha)^{-r},$$

because generalized condition numbers of both $A/\|A \#_\alpha B\|$ and $B/\|A \#_\alpha B\|$ coincides with $\frac{M}{\|A \#_\alpha B\|}/\frac{m}{\|A \#_\alpha B\|} = M/m = h$. Hence we have the desired inequality:

$$\|A^r \#_\alpha B^r\| \le K(h^2,\alpha)^{-r}\|A \#_\alpha B\|^r$$

for all $0 < r < 1$. Therefore the proof is complete. □

By Theorem 5.24, we have the following converse inequality of the Ando-Hiai one for the case $r > 1$.

Corollary 5.2 *Let A and B be positive operators such that $mI_H \le A, B \le MI_H$ for some scalars $0 < m < M$, $h = \frac{M}{m}$ and $0 \le \alpha \le 1$. Then*

$$K(h^{2r},\alpha)\|A \#_\alpha B\|^r \le \|A^r \#_\alpha B^r\| (\le \|A \#_\alpha B\|^r) \qquad \text{for all } r > 1. \tag{5.30}$$

Proof. For $r > 1$, we have $0 < \frac{1}{r} < 1$ and by (5.28) in Theorem 5.24

$$\|A^{\frac{1}{r}} \#_\alpha B^{\frac{1}{r}}\| \le K(h^2,\alpha)^{-\frac{1}{r}}\|A \#_\alpha B\|^{\frac{1}{r}}.$$

Replacing A and B by A^r and B^r respectively, and a generalized condition number of A^r and B^r is h^r, it follows that

$$\|A \#_\alpha B\| \le K(h^{2r},\alpha)^{-\frac{1}{r}}\|A^r \#_\alpha B^r\|^{\frac{1}{r}}$$

and by taking r-th power on both sides we have the desired inequality (5.30). □

In the remainder of the section, we investigate the Ando-Hiai inequality without the framework of operator mean. The following theorem corresponds to (5.28) in Theorem 5.24 in the case $\alpha > 1$.

Theorem 5.25 *Let A and B be positive operators such that $0 < mI_H \le A, B \le MI_H$ for some scalars $0 < m < M$, $h = \frac{M}{m}$ and $\alpha > 1$. Then*

$$K(h,r)^\alpha K(h^2,\alpha)^{-r}\|A \natural_\alpha B\|^r \le \|A^r \natural_\alpha B^r\| \le K(h^{2r},\alpha)\|A \natural_\alpha B\|^r \tag{5.31}$$

for all $0 < r < 1$, where $K(h,\alpha)$ is defined by (2.31)*.*

Proof. Since $\|B^pA^pB^p\| \le \|(BAB)^p\|$ for all $p \in [0,1]$ and the opposite inequality holds for all $p > 1$ by Theorem 5.9, for each $\alpha > 1$, we have

$$\begin{aligned}
\|A^r \natural_\alpha B^r\| &= \|A^{\frac{r}{2}}\left(A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}\right)^\alpha A^{\frac{r}{2}}\| \\
&\le K(h^{2r},\alpha)\|A^{\frac{r}{2\alpha}}A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}A^{\frac{r}{2\alpha}}\|^\alpha \qquad \text{by } \alpha > 1 \\
&= K(h^{2r},\alpha)\|A^{\frac{r-r\alpha}{2\alpha}}B^rA^{\frac{r-r\alpha}{2\alpha}}\|^\alpha \\
&\le K(h^{2r},\alpha)\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^{r\alpha} \qquad \text{by } 0 < r < 1 \\
&= K(h^{2r},\alpha)\|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^{r\alpha} \\
&\le K(h^{2r},\alpha)\|A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha A^{\frac{1}{2}}\|^r \qquad \text{by } \alpha > 1 \\
&= K(h^{2r},\alpha)\|A \natural_\alpha B\|^r
\end{aligned}$$

and hence we have the right-hand side of (5.31).

Conversely, we have

$$\begin{aligned}
\|A^r \natural_\alpha B^r\| &= \|A^{\frac{r}{2}}\left(A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}\right)^\alpha A^{\frac{r}{2}}\| \\
&\geq \|A^{\frac{r}{2\alpha}}A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}}A^{\frac{r}{2\alpha}}\|^\alpha \qquad \text{by } \alpha > 1 \\
&= \|A^{\frac{r-r\alpha}{2\alpha}}B^rA^{\frac{r-r\alpha}{2\alpha}}\|^\alpha \\
&\geq \left(K(h,r)\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^r\right)^\alpha \qquad \text{by } 0 < r < 1 \\
&= K(h,r)^\alpha\|A^{\frac{1-\alpha}{2\alpha}}BA^{\frac{1-\alpha}{2\alpha}}\|^{r\alpha} \\
&= K(h,r)^\alpha\|A^{\frac{1}{2\alpha}}A^{-\frac{1}{2}}BA^{-\frac{1}{2}}A^{\frac{1}{2\alpha}}\|^{r\alpha} \\
&\geq K(h,r)^\alpha\left(K(h^2,\alpha)^{-1}\|A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^\alpha A^{\frac{1}{2}}\|\right)^r \qquad \text{by } \alpha > 1 \\
&= K(h,r)^\alpha K(h^2,\alpha)^{-r}\|A \natural_\alpha B\|^r
\end{aligned}$$

and hence we have the left-hand side of (5.31). □

By Theorem 5.25, we have the following complement of the Ando-Hiai inequality in the case $\alpha > 1$.

Theorem 5.26 *Let A and B be positive operators such that* $0 < mI_H \leq A, B \leq MI_H$ *for some scalars* $0 < m < M$, $h = \frac{M}{m}$ *and* $\alpha > 1$. *Then*

$$\|A^r \natural_\alpha B^r\| \leq K(h^{2r},\alpha)\|A \natural_\alpha B\|^r \qquad \text{for all } 0 < r < 1 \tag{5.32}$$

or equivalently

$$A \natural_\alpha B \leq I_H \implies A^r \natural_\alpha B^r \leq K(h^{2r},\alpha) \qquad \text{for all } 0 < r < 1, \tag{5.33}$$

where $K(h,\alpha)$ *is defined by* (2.31).

The following corollary is a complementary result for Theorem 5.25.

Corollary 5.3 *Let A and B be positive operators such that* $0 < mI_H \leq A, B \leq MI_H$ *for some scalars* $0 < m < M$, $h = \frac{M}{m}$ *and* $\alpha > 1$. *Then*

$$K(h^2,\alpha)^{-r}\|A \natural_\alpha B\|^r \leq \|A^r \natural_\alpha B^r\| \leq K(h,r)^\alpha K(h^{2r},\alpha)\|A \natural_\alpha B\|^r \tag{5.34}$$

for all $r > 1$.

Next, we show converse norm inequalities for the α-geometric mean and the chaotically geometric one.

Theorem 5.27 *If A and B are positive operators such that* $0 < mI_H \leq A, B \leq MI_H$ *for some scalars* $0 < m < M$ *and* $h = \frac{M}{m}$, *then*

$$K(h^2,\alpha)\|A \diamondsuit_\alpha B\| \leq \|A \#_\alpha B\| \qquad \text{for all } 0 < \alpha < 1. \tag{5.35}$$

$$S(h)^{-\alpha}K(h^2,\alpha)^{-1}\|A \natural_\alpha B\| \leq \|A \diamondsuit_\alpha B\| \leq h^{2(\alpha-1)}\|A \natural_\alpha B\| \qquad \text{for all } \alpha > 1, \tag{5.36}$$

where the Specht ratio $S(h)$ *is defined by* (2.35).

Proof. By (5.28) in Theorem 5.24, it follows that for each $0 < \alpha < 1$

$$\|A^r \#_\alpha B^r\| \le K(h^2,\alpha)^{-r}\|A \#_\alpha B\|^r \qquad \text{for all } 0 < r < 1.$$

By taking $\frac{1}{r}$-th power on both sides, we have

$$\|(A^r \#_\alpha B^r)^{\frac{1}{r}}\| \le K(h^2,\alpha)^{-1}\|A \#_\alpha B\|$$

and hence we have the desired inequality (5.35)

$$\|A \diamondsuit_\alpha B\| \le K(h^2,\alpha)^{-1}\|A \#_\alpha B\|$$

by the formula (5.26) in Theorem 5.21.

Next, since $K(h^{2r},\alpha) \le (h^{2r})^{\alpha-1}$ by Theorem 2.13, it follows from (5.31) in Theorem 5.25 that for each $\alpha > 1$

$$\begin{aligned} K(h,r)^{\frac{\alpha}{r}} K(h^2,\alpha)^{-1}\|A \natural_\alpha B\| \le \|A^r \natural_\alpha B^r\|^{\frac{1}{r}} &\le K(h^{2r},\alpha)^{\frac{1}{r}}\|A \natural_\alpha B\| \\ &\le h^{2(\alpha-1)}\|A \natural_\alpha B\| \qquad \text{for all } 0 < r < 1. \end{aligned}$$

On the other hand, since $K(h^r,\frac{1}{r}) = K(h,r)^{-\frac{1}{r}}$ in the case of $p = 1$ by (i) of Theorem 2.13 and $K(h^r,\frac{1}{r}) \to S(h)$ as $r \to 0$ by (i) of Theorem 2.17, we have

$$\lim_{r\to 0} K(h,r)^{\frac{\alpha}{r}} = \lim_{r\to 0} K\left(h^r,\frac{1}{r}\right)^{-\alpha} = S(h)^{-\alpha}.$$

By Theorem 5.21 we have $(A^r \natural_\alpha B^r)^{\frac{1}{r}} \to A \diamondsuit_\alpha B$ as $r \to 0$ and hence

$$S(h)^{-\alpha}K(h^2,\alpha)^{-1}\|A \natural_\alpha B\| \le \|A \diamondsuit_\alpha B\| \le h^{2(\alpha-1)}\|A \natural_\alpha B\|.$$

□

Finally, we show a slight improvement of Theorem 5.23 for the chaotically geometric mean and its converse. The following lemma shows the Golden-Thompson type inequality for the operator norm and its converse.

Lemma 5.6 *Let A and B be self-adjoint operators such that $mI_H \le B \le MI_H$ for some scalars $m < M$. Then*

$$S(e^{M-m})^{-1}\|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\| \le \|e^{A+B}\| \le \|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\|,$$

where $S(e^{M-m})$ is the Specht ratio defined by (2.35).

Proof. Since $0 < e^m \le e^B \le e^M$ and a generalized condition number of e^B is e^{M-m}, it follows from Theorem 5.9 and Theorem 5.11 that

$$K(e^{M-m},p)\|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\|^p \le \|e^{\frac{pA}{2}}e^{pB}e^{\frac{pA}{2}}\| \le \|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\|^p \qquad \text{for all } p \in [0,1].$$

Taking $\frac{1}{p}$-th power of both sides, we have

$$K(e^{M-m},p)^{\frac{1}{p}}\|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\| \leq \|e^{\frac{pA}{2}}e^{pB}e^{\frac{pA}{2}}\|^{\frac{1}{p}} \leq \|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\|. \tag{5.37}$$

It follows from (i) of Theorem 2.13 and (i) of Theorem 2.17 that

$$K\left(e^{M-m},p\right)^{\frac{1}{p}} = K\left(e^{pM-pm},\frac{1}{p}\right)^{-1} \to S(e^{M-m})^{-1} \qquad \text{as } p \to 0.$$

By the Lie-Trotter formula, we have $\|e^{\frac{pA}{2}}e^{pB}e^{\frac{pA}{2}}\|^{\frac{1}{p}} \to \|e^{A+B}\|$ as $p \to 0$ and hence by (5.37) it follows that

$$S(e^{M-m})^{-1}\|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\| \leq \|e^{A+B}\| \leq \|e^{\frac{A}{2}}e^{B}e^{\frac{A}{2}}\|,$$

as desired. □

By Lemma 5.6, we have the following theorem which is a slight improvement of Theorem 5.23.

Theorem 5.28 *Let A and B be strictly positive operators such that $0 < mI_H \leq B \leq MI_H$ for some scalars $0 < m < M$, $h_B = \frac{M}{m}$. Then for each real number $\alpha \in \mathbb{R}$*

$$S(h_B^{\alpha})^{-1}\|A^{\frac{1-\alpha}{2}}B^{\alpha}A^{\frac{1-\alpha}{2}}\| \leq \|A \diamondsuit_{\alpha} B\| \leq \|A^{\frac{1-\alpha}{2}}B^{\alpha}A^{\frac{1-\alpha}{2}}\|,$$

where $S(h)$ is the Specht ratio defined by (2.35).

Proof. For each $\alpha > 0$, replacing A and B by $(1-\alpha)\log A$ and $\alpha \log B$ respectively in Lemma 5.6, we have the desired inequality since $\alpha \log m \leq \alpha \log B \leq \alpha \log M$ and $e^{\alpha \log M - \alpha \log m} = h_B^{\alpha}$. In the case of $\alpha < 0$, we have $\alpha \log M \leq \alpha \log B \leq \alpha \log m$ and $e^{\alpha \log m - \alpha \log M} = h_B^{-\alpha}$. By the property of the Specht ratio in Theorem 2.16, it follows that $S(h_B^{-\alpha}) = S(h_B^{\alpha})$ and hence we have this theorem. □

The following corollary is a complementary result for Theorem 5.23.

Corollary 5.4 *Let A and B be positive operators such that $mI_H \leq A \leq MI_H$ for some scalars $0 < m < M$, $h_A = \frac{M}{m}$. Then for each real number $\alpha \in \mathbb{R}$*

$$S(h_A^{1-\alpha})^{-1}\|A^{\frac{1-\alpha}{2}}B^{\alpha}A^{\frac{1-\alpha}{2}}\| \leq \|A \diamondsuit_{\alpha} B\| \leq \|A^{\frac{1-\alpha}{2}}B^{\alpha}A^{\frac{1-\alpha}{2}}\|.$$

Proof. If we apply $B \diamondsuit_{1-\alpha} A$ to Theorem 5.23, then it follows that

$$S(h_A^{1-\alpha})^{-1}\|B^{\frac{\alpha}{2}}A^{1-\alpha}B^{\frac{\alpha}{2}}\| \leq \|B \diamondsuit_{1-\alpha} A\|$$

and hence we have this corollary. □

Remark 5.8 *Let A and B be positive operators such that $mI_H \leq A, B \leq MI_H$ for some scalars $0 < m < M$, $h = \frac{M}{m}$. Since $\|A^{\frac{1-\alpha}{2}} B^\alpha A^{\frac{1-\alpha}{2}}\| \leq \|A^{1-\alpha} B^\alpha\|$, the expression in Theorem 5.22 implies*

$$K(h^2, \alpha)\|A^{\frac{1-\alpha}{2}} B^\alpha A^{\frac{1-\alpha}{2}}\| \leq \|A \diamondsuit_\alpha B\| \qquad \text{for all } \alpha \in [0,1]. \tag{5.38}$$

By combing Theorem 5.28 and Corollary 5.4, we have

$$\max\{S(h^\alpha)^{-1}, S(h^{1-\alpha})^{-1}\} \|A^{\frac{1-\alpha}{2}} B^\alpha A^{\frac{1-\alpha}{2}}\| \leq \|A \diamondsuit_\alpha B\| \qquad \text{for all } \alpha \in \mathbb{R}. \tag{5.39}$$

Then (5.39) is an improvement of (5.38). As a matter of fact, we have

$$K(h^2, \alpha) \leq S(h^\alpha)^{-1} \qquad \text{for all } 0 \leq \alpha \leq \tfrac{1}{2}. \tag{i}$$

$$K(h^2, \alpha) \leq S(h^{1-\alpha})^{-1} \qquad \text{for all } \tfrac{1}{2} \leq \alpha \leq 1. \tag{ii}$$

To prove (i), *it is sufficient to show $K(h,\alpha)^{-1} \geq S(h^{\frac{\alpha}{2}})$ for all $0 \leq \alpha \leq \frac{1}{2}$. By* (ii) *of Theorem 2.12,* (i) *of Theorem 2.13 and* (ii) *of Theorem 2.17, we have*

$$\begin{aligned} K(h,\alpha)^{-1} &= K(h, 1-\alpha)^{-1} = K\left(h^{1-\alpha}, \frac{1}{1-\alpha}\right)^{1-\alpha} \\ &= K\left(h^{1-\alpha}, \frac{\alpha + 1 - \alpha}{1-\alpha}\right)^{1-\alpha} \geq S(h^\alpha)^{1-\alpha}. \end{aligned}$$

Since $S(h^s)^{\frac{1}{s}}$ is decreasing for $0 \leq s \leq 1$ by Lemma 4.4, it follows that $S(h^\alpha) \geq S(h^{\frac{\alpha}{2}})^2$ and hence we have

$$S(h^\alpha)^{1-\alpha} \geq S(h^{\frac{\alpha}{2}})^{2(1-\alpha)} \geq S(h^{\frac{\alpha}{2}})$$

since $0 \leq \alpha \leq \frac{1}{2}$. Therefore, it follows that $K(h^2,\alpha) \leq S(h^\alpha)^{-1}$ for all $0 \leq \alpha \leq \frac{1}{2}$. Similarly, we have (ii). *Therefore we have*

$$K(h^2, \alpha) \leq \max\left\{S(h^\alpha)^{-1}, S(h^{1-\alpha})^{-1}\right\} \qquad \text{for all } \alpha \in [0,1].$$

5.5 Complement of the Ando-Hiai inequality

The following theorem is an operator norm version of a generalized Ando-Hiai inequality (GAH) in Theorem 3.5.

Theorem 5.29 *Let A and B be positive operators, and let $\alpha \in [0,1]$. Then*

$$\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \leq \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \qquad \text{for all } r, s \geq 1, \tag{5.40}$$

or equivalently

$$\|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \leq \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \qquad \text{for all } 0 < r,s \leq 1. \tag{5.41}$$

For each $\alpha \in [0,1]$, we have no relation between $\|A\#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}$ and $\|A^r\#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\|$ for $r,s > 0$ without constraint $(r-1)(s-1) \geq 0$. In this section, we show a complement of the generalized Ando-Hiai inequality. We attempt to find upper and lower bounds for $\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\|$ by means of scalar multiples of $\|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}$, that is, for each $0 < r \leq 1$ and $s \geq 1$ there exist constants β and γ such that

$$\beta\, \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \leq \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \leq \gamma\, \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}$$

for two positive operators A and B.

First of all, in the case of $r \geq 1$ and $s \geq 1$, we estimate a lower bound of the generalized Ando-Hiai inequality (5.40):

Theorem 5.30 *Let A and B be positive operators on a Hilbert space H such that $m_1 I_H \leq A \leq M_1 I_H$ and $m_2 I_H \leq B \leq M_2 I_H$ for some scalars $0 < m_i \leq M_i$ $(i = 1,2)$, and let $\alpha \in [0,1]$. Put $h_i = \frac{M_i}{m_i}$ for $i = 1,2$. Then for each $r \geq 1$ and $s \geq 1$*

$$\begin{aligned} &K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \\ &\qquad \leq \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \;\left(\leq \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \right), \end{aligned}$$

or equivalently for each $0 < r \leq 1$ and $0 < s \leq 1$

$$\begin{aligned} &\left(\|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \leq \right) \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \\ &\qquad \leq K(h_1 h_2, \alpha)^{\frac{-rs}{(1-\alpha)s+\alpha r}} \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}. \end{aligned}$$

Proof. It follows that the generalized condition number of $A^{-\frac{r}{2}} B^s A^{-\frac{r}{2}}$ is $h_1^r h_2^s$ since $m_2^r / M_1^r I_H \leq A^{-\frac{r}{2}} B^s A^{-\frac{r}{2}} \leq M_2^s / m_1^r I_H$. By Theorem 5.11, we have

$$\begin{aligned} \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| &= \|A^{\frac{r}{2}} \left(A^{-\frac{r}{2}} B^s A^{-\frac{r}{2}}\right)^{\frac{\alpha r}{(1-\alpha)s+\alpha r}} A^{\frac{r}{2}}\| \\ &\geq K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A^{\frac{(1-\alpha)s}{2\alpha}} B^s A^{\frac{(1-\alpha)s}{2\alpha}}\|^{\frac{\alpha r}{(1-\alpha)s+\alpha r}} \quad \text{by } \tfrac{\alpha r}{(1-\alpha)s+\alpha r} \in [0,1] \text{ and (5.17)} \\ &\geq K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A^{\frac{1-\alpha}{2\alpha}} B A^{\frac{1-\alpha}{2\alpha}}\|^{\frac{\alpha rs}{(1-\alpha)s+\alpha r}} \quad \text{by } s \geq 1 \text{ and (5.18)} \\ &\geq K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A^{\frac{1}{2\alpha}} A^{-\frac{1}{2}} B A^{-\frac{1}{2}} A^{\frac{1}{2\alpha}}\|^{\frac{\alpha rs}{(1-\alpha)s+\alpha r}} \\ &\geq K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A^{\frac{1}{2}} \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)^{\alpha} A^{\frac{1}{2}}\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \quad \text{by } 0 \leq \alpha \leq 1 \text{ and (5.16)} \\ &= K\left(h_1^r h_2^s, \tfrac{\alpha r}{(1-\alpha)s+\alpha r}\right) \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}. \end{aligned}$$

□

If we put $r = s$ in Theorem 5.30, then we have the following converse of the Ando-Hiai inequality (5.28) in Theorem 5.24.

Corollary 5.5 *Suppose that the same conditions of Theorem 5.30 hold. Then for each* $r \geq 1$

$$K(h_1^r h_2^r, \alpha)\ \|A \#_\alpha B\|^r \ \leq\ \|A^r \#_\alpha B^r\|\ (\ \leq\ \|A \#_\alpha B\|^r\ ),$$

or equivalently for each $0 < r \leq 1$

$$(\ \|A \#_\alpha B\|^r\ \leq\ )\ \|A^r \#_\alpha B^r\|\ \leq\ K(h_1 h_2, \alpha)^{-r}\ \|A \#_\alpha B\|^r.$$

We remark that in Corollary 5.5 the constant $K(h_1^r h_2^r, \alpha) = 1$ in the case of $\alpha = 0, 1$ and $K(h_1^r h_2^r, \alpha) \neq 1$ in the case of $r = 1$ and $0 < \alpha < 1$.

Remark 5.9 *For* $\alpha \in [0,1]$*, the generalized Ando-Hiai inequality*

$$\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \leq \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}$$

does not always hold for $0 < r \leq 1$ *and* $s \geq 1$*. In fact, put* $A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}$ *and* $B = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$*. Then for* $\alpha = \frac{1}{2}$ *we have* $\|A^{\frac{1}{2}} \#_{\frac{1}{5}} B^2\| = 4.798011 > \|A \#_{\frac{1}{2}} B\|^{\frac{4}{5}} = 4.795148$ *in the case of* $r = \frac{1}{2}$ *and* $s = 2$*. Also,* $\|A^{\frac{1}{2}} \#_{\frac{1}{9}} B^4\| = 5.514677 < \|A \#_{\frac{1}{2}} B\|^{\frac{8}{9}} = 5.707511$ *in the case of* $r = \frac{1}{2}$ *and* $s = 4$*.*

At the end of this section, we present a complementary inequality to the generalized Ando-Hiai inequality for the case $0 < r \leq 1$ and $s \geq 1$. The following theorem gives estimates of both upper and lower bounds of $\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\|$ by means of scalars multiples of $\|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}$.

Theorem 5.31 *Let A and B be positive operators on a Hilbert space H such that* $m_1 I_H \leq A \leq M_1 I_H$ *and* $m_2 I_H \leq B \leq M_2 I_H$ *for some scalars* $0 < m_i \leq M_i$ $(i = 1,2)$*, and let* $\alpha \in [0,1]$*. Put* $h_i = \frac{M_i}{m_i}$ *for* $i = 1,2$*. Then for each* $0 < r \leq 1$ *and* $s \geq 1$

$$\begin{aligned} &K\left(h_1^s, \tfrac{r}{s}\right)^{\frac{(1-\alpha)s}{(1-\alpha)s+\alpha r}} K\left(h_1^{s\alpha} h_2^{s\alpha}, \tfrac{r}{(1-\alpha)s+\alpha r}\right) h_2^{\frac{(1-\alpha)s(r-s)}{(1-\alpha)s+\alpha r}}\ \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \\ &\leq \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| \\ &\leq h_2^{\frac{(1-\alpha)s(s-r)}{(1-\alpha)s+\alpha r}}\ \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}. \end{aligned}$$

In order to prove this theorem, we need the following lemma.

Lemma 5.7 *Let A and B be positive operators on a Hilbert space H such that* $m_1 I_H \leq A \leq M_1 I_H$ *and* $m_2 I_H \leq B \leq M_2 I_H$ *for some scalars* $0 < m_i \leq M_i$ $(i = 1,2)$*, and let* $\alpha \in [0,1]$*. Put* $h_i = \frac{M_i}{m_i}$ *for* $i = 1,2$*. Then for each* $0 < s \leq 1$

$$K(h_2, s) \|A^{-2}\|^{s-1} (ABA)^s \leq AB^s A\ \leq\ \|A^2\|^{1-s} (ABA)^s. \tag{5.42}$$

Proof. Since $f(t) = t^s$ for $0 < s \le 1$ is operator concave, it follows that

$$AB^sA = \|A^2\| \frac{A}{\|A\|} B^s \frac{A}{\|A\|} \le \|A^2\| \left(\frac{A}{\|A\|} B \frac{A}{\|A\|} \right)^s = \|A^2\|^{1-s}(ABA)^s,$$

and we have the right-hand side of (5.42). Next, we show the left-hand side of (5.42). By (2.11) in Theorem 2.11, we have

$$\langle A^sB^sA^sx,x\rangle \ge K(h_2,s)\langle (ABA)^sA^{s-1}x, A^{s-1}x\rangle \|A^sx\|^{2-2s}\|A^{s-1}x\|^{2s-2}.$$

Since $\langle A^{2s}x,x\rangle \ge \|A^{-2s}\|^{-1}$, it follows that

$$\begin{aligned} \|A^sx\|^{2s-2}\|A^{s-1}x\|^{2-2s} &= \langle A^{2s}x,x\rangle^{s-1}\langle A^{2s-2}x,x\rangle^{1-s} \\ &\le \|A^{-2}\|^{s(1-s)}\|A^{-2}\|^{(1-s)^2} = \|A^{-2}\|^{1-s} \end{aligned}$$

and hence

$$K(h_2,s)A^{s-1}(ABA)^sA^{s-1} \le \|A^{-2}\|^{1-s}A^sB^sA^s,$$

as desired. □

Proof of Theorem 5.31. Replacing A by $A^{-\frac{1}{2}}$ in Lemma 5.7, we have

$$A^{-\frac{1}{2}}B^sA^{-\frac{1}{2}} \le \|A^{-1}\|^{1-s}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^s$$

for all $0 < s \le 1$. By $\frac{\alpha}{(1-\alpha)s+\alpha} \in [0,1]$ and the Löwner-Heinz inequality, we have

$$\left(A^{-\frac{1}{2}}B^sA^{-\frac{1}{2}}\right)^{\frac{\alpha}{(1-\alpha)s+\alpha}} \le \left(\|A^{-1}\|^{1-s}\right)^{\frac{\alpha}{(1-\alpha)s+\alpha}} (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{\alpha s}{(1-\alpha)s+\alpha}}.$$

Therefore, it follows that

$$\|A \,\#_{\frac{\alpha}{(1-\alpha)s+\alpha}}\, B^s\| \le \|A^{-1}\|^{\frac{\alpha(1-s)}{(1-\alpha)s+\alpha}} \|A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{\alpha s}{(1-\alpha)s+\alpha}}A^{\frac{1}{2}}\|.$$

Next, we estimate the norm of the right-hand side in the expression above. Since $0 < \frac{\alpha s}{(1-\alpha)s+\alpha} \le 1$, it follows from (5.17) in Theorem 5.11 that

$$\begin{aligned} &\|A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{\alpha s}{(1-\alpha)s+\alpha}}A^{\frac{1}{2}}\| \\ &\le \|A^{\frac{(1-\alpha)s+\alpha}{2s}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{(1-\alpha)s+\alpha}{2s}}\|^{\frac{s}{(1-\alpha)s+\alpha}} \\ &= \|A^{\frac{\alpha(1-s)}{2s}}A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}A^{\frac{\alpha(1-s)}{2s}}\|^{\frac{s}{(1-\alpha)s+\alpha}} \\ &\le \|A \,\#_\alpha\, B\|^{\frac{s}{(1-\alpha)s+\alpha}} \|A\|^{\frac{\alpha(1-s)}{(1-\alpha)s+\alpha}} \end{aligned}$$

and since $\|A\|\|A^{-1}\| \le \frac{M_1}{m_1} = h_1$ by $m_1 \le A \le M_1$, we have

$$\|A \,\#_{\frac{\alpha}{(1-\alpha)s+\alpha}}\, B^s\| \le h_1^{\frac{\alpha(1-s)}{(1-\alpha)s+\alpha}} \|A \,\#_\alpha\, B\|^{\frac{s}{(1-\alpha)s+\alpha}} \tag{5.43}$$

for all $0 < s \le 1$. Since

$$A \#_\alpha B = B \#_{1-\alpha} A \quad \text{and} \quad A^r \#_{\frac{\alpha r}{1-\alpha+\alpha r}} B = B \#_{\frac{1-\alpha}{1-\alpha+\alpha r}} A^r,$$

replacing A, B, α and s by B, A, $1-\alpha$ and r in (5.43), it follows that for each $0 < r \le 1$

$$\|A^r \#_{\frac{\alpha r}{1-\alpha+\alpha r}} B\| \le h_2^{\frac{(1-\alpha)(1-r)}{1-\alpha+\alpha r}} \|A \#_\alpha B\|^{\frac{r}{1-\alpha+\alpha r}}. \tag{5.44}$$

For each $0 < r \le 1$ and $s \ge 1$, replacing A, B and r by A^s, B^s and $\frac{r}{s}$ respectively in (5.44), we have

$$\begin{aligned}
\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| &= \|(A^s)^{\frac{r}{s}} \#_{\frac{\alpha \frac{r}{s}}{(1-\alpha)+\alpha\frac{r}{s}}} B^s\| \\
&\le h_2^{s\frac{(1-\alpha)(1-\frac{r}{s})}{(1-\alpha)+\alpha\frac{r}{s}}} \|A^s \#_\alpha B^s\|^{\frac{\frac{r}{s}}{(1-\alpha)+\alpha\frac{r}{s}}} \qquad \text{by } 0 \le \tfrac{r}{s} \le 1 \text{ and (5.44)} \\
&= h_2^{\frac{(1-\alpha)s(s-r)}{(1-\alpha)s+\alpha r}} \|A^s \#_\alpha B^s\|^{\frac{r}{(1-\alpha)s+\alpha r}} \\
&\le h_2^{\frac{(1-\alpha)s(s-r)}{(1-\alpha)s+\alpha r}} \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \qquad \text{by } s \ge 1 \text{ and (AH)},
\end{aligned}$$

and hence we have the right-hand side in Theorem 5.31.

Next, we show the left-hand side in Theorem 5.31. By using the left-hand side of (5.42) in Lemma 5.7 and (5.17) in Theorem 5.11 it similarly follows that

$$K(h_2,s)^{\frac{\alpha}{(1-\alpha)s+\alpha}} K\left(h_1^\alpha h_2^\alpha, \tfrac{s}{(1-\alpha)s+\alpha}\right) h_1^{\frac{\alpha(s-1)}{(1-\alpha)s+\alpha}} \|A \#_\alpha B\|^{\frac{s}{(1-\alpha)s+\alpha}} \le \|A \#_{\frac{\alpha}{(1-\alpha)s+\alpha}} B^s\|$$

holds for each $0 < s \le 1$ and this implies that

$$\|A^r \#_{\frac{\alpha r}{1-\alpha+\alpha r}} B\| \ge K(h_1,r)^{\frac{1-\alpha}{\alpha r+1-\alpha}} K\left(h_1^\alpha h_2^\alpha, \tfrac{r}{\alpha r+1-\alpha}\right) h_2^{\frac{(1-\alpha)(r-1)}{\alpha r+1-\alpha}} \|A \#_\alpha B\|^{\frac{r}{\alpha r+1-\alpha}} \tag{5.45}$$

holds for each $0 < r \le 1$. Therefore for $s \ge 1$ we have

$$\begin{aligned}
\|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\| &= \|(A^s)^{\frac{r}{s}} \#_{\frac{\alpha\frac{r}{s}}{1-\alpha+\alpha\frac{r}{s}}} B^s\| \\
&\ge K(h_1^s, \tfrac{r}{s})^{\frac{(1-\alpha)s}{(1-\alpha)s+\alpha r}} K\left(h_1^{s\alpha} h_2^{s\alpha}, \tfrac{r}{(1-\alpha)s+\alpha r}\right) h_2^{\frac{(1-\alpha)s(r-s)}{(1-\alpha)s+\alpha r}} \|A^s \#_\alpha B^s\|^{\frac{r}{(1-\alpha)s+\alpha r}} \\
&\ge K(h_1^s, \tfrac{r}{s})^{\frac{(1-\alpha)s}{(1-\alpha)s+\alpha r}} K\left(h_1^{s\alpha} h_2^{s\alpha}, \tfrac{r}{(1-\alpha)s+\alpha r}\right) h_2^{\frac{(1-\alpha)s(r-s)}{(1-\alpha)s+\alpha r}} \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}},
\end{aligned}$$

as desired. Hence the proof is completed. □

We remark that in the case of $r = s = 1$, both bounds of the inequalities in Theorem 5.31 are equal to 1.

Finally, we show a complementary result of Theorem 5.31:

Corollary 5.6 *Suppose that the conditions of Theorem 5.31 hold. Then for each* $r \geq 1$ *and* $0 < s \leq 1$

$$h_2^{\frac{(1-\alpha)s(s-r)}{(1-\alpha)s+\alpha r}} \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}} \leq \|A^r \#_{\frac{\alpha r}{(1-\alpha)s+\alpha r}} B^s\|$$
$$\leq K\left(h_1^{\frac{r}{s}}, \frac{s}{r}\right)^{\frac{-(1-\alpha)rs}{(1-\alpha)s+\alpha r}} K\left((h_1^{\frac{r}{s}} h_2)^\alpha, \frac{r}{(1-\alpha)s+\alpha r}\right)^s h_2^{\frac{(1-\alpha)s(r-s)}{(1-\alpha)s+\alpha r}} \|A \#_\alpha B\|^{\frac{rs}{(1-\alpha)s+\alpha r}}.$$

5.6 Converses of Hölder's inequality

Let a_i and b_i be positive real numbers for $i = 1, \cdots, n$. Hölder's inequality says that

$$\sum_{i=1}^{n} a_i^{\frac{1}{p}} b_i^{\frac{1}{q}} \leq \left(\sum_{i=1}^{n} a_i\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i\right)^{\frac{1}{q}} \tag{5.46}$$

for $p, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. When $p = q = 2$ in (5.46), we have the Cauchy-Schwarz inequality

$$\sum_{i=1}^{n} \sqrt{a_i b_i} \leq \sqrt{\sum_{i=1}^{n} a_i} \sqrt{\sum_{i=1}^{n} b_i}. \tag{5.47}$$

These inequalities can be extended to operators by means of the subadditivity of the operator geometric mean [124, Theorem 5.7]: Let $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$ be positive invertible operators on a Hilbert space. The following inequality is regarded as Hölder's inequality for operators

$$\sum_{i=1}^{n} A_i \#_\alpha B_i \leq \left(\sum_{i=1}^{n} A_i\right) \#_\alpha \left(\sum_{i=1}^{n} B_i\right) \tag{5.48}$$

for all $\alpha \in [0,1]$. In particular, in the case of $\alpha = \frac{1}{2}$ we have the Cauchy-Schwarz operator inequality:

$$\sum_{i=1}^{n} A_i \# B_i \leq \left(\sum_{i=1}^{n} A_i\right) \# \left(\sum_{i=1}^{n} B_i\right). \tag{5.49}$$

By using the Mond-Pečarić method for concave functions, we shall show converses of operator Hölder's inequality (5.48).

Theorem 5.32 *Let* A_i *and* B_i *be positive invertible operators such that* $mA_i \leq B_i \leq MA_i$ *for some scalars* $0 < m \leq M$ *and* $i = 1, 2, \cdots, n$. *Then for each* $\alpha \in [0,1]$

$$\left(\sum_{i=1}^{n} A_i\right) \#_\alpha \left(\sum_{i=1}^{n} B_i\right) \leq K(m, M, \alpha)^{-1} \sum_{i=1}^{n} A_i \#_\alpha B_i \tag{5.50}$$

and

$$\left(\sum_{i=1}^{n} A_i\right) \#_\alpha \left(\sum_{i=1}^{n} B_i\right) - \sum_{i=1}^{n} A_i \#_\alpha B_i \le -C(m,M,\alpha) \sum_{i=1}^{n} A_i, \tag{5.51}$$

where the generalized Kantorovich constant $K(m,M,\alpha)$ is defined by (2.29) *and the Kantorovich constant for the difference $C(m,M,\alpha)$ is defined by* (2.38)*.*

If we put $\alpha = \frac{1}{2}$ in Theorem 5.32, then we have the following converses of the Cauchy-Schwarz operator inequalities.

Corollary 5.7 *Let A_i and B_i be positive invertible operators such that $mA_i \le B_i \le MA_i$ for some scalars $0 < m \le M$ and $i = 1,2,\cdots,n$. Then*

$$\left(\sum_{i=1}^{n} A_i\right) \# \left(\sum_{i=1}^{n} B_i\right) \le \frac{\sqrt{M}+\sqrt{m}}{2\sqrt[4]{mM}} \sum_{i=1}^{n} A_i \# B_i$$

and

$$\left(\sum_{i=1}^{n} A_i\right) \# \left(\sum_{i=1}^{n} B_i\right) - \sum_{i=1}^{n} A_i \# B_i \le \frac{(\sqrt{M}-\sqrt{m})^2}{4(\sqrt{M}+\sqrt{m})} \sum_{i=1}^{n} A_i.$$

To prove our results, we need the following Lemmas, also see [124, Corollary 5.33].

Lemma 5.8 *Let A and B be positive operators such that $mA \le B \le MA$ for some scalars $0 < m \le M$ and let $\Phi : B(H) \mapsto B(K)$ be a positive linear mapping. Then for each $\alpha \in [0,1]$*

$$\Phi(A) \#_\alpha \Phi(B) \le K(m,M,\alpha)^{-1} \Phi(A \#_\alpha B), \tag{5.52}$$

where the generalized Kantorovich constant $K(m,M,\alpha)$ is optimal.

Proof. Put $C = A^{-1/2}BA^{-1/2}$. If we put

$$\lambda_0 = \frac{\alpha}{1-\alpha}\frac{M^{1-\alpha}-m^{1-\alpha}}{m^{-\alpha}-M^{-\alpha}} \qquad \text{and} \qquad \mu_0 = \frac{\alpha(M-m)}{M^\alpha - m^\alpha},$$

then

$$\alpha t^{1-\alpha} + (1-\alpha)\lambda_0 t^{-\alpha} \le \mu_0 \qquad \text{for all } t \in [m,M].$$

Since $mI \le C \le MI$, we get

$$\alpha C + (1-\lambda_0)I \le \mu_0 C^\alpha$$

and hence

$$\alpha B + (1-\alpha)\lambda_0 A \le \mu_0 A \#_\alpha B.$$

This implies

$$(1-\alpha)\lambda_0 \Phi(A) + \alpha\Phi(B) \le \mu_0 \Phi(A \#_\alpha B).$$

By the weighted arithmetic-geometric mean inequality, it follows that

$$(1-\alpha)\lambda_0 \Phi(A) + \alpha\Phi(B) \ge \lambda_0^{1-\alpha}\, \Phi(A) \#_\alpha \Phi(B).$$

On the other hand, since $\frac{\mu_0}{\lambda_0^{1-\alpha}} = K(m,M,\alpha)^{-1}$ by an easy calculation, we have the desired inequality (5.52).

Moreover, the generalized Kantorovich constant $K(m,M\alpha)$ is optimal in the sense that for each $\alpha \in [0,1]$ there exist two positive operators A, B such that $mA \le B \le MA$ for some scalars $0 < m \le M$ and a positive linear mapping Φ such that

$$\Phi(A) \,\#_\alpha\, \Phi(B) = K(m,M,\alpha)^{-1}\Phi(A \,\#_\alpha\, B).$$

As a matter of fact, let $\Phi : M_2(\mathbb{C}) \mapsto \mathbb{C}$ be a positive linear mapping defined by

$$\Phi(X) = rx_{11} + (1-r)x_{22} \qquad \text{for} \quad X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \qquad \text{with } 0 < r < 1$$

and A and B positive definite matrices such as

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} M & 0 \\ 0 & m \end{pmatrix}.$$

Then it is clear that the sandwich condition $mA \le B \le MA$ holds. If we put

$$r = \frac{\alpha m^\alpha (M-m) - m(M^\alpha - m^\alpha)}{(1-\alpha)(M-m)(M^\alpha - m^\alpha)},$$

then we have $0 < r < 1$. Therefore it follows that

$$\begin{aligned} \frac{\Phi(A) \,\#_\alpha\, \Phi(B)}{\Phi(A \,\#_\alpha\, B)} &= \frac{(m + r(M-m))^\alpha}{m^\alpha + r(M^\alpha - m^\alpha)} \\ &= \left(\frac{\alpha(Mm^\alpha - mM^\alpha)}{(1-\alpha)(M^\alpha - m^\alpha)} \right)^\alpha \frac{(1-\alpha)(M-m)}{m^\alpha M - mM^\alpha} \\ &= K(m,M,\alpha)^{-1}. \end{aligned}$$

□

Lemma 5.9 *Let A and B be positive operators such that $mA \le B \le MA$ for some scalars $0 < m \le M$ and let $\Phi : B(H) \mapsto B(K)$ be a positive linear mapping. Then for each $\alpha \in [0,1]$*

$$\Phi(A) \,\#_\alpha\, \Phi(B) - \Phi(A \,\#_\alpha\, B) \le -C(m,M,\alpha)\Phi(A), \tag{5.53}$$

where the Kantorovich constant for the difference $C(m,M,\alpha)$ is optimal.

Proof. Put $C = A^{-1/2}BA^{-1/2}$ and for each $\alpha \in [0,1]$

$$\lambda = \left(\frac{M^\alpha - m^\alpha}{\alpha(M-m)} \right)^{\frac{1}{\alpha - 1}}.$$

Since $mI \le C \le MI$, it follows that

$$(1-\alpha)\lambda^\alpha I + \alpha\lambda^{\alpha-1}C \le C^\alpha - C(m,M,\alpha)I$$

and hence

$$(1-\alpha)\lambda^{\alpha}A+\alpha\lambda^{\alpha-1}B \le A \#_{\alpha} B - C(m,M,\alpha)A.$$

This implies

$$(1-\alpha)\lambda^{\alpha}\Phi(A)+\alpha\lambda^{\alpha-1}\Phi(B) \le \Phi(A \#_{\alpha} B) - C(m,M,\alpha)\Phi(A).$$

On the other hand, by the weighted arithmetic-geometric mean inequality

$$(1-\alpha)\lambda^{\alpha}\Phi(A)+\alpha\lambda^{\alpha-1}\Phi(B) \ge \Phi(A) \#_{\alpha} \Phi(B)$$

and hence we have the desired inequality (5.53).

Moreover, the Kantorovich constant for the difference $C(m,M,\alpha)$ is optimal in the sense that for each $\alpha \in [0,1]$ there exist two positive operators A, B such that $mA \le B \le MA$ for some scalars $0 < m \le M$ and a positive linear mapping Φ such that

$$\Phi(A) \#_{\alpha} \Phi(B) - \Phi(A \#_{\alpha} B) = -C(m,M,\alpha)\Phi(A).$$

As a matter of fact, put A,B and Φ be as in Lemma 5.8. If we put

$$r = \frac{1}{M-m}\left(\frac{M^{\alpha}-m^{\alpha}}{\alpha(M-m)}\right)^{\frac{1}{\alpha-1}} - \frac{m}{M-m},$$

then we have $0 < r < 1$ and

$$\begin{aligned}\Phi(A) \#_{\alpha} \Phi(B) - \Phi(A \#_{\alpha} B) &= (m+r(M-m))^{\alpha} - m^{\alpha} - r(M^{\alpha}-m^{\alpha}) \\ &= -C(m,M,\alpha) \\ &= -C(m,M,\alpha)\Phi(A)\end{aligned}$$

as desired. □

Proof of Theorem 5.32. If we put $\mathscr{A} = \mathrm{diag}(A_1,\cdots,A_n), \mathscr{B} = \mathrm{diag}(B_1,\cdots,B_n)$ and $\Phi(\mathscr{A}) = Z^*\mathscr{A}Z$ where $Z^* = (I,\cdots,I)$ in Lemma 5.8 and Lemma 5.9 respectively, then a sandwich condition $m\mathscr{A} \le \mathscr{B} \le M\mathscr{A}$ is satisfied and we have Theorem 5.32. □

5.7 Notes

Kantorovich type converse inequalities for operator norm and spectral radius are firstly discussed by Bourin [26] and afterward generalized by J.I. Fujii, Seo and Tominaga [76].

The results in Section 5.2 are due to [102], in Section 5.3 [235], in Section 5.4 [236, 263], in Section 5.5 [268] and in Section 5.6 [167, 28, 93, 266].

Matrix Young inequality is due to Ando [10].

Chapter 6

Geometric Mean

This chapter is devoted to the geometric mean of n operators due to Ando-Li-Mathias and Lowson-Lim. We present an alternative proof of the power convergence of the symmetrization procedure on the weighted geometric mean due to Lawson and Lim. We show a converse of the weighted arithmetic-geometric mean inequality of n operators.

6.1 Introduction

First of all, we present a definition for the geometric mean of three or more positive invertible operators on a Hilbert space. For positive invertible operators A and B on a Hilbert space H, the geometric mean $A \# B$ of A and B is defined by

$$A \# B = A^{\frac{1}{2}} \left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right)^{\frac{1}{2}} A^{\frac{1}{2}}.$$

As an extension of $A \# B$, for any positive integer $n \geq 2$, the geometric mean $G(A_1, A_2, \cdots, A_n)$ of any n-tuple of positive invertible operators $A_1, A_2, \cdots, A_n$ on a Hilbert space H is defined by induction as follows:

(i) $G(A_1, A_2) = A \# B$.

(ii) Assume that the geometric mean of any $(n-1)$-tuple of operators is defined. Let

$$G((A_j)_{j \neq i}) = G(A_1, \cdots, A_{i-1}, A_{i+1}, \cdots, A_n)$$

and let sequences $\{A_i^{(r)}\}_{r=1}^{\infty}$ be $A_i^{(1)} = A_i$ and $A_i^{(r+1)} = G((A_j^{(r)})_{j\neq i})$. Then there exists $\lim_{r\to\infty} A_i^{(r)}$ uniformly and it does not depend on i. Hence the geometric mean of n operators is defined by

$$\lim_{r\to\infty} A_i^{(r)} = G(A_1, A_2, \cdots, A_n) \qquad \text{for } i = 1, \cdots, n.$$

For positive invertible operators A and B, let

$$R(A,B) = \max\left\{r(A^{-1}B), r(B^{-1}A)\right\},$$

where $r(T)$ means the spectral radius of T. Then

$$R(G(A_1,A_2,\cdots,A_n), G(B_1,B_2,\cdots,B_n)) \le \left\{\prod_{i=1}^{n} R(A_i,B_i)\right\}^{\frac{1}{n}}. \tag{6.1}$$

In particular,

$$R(A_i^{(2)}, A_j^{(2)}) \le R(A_i,A_j)^{\frac{1}{n-1}}. \tag{6.2}$$

We have the following converse of the arithmetic-geometric mean inequality.

Lemma 6.1 *Let A_1 and A_2 be positive operators such that $mI_H \le A_1, A_2 \le MI_H$ for some scalars $M > m > 0$. Then*

$$\frac{A_1 + A_2}{2} \le \frac{M+m}{2\sqrt{Mm}} G(A_1,A_2) = \frac{M+m}{2\sqrt{Mm}} A_1 \# A_2.$$

Proof. If we put $C = A_1^{-\frac{1}{2}} A_2 A_1^{-\frac{1}{2}}$, then we have $\sqrt{\frac{m}{M}} I_H \le C^{\frac{1}{2}} \le \sqrt{\frac{M}{m}} I_H$. Since

$$\max\left\{t + \frac{1}{t} : \sqrt{\frac{m}{M}} \le t \le \sqrt{\frac{M}{m}}\right\} = \sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}},$$

it follows that

$$\frac{1}{2}\left(C^{\frac{1}{2}} + C^{-\frac{1}{2}}\right) \le \frac{1}{2}\left(\sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}}\right)$$

and hence we have

$$\frac{I_H + C}{2} \le \frac{M+m}{2\sqrt{Mm}} C^{\frac{1}{2}}.$$

Multiplying both sides by $A_1^{\frac{1}{2}}$, we have

$$\frac{A_1 + A_2}{2} \le \frac{M+m}{2\sqrt{Mm}} A_1 \# A_2.$$

□

For any positive integer $n \ge 2$, we show a converse of the arithmetic-geometric mean inequality of n operators, which is an extension of Lemma 6.1:

Theorem 6.1 *For any positive integer $n \geq 2$, let $A_1, \cdots, A_n$ be positive invertible operators such that $mI_H \leq A_i \leq MI_H$ for $i = 1, 2, \cdots, n$ for some scalars $0 < m < M$. Then*

$$\frac{A_1 + \cdots + A_n}{n} \leq \left(\frac{(M+m)^2}{4Mm} \right)^{\frac{n-1}{2}} G(A_1, \cdots, A_n).$$

Proof. We will prove it by induction on n. In the case of $n = 2$, it holds by Lemma 6.1. Assume that Theorem 6.1 holds for $n-1$. For positive integer r, we define $A_i^{(r)}, h_r$ and K_r as follows:

$$A_i^{(0)} = A_i \quad \text{and} \quad A_i^{(r)} = G\left(\left(A_j^{(r-1)} \right)_{j \neq i} \right),$$

$$h_0 = h \quad \text{and} \quad h_r = \max_{i,j} R\left(A_i^{(r)}, A_j^{(r)} \right),$$

$$K_r = \frac{1 + h_r}{2\sqrt{h_r}}.$$

Then by the induction hypothesis on n, we have

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} A_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}\sum_{j \neq i} A_j \right) &\leq \frac{1}{n}\sum_{i=1}^{n} K_0^{n-2} A_i^{(1)} \\
&= K_0^{n-2}\frac{1}{n}\sum_{i=1}^{n} A_i^{(1)} \leq (K_0K_1)^{n-2}\frac{1}{n}\sum_{i=1}^{n} A_i^{(2)} \\
&\leq \cdots \\
&\leq (K_0K_1\cdots K_r)^{n-2}\frac{1}{n}\sum_{i=1}^{n} A_i^{(r+1)}.
\end{aligned}$$

Since

$$\lim_{r \to \infty} A_i^{(r)} = G(A_1, A_2, \cdots, A_n) \quad \text{for } i = 1, 2, \cdots, n,$$

we have

$$\lim_{r \to \infty} \frac{1}{n}\sum_{i=1}^{n} A_i^{(r+1)} = G(A_1, A_2, \cdots, A_n).$$

So we have only to prove the following inequality:

$$\limsup_{r \to \infty} K_0K_1\cdots K_r \leq K_0^{\frac{n-1}{n-2}}.$$

By (6.2), we have

$$1 \leq h_r \leq h_{r-1}^{\frac{1}{n-1}} \leq \cdots \leq h_0^{\left(\frac{1}{n-1}\right)^r}.$$

Since

$$\frac{1}{2}\left(\frac{1}{x} + x \right) \leq \frac{1}{2}\left(\frac{1}{y^\alpha} + y^\alpha \right) \leq \left\{ \frac{1}{2}\left(\frac{1}{y} + y \right) \right\}^\alpha$$

holds for $1 \leq x \leq y^{\alpha}$ and $\alpha \in (0,1]$, we have

$$K_r = \frac{1+h_r}{2\sqrt{h_r}} = \frac{1}{2}\left(\frac{1}{\sqrt{h_r}} + \sqrt{h_r}\right) \leq \left\{\frac{1}{2}\left(\frac{1}{\sqrt{h_0}} + \sqrt{h_0}\right)\right\}^{\left(\frac{1}{n-1}\right)^r} = K_0^{\left(\frac{1}{n-1}\right)^r}.$$

Therefore we obtain

$$K_0K_1\cdots K_r \leq K_0^{1+\frac{1}{n-1}+\cdots+\left(\frac{1}{n-1}\right)^r} \to K_0^{\frac{n-1}{n-2}} \quad \text{as } r \to \infty.$$

Hence we have

$$\frac{A_1+A_2+\cdots+A_n}{n} \leq \left(\frac{1+h}{2\sqrt{h}}\right)^{n-1} G(A_1,A_2,\cdots,A_n).$$

By putting $h = \frac{M}{m}$, we obtain this Theorem 6.1. □

The following result is a noncommutative variant of the Greub-Rheinboldt inequality, which is equivalent to the Kantorovich inequality.

Lemma 6.2 *Let A and B be positive operators such that $mI_H \leq A, B \leq MI_H$ for some scalars $0 < m < M$. Then*

$$\sqrt{\langle Ax,x\rangle\langle Bx,x\rangle} \leq \frac{M+m}{2\sqrt{Mm}}\langle A \# Bx,x\rangle$$

for every unit vector $x \in H$.

Proof. By the Kantorovich inequality, if C is a positive operator such that $aI_H \leq C \leq bI_H$ for some scalars $0 < a < b$, then

$$\langle C^2x,x\rangle \leq \frac{(a+b)^2}{4ab}\langle Cx,x\rangle^2 \tag{6.3}$$

for every unit vector $x \in H$. Replacing C by $(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}$ and x by $\frac{A^{\frac{1}{2}}}{\|A^{\frac{1}{2}}x\|}$ in (6.3), we have

$$\sqrt{\frac{m}{M}}I_H \leq (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}} \leq \sqrt{\frac{M}{m}}I_H$$

and hence

$$\langle Bx,x\rangle\langle Ax,x\rangle \leq \frac{1}{4}\left(\sqrt{\frac{M}{m}} + \sqrt{\frac{m}{M}}\right)\langle A \# Bx,x\rangle^2.$$

□

For any positive integer $n \geq 2$, we show a noncommutative variant of the Greub-Rheinboldt inequality of n operators, which is an extension of Lemma 6.2:

Theorem 6.2 *For any positive integer $n \geq 2$, let $A_1, \cdots, A_n$ be positive invertible operators on a Hilbert space H such that $0 < mI_H \leq A_i \leq MI_H$, $i = 1, 2, \cdots, n$, for some scalars $0 < m < M$. Then*

$$\sqrt[n]{\langle A_1x,x\rangle\langle A_2x,x\rangle\cdots\langle A_nx,x\rangle} \leq \left(\frac{(M+m)^2}{4Mm}\right)^{\frac{n-1}{2}} \langle G(A_1,\cdots,A_n)x,x\rangle$$

for all $x \in H$.

Proof. By using Theorem 6.1 and arithmetic-geometric mean inequality, we have

$$\begin{aligned}\left(\prod_{i=1}^{n}\langle A_ix,x\rangle\right)^{\frac{1}{n}} &\leq \frac{1}{n}\sum_{i=1}^{n}\langle A_ix,x\rangle \\ &= \left\langle \frac{1}{n}\sum_{i=1}^{n}A_ix,x\right\rangle \\ &\leq \left(\frac{(M+m)^2}{4Mm}\right)^{\frac{n-1}{2}} \langle G(A_1,A_2,\cdots,A_n)x,x\rangle.\end{aligned}$$

This completes the proof. □

We recall the Specht theorem: For $x_1, \cdots, x_n \in [m, M]$ with $M \geq m > 0$,

$$\frac{x_1+\cdots+x_n}{n} \leq S(h)\ \sqrt[n]{x_1\cdots x_n}, \tag{6.4}$$

where $h = \frac{M}{m} (\geq 1)$ and the Specht ratio $S(h)$ is defined by (2.35).

We recall the t-geometric mean for $t \in [0,1]$:

$$A \#_t B = A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}}$$

for positive invertible operators A and B. The following theorem is a noncommutative version of the Specht theorem (6.4) in the case of $n = 2$.

Theorem 6.3 *Let A_1 and A_2 be positive invertible operators such that $mI_H \leq A_1, A_2 \leq MI_H$ for some scalars $0 < m < M$ and put $h = \frac{M}{m}$. Then*

$$(1-t)A_1 + tA_2 \leq S(h)\ A_1 \#_t A_2 \qquad \text{for all } t \in [0,1].$$

To prove Theorem 6.3, we need the following converse ratio type inequality of Young's inequality.

Lemma 6.3 *Let a be a positive number. Then the inequality*

$$S(a)a^{1-t} \geq (1-t)a + t \tag{6.5}$$

holds for all $t \in [0,1]$. Consequently, for $a, b > 0$, the inequality

$$S\left(\frac{a}{b}\right)a^{1-t}b^t \geq (1-t)a + tb \tag{6.6}$$

holds for all $t \in [0,1]$.

Proof. Let $a \neq 1$. We put

$$f_a(t) = \frac{(1-t)a+t}{a^{1-t}} = \left(\frac{1-a}{a}t+1\right)a^t.$$

Then we obtain the constant $S(a) = \frac{(a-1)a^{\frac{1}{a-1}}}{e\log a}$ as the maximum of $f_a(t)$ for $t \in [0,1]$. Indeed, we have by the elementary differential calculus

$$f_a'(t) = \left\{\frac{1-a}{a} + \left(\frac{1-a}{a}t+1\right)\log a\right\}a^t,$$

and so the equation $f_a'(t) = 0$ has the following unique solution $t = t_0$:

$$t_0 = \frac{a}{a-1} - \frac{1}{\log a} \in [0,1].$$

In fact, the Klein inequality ensures $t_0 \in [0,1]$. Furthermore it is easily seen that

$$f_a'(t) > 0 \quad \text{for } t < t_0 \quad \text{and} \quad f_a'(t) < 0 \quad \text{for} \quad t > t_0.$$

Therefore, the maximum of $f_a(t)$ takes at $t = t_0$ and we have

$$\max_{0\le t\le 1} f_a(t) = f_a(t_0) = S(a).$$

In the case of $a = 1$, the inequality (6.5) is clear since $S(1) = 1$. Finally, the inequality (6.6) is obtained replacing a by $\frac{a}{b}$ in (6.5). □

Proof of Theorem 6.3. Let C be a positive operator such that $mI_H \le C \le MI_H$ for some scalars $0 < m < M$. Then we have

$$\max_{m\le a\le M} S(a)C^{1-t} \ge (1-t)C + tI_H$$

for all $t \in [0,1]$. Moreover, since $S(a)$ is decreasing for $0 < a < 1$ and increasing for $a > 1$, the maximum of $S(a)$ in $a \in [m,M]$ is given by $\max\{S(m),S(M)\}$ and hence

$$\max\{S(m),S(M)\}C^{1-t} \ge (1-t)C + tI_H.$$

Here we replace C by $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Then we have $\frac{m}{M}I_H \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \le \frac{M}{m}I_H$, i.e. $\frac{1}{h}I_H \le A^{-\frac{1}{2}}BA^{-\frac{1}{2}}b \le hI_H$. Hence it follows that for any $t \in [0,1]$

$$S(h)(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{1-t} \ge (1-t)A^{-\frac{1}{2}}BA^{-\frac{1}{2}} + tI_H$$

by $S(h) = S(\frac{1}{h})$. Multiplying both sides by $A^{\frac{1}{2}}$, we have

$$S(h)A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{1-t}A^{\frac{1}{2}} \ge (1-t)B + tA.$$

□

6.2 Weighted geometric mean

In this section, we present the construction of the weighted geometric mean of n operators, which extend to the geometric mean $G(A_1,\cdots,A_n)$ of n operators due to Ando-Li-Mathias. For two positive invertible operators A and B, the weighted (power) arithmetic, geometric and harmonic means for $t\in[0,1]$ are defined by

the weighted arithmetic mean $\quad A\ \nabla_t\ B := (1-t)A+tB,$

the weighted geometric mean $\quad A\ \#_t\ B := A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^t A^{\frac{1}{2}},$

the weighted harmonic mean $\quad A\ !_t\ B := \left((1-t)A^{-1}+tB^{-1}\right)^{-1}.$

We need some preparations to define weighted means of n operators. We use a limiting process to define a weighted means of n operators. In proving convergence we use the following Thompson metric on the convex cone Ω of positive invertible operators:

$$d(A,B)=\max\{\log M(A/B),\log M(B/A)\}$$

where $M(A/B)=\inf\{\lambda>0: A\le\lambda B\}$. We remark that Ω is a complete metric space with respect to this metric and the corresponding metric topology on Ω agrees with the relative norm topology.

Lemma 6.4 *Let A and B be positive invertible operators. Then the following estimates coincide with the Thompson metric:*

$$\begin{aligned} d(A,B) &= \max\{\log\|B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\|,\log\|A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|\} \\ &= \log(\max\{r(B^{-1}A),r(A^{-1}B)\}) \\ &= \|\log B^{-\frac{1}{2}}AB^{-\frac{1}{2}}\| = \|\log A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\|. \end{aligned}$$

The Thompson metric has many nice properties, cf. [13]:

Lemma 6.5 *Let A,B and C be positive invertible operators. Then*

(i) $d(A,C)\le d(A,B)+d(B,C)$.

(ii) $d(A,B)\ge 0$ *and* $d(A,B)=0 \iff A=B$.

(iii) $\exp(-d(A,B))A\le B\le\exp(d(A,B))A$.

(iv) $d(T^*AT,T^*BT)=d(A,B)$ *for every invertible operator T.*

(v) $d(A^t,B^t)\le t\,d(A,B)$ *for all $t\in[0,1]$.*

(vi) $\|A-B\|\le(\exp d(A,B)-1)\|A\|$.

Next, we estimate the distance among weighted geometric means.

Lemma 6.6 $d(A \#_t C, B \#_t D) \le (1-t)\, d(A,B) + t\, d(C,D)$ *for all* $t \in [0,1]$.

Proof. Note that

$$T^*(X)^t T = \|T^*T\| \left(\frac{T^*}{\|T^*\|} X^t \frac{T}{\|T\|} \right) \le \|T^*T\| \left(\frac{T^*}{\|T^*\|} X \frac{T}{\|T\|} \right)^t = \|T^*T\|^{1-t} (T^*XT)^t$$

by Jensen's operator inequality. So we have

$$\begin{aligned}
&\log r\left((A\#_t C)^{-1}(B\#_t D)\right) = \log r\left((A^{-1}\#_t C^{-1})(B\#_t D)\right) \\
&= \log r\left(A^{-1/2}(A^{1/2}C^{-1}A^{1/2})^t A^{-1/2} B^{1/2}(B^{-1/2}DB^{-1/2})^t B^{1/2}\right) \\
&= \log \|(A^{1/2}C^{-1}A^{1/2})^{t/2} A^{-1/2} B^{1/2}(B^{-1/2}DB^{-1/2})^t B^{1/2} A^{-1/2}(A^{1/2}C^{-1}A^{1/2})^{t/2}\| \\
&\le \log \|A^{-1/2}BA^{-1/2}\|^{1-t} \|(A^{1/2}C^{-1}A^{1/2})^{t/2}(A^{-1/2}DA^{-1/2})^t (A^{1/2}C^{-1}A^{1/2})^{t/2}\| \\
&\le \log \|A^{-1/2}BA^{-1/2}\|^{1-t} \|(A^{1/2}C^{-1}A^{1/2})^{1/2}(A^{-1/2}DA^{-1/2})(A^{1/2}C^{-1}A^{1/2})^{1/2}\|^t \\
&= \log r(A^{-1}B)^{1-t} r\left((A^{1/2}C^{-1}A^{1/2})(A^{-1/2}DA^{-1/2})\right)^t \\
&= \log r(A^{-1}B)^{1-t} r\left(C^{-1}D\right)^t = (1-t)\log r(A^{-1}B) + t \log r\left(C^{-1}D\right) \\
&= (1-t)\, d(A,B) + t\, d(C,D).
\end{aligned}$$

□

We present the definition of the weighted geometric mean $G[n,t]$ with $t \in [0,1]$ for an n-tuple of positive invertible operators $A_1, A_2, \cdots, A_n$. Let $G[2,t](A_1,A_2) = A_1 \#_t A_2$. For $n \ge 3$, $G[n,t]$ is defined inductively as follows: Put $A_i^{(1)} = A_i$ for all $i = 1, \cdots, n$ and

$$A_i^{(r)} = G[n-1,t]((A_j^{(r-1)})_{j \ne i}) = G[n-1,t](A_1^{(r-1)}, \cdots, A_{i-1}^{(r-1)}, A_{i+1}^{(r-1)}, \cdots, A_n^{(r-1)})$$

inductively for r. If sequences $\{A_i^{(r)}\}$ have the same limit $\lim_{r\to\infty} A_i^{(r)}$ for all $i = 1, \cdots, n$ in the Thompson metric, then we define

$$G[n,t](A_1, \cdots, A_n) = \lim_{r\to\infty} A_i^{(r)}.$$

To show that sequences $\{A_i^{(r)}\}$ converge, we investigate the construction of the weighted arithmetic mean due to Lawson and Lim: Let $A[2,t](A_1,A_2) = (1-t)A_1 + tA_2$. For $n \ge 3$, $A[n,t]$ is defined inductively as follows: Put $\widetilde{A_i^{(1)}} = A_i$ for all $i = 1, \cdots, n$ and

$$\widetilde{A_i^{(r)}} = A[n-1,t]((\widetilde{A_j^{(r-1)}})_{j \ne i}) = A[n-1,t](\widetilde{A_1^{(r-1)}}, \cdots, \widetilde{A_{i-1}^{(r-1)}}, \widetilde{A_{i+1}^{(r-1)}}, \cdots, \widetilde{A_n^{(r-1)}})$$

inductively for r. Then we see that sequences $\{\widetilde{A_i^{(r)}}\}$ have the same limit $\lim_{r\to\infty} \widetilde{A_i^{(r)}}$ for all $i = 1, \cdots, n$ because it is just the problems on weights. If we put

$$A[n,t](A_1, \cdots, A_n) = \lim_{r\to\infty} \widetilde{A_i^{(r)}},$$

then it is expressed by

$$A[n,t](A_1,\cdots,A_n) = t[n]_1 A_1 + \cdots + t[n]_n A_n \tag{6.7}$$

where $t[n]_i \geq 0$ for $i = 1,\cdots,n$ and $\sum_{i=1}^n t[n]_i = 1$. Similarly, we can define the weighted harmonic mean $H[n,t](A_1,\cdots,A_n)$ as

$$H[n,t](A_1,\cdots,A_n) = \left(t[n]_1 A_1^{-1} + \cdots + t[n]_n A_n^{-1}\right)^{-1}.$$

We remark that the coefficient $\{t[n]_i\}$ depends on n only. For example, in the case of $n = 2,3$, it follows that

$$\begin{aligned}
&A[2,t](A_1,A_2) = (1-t)A_1 + tA_2,\\
&t[2]_1 = 1-t \quad \text{and} \quad t[2]_2 = t,\\
&A[3,t](A_1,A_2,A_3) = \frac{1-t}{2-t}A_1 + \frac{1-t+t^2}{2+t-t^2}A_2 + \frac{t}{1+t}A_3,\\
&t[3]_1 = \frac{1-t}{2-t}, \quad t[3]_2 = \frac{1-t+t^2}{(2-t)(1+t)} \quad \text{and} \quad t[3]_3 = \frac{t}{1+t}.
\end{aligned}$$

For the sake of convenience, we show the general term of the coefficient $\{t[n]_i\}$:

Lemma 6.7 *For any positive integer $n \geq 2$*

$$t[n]_{n-m} = \frac{m(m+1) + 2m(n-2m-2)t + (n^2-(4m+1)n+4m(m+1))t^2}{(n-1)(m+(n-2m)t)(m+1+(n-2(m+1))t)} \tag{6.8}$$

for $m = 0,1,\cdots,n-1$.

Proof. We prove this lemma by induction on both n and m. First of all, we show

$$t[n]_n = \frac{t}{1+(n-2)t} \tag{6.9}$$

for any integer $n \geq 2$. Suppose that the expression (6.9) holds for n and put $\omega = t[n]_n \in (0,1)$. Noticing that $t[n+1]_{n+1} = A[n+1,t](O,\cdots,O,I_H)$, we consider the case $A_1 = \cdots = A_n = O$ and $A_{n+1} = I_H$. In this case, for $i = 1,\cdots,n$, all $\widetilde{A_i^{(r)}}$ are equal and hence we can write $a_r I_H = \widetilde{A_i^{(r)}}$, and also put $b_r I_H = \widetilde{A_{n+1}^{(r)}}$. Then simple observation shows

$$a_1 = 0,\ b_1 = 1,\ a_{r+1} = (1-\omega)a_r + \omega b_r \quad \text{and} \quad b_{r+1} = a_r,$$

and hence we have

$$a_1 = 0,\ a_{r+1} - a_r = -\omega(a_r - a_{r-1}).$$

It follows that

$$a_{r+1} = a_1 - \sum_{k=1}^{r} (-\omega)^k = \omega\frac{1-(-\omega)^r}{1+\omega} \longrightarrow \frac{t[n]_n}{1+t[n]_n} = \frac{t}{1+(n-1)t} \quad \text{as } r \to \infty.$$

It follows from $t[n+1]_{n+1} = \lim_{r\to\infty} a_r$ that (6.9) holds for any integer $n \geq 2$ by induction.

Replacing t by $1-t$ in (6.9), we have

$$t[n]_1 = \frac{1-t}{(n-1)-(n-2)t}. \tag{6.10}$$

By a similar consideration, we have the following recurrence formula:

$$t[n+1]_k = 1 - \frac{\sum\limits_{j<k} t[n]_j}{1+t[n]_{k-1}} - \frac{\sum\limits_{j>k-1} t[n]_j}{1+t[n]_k} \tag{6.11}$$

for $k = 2, \cdots, n$.

Now, we show (6.8) by induction. Since $t[2]_2 = t$, it follows that (6.8) holds for $n = 2$. Inductively, let $n \geq 2$ be an integer such that (6.8) holds. Then it follows from (6.9) and induction that

$$\sum_{j>n-m-1} t[n]_j = \frac{(m+1)(m+(n-2m-1)t)}{(n-1)(m+1+(n-2m-2)t)}$$

for all m such that $m < n$. If $m = n$, then it follows from (6.10) that $t[n+1]_1 = \frac{1-t}{n+(1-n)t}$. If $m < n$, then it follows from (6.9) and (6.11) that

$$t[n+1]_{n-m} = \frac{m(m+1)+2m(n-2m-1)t+(n^2-(4m-1)n+4m^2)t^2}{n(m+(n+1-2m)t)(m+1+(n-2m-1)t)}.$$

In the case of $n+1$, it follows that (6.8) holds for all $m < n+1$. Therefore, (6.8) holds for all $n \geq 2$ by induction. □

To confirm that the above weighted geometric mean can be always defined, we observe properties of the weighted geometric mean.

Lemma 6.8 *For any positive integer $n \geq 2$, let $A_1, \cdots, A_n$ and $B_1, \cdots, B_n$ be positive invertible operators. Assume that $G[n,t]$ is defined for $n \leq n_0$ for some n_0. Then*

$$d(G[n,t](A_1,\cdots,A_n), G[n,t](B_1,\cdots,B_n)) \leq A[n,t](d(A_1,B_1),\cdots,d(A_n,B_n)) \tag{6.12}$$

holds for $n \leq n_0$.

Proof. It follows from Lemma 6.6 that the inequality (6.12) holds for $n = 2$:

$$d(A_1 \#_t A_2, B_1 \#_t B_2) \leq (1-t)\, d(A_1,B_1) + t\, d(A_2,B_2) \quad \text{for all } t \in [0,1].$$

Assume (6.12) holds for $n = N < n_0$. For $(N+1)$-tuples $(A_1,\cdots,A_{N+1})$ and $(B_1,\cdots,B_{N+1})$, it follows by induction that

$$\begin{aligned} d(A_J^{(r+1)}, B_J^{(r+1)}) &= d\ (G[N,t]((A_j^{(r)})_{j\neq J}), G[N,t]((B_j^{(r)})_{j\neq J})\ ) \\ &\leq A[N,t]\ ((d(A_j^{(r)}, B_j^{(r)})_{j\neq J})\ ). \end{aligned}$$

Note that this process is parallel to that of the definition for the weighted arithmetic mean: For a fixed J, put weights $w_j^{(r)}$ inductively with

$$\begin{aligned}\widetilde{A_J^{(r+1)}} &= A[N,t]((\widetilde{A_j^{(r)}})_{j\neq J}) = \sum_{j=1}^{N+1} w_j^{(1)}\widetilde{A_j^{(r)}} \\ &= \sum_{j=1}^{N+1} w_j^{(2)}\widetilde{A_j^{(r-1)}} = \cdots = \sum_{j=1}^{N+1} w_j^{(r)} A_j.\end{aligned}$$

Then we have

$$\begin{aligned}\widetilde{A_J^{(r+1)}}\nabla_t\widetilde{B_J^{(r+1)}} &= \left(\sum_{j=1}^{N+1} w_j^{(r)} A_j\right)\nabla_t\left(\sum_{j=1}^{N+1} w_j^{(r)} B_j\right) \\ &= \sum_{j=1}^{N+1} w_j^{(r)}(A_j\nabla_t B_j).\end{aligned}$$

The left hand in the above equation converges to

$$A[N+1,t](A_1\nabla_t B_1,\cdots,A_{N+1}\nabla_t B_{N+1}) = \sum_{k=1}^{N+1} t[N+1]_k A_k\nabla_t B_k \quad \text{as } r\to\infty,$$

which implies

$$w_k^{(r)} \longrightarrow t[N+1]_k.$$

Then the same weights appear in the successive relations for $d(A_J^{(r+1)},B_J^{(r+1)})$ as that for $\widetilde{A_J^{(r+1)}}\nabla_t\widetilde{B_J^{(r+1)}}$:

$$\begin{aligned}d(A_J^{(r+1)},B_J^{(r+1)}) &\le A[N,t]((d(A_k^{(r)},B_k^{(r)}))_{k\neq J}) = \sum_{k=1}^{N+1} w_k^{(1)} d(A_k^{(r)},B_k^{(r)}) \\ &\le \sum_{k=1}^{N+1} w_k^{(2)} d(A_k^{(r-1)},B_k^{(r-1)}) \le \cdots \le \sum_{k=1}^{N+1} w_k^{(r)} d(A_k,B_k),\end{aligned}$$

so that, taking limit as $r\to\infty$, we have

$$\begin{aligned}&d(G[N+1,t](A_1,\cdots,A_{N+1}),G[N+1,t](B_1,\cdots,B_{N+1})) \\ &\qquad\le \sum_{k=1}^{N+1} t[N+1]_k d(A_k,B_k) = A[N+1,t](d(A_1,B_1),\cdots,d(A_{N+1},B_{N+1})).\end{aligned}$$

Thus (6.12) holds for all $n\le n_0$. □

Remark 6.1 *If $t=\frac{1}{2}$, then we have $t[n]_i=\frac{1}{n}$ for $i=1,\cdots,n$ and hence Lemma 6.8 is a generalization of* (6.1)*. In fact,*

$$R(G(A_1,\cdots,A_n),G(B_1,\cdots,B_n)) \le \left\{\prod_{i=1}^{n} R(A_i,B_i)\right\}^{\frac{1}{n}}.$$

Taking the logarithm of both sides of this inequality, we have

$$\log R(G(A_1,\cdots,A_n),G(B_1,\cdots,B_n)) \le \frac{1}{n}\left(\log R(A_1,B_1)+\cdots+\log R(A_n,B_n)\right),$$

that is,

$$d(G(A_1,\cdots,A_n),G(B_1,\cdots,B_n)) \le \frac{1}{n}\sum_{i=1}^{n} d(A_i,B_i).$$

Now we confirm that $G[n,t](A_1,\cdots,A_n)$ is defined for all n:

Theorem 6.4 *For any positive integer $n \ge 2$ and $0<t<1$, the weighted geometric mean $G[n,t]$ can be defined for all n-tuples of positive invertible operators and*

$$d\ (G[n,t](A_1,\cdots,A_n),G[n,t](B_1,\cdots,B_n)\ ) \le A[n,t](d(A_1,B_1),\cdots,d(A_n,B_n))$$

holds.

Proof. For $n=2$, $G[2,t](A_1,A_2)=A_1\#_tA_2$ is defined. Assume that $G[n,t]$ is defined for $n\le N$. Take $(N+1)$-tuples $(A_1,\cdots,A_{N+1})$ and $(B_1,\cdots,B_{N+1})$ of positive invertible operators. By Lemma 6.8, we have

$$\begin{aligned} d\ (G[n,t](A_{i(1)},\cdots,A_{i(n)}),G[n,t](B_{i(1)},\cdots,B_{i(n)})\ ) \\ \le A[n,t](d(A_{i(1)},B_{i(1)}),\cdots,d(A_{i(n)},B_{i(n)})) \end{aligned}$$

for all $n\le N$. Take the sequence $\{A_i^{(r)}\}$ for $(A_1,\cdots,A_{N+1})$ to define $G[N+1,t]$. To show the existence of the weighted geometric mean, we have only to show that $\{A_J^{(r)}\}_{r=1}^{\infty}$ for a fixed J is a Cauchy sequence in the Thompson metric. Then the above inequality shows

$$\begin{aligned} d(A_J^{(r+1)},A_J^{(r)}) &= d\ (G[N,t]((A_j^{(r)})_{j\ne J}),G[N,t]((A_J^{(r)}))\ ) \\ &\le A[N,t]((d(A_j^{(r)},A_J^{(r)}))_{j\ne J}) \\ &= A[N,t]((d(G[N,t]((A_i^{(r-1)})_{i\ne j}),G[N,t]((A_i^{(r-1)})_{i\ne J}))) \\ &= A[N,t]((d(G[N,t]((A_{j(i)}^{(r-1)})),G[N,t]((A_{J(i)}^{(r-1)}))) \\ &\le A[N,t]\big(A[N,t](d(A_{j(i)}^{(r-1)},A_{J(i)}^{(r-1)})\big) \end{aligned}$$

Since d is a metric, then $d(A_{j(i)}^{(r-1)},A_{J(i)}^{(r-1)})=0$ when $j(i)=J(i)$. Moreover a direct computation shows the above last form can be expressed by only the terms $d(A_k^{(r-1)},A_{k+1}^{(r-1)})$ $(k=1,\cdots,N)$. There exist positive numbers v_k which do not depend on r with

$$d(A_J^{(r+1)},A_J^{(r)}) \le \sum_{k=1}^{N} v_k d(A_k^{(r-1)},A_{k+1}^{(r-1)}).$$

Since Lemma 6.8 implies

$$\begin{aligned} d(A_k^{(r)}, A_{k+1}^{(r)}) &= d\Big(G[N,t](A_1^{(r-1)}, \cdots, A_{k-1}^{(r-1)}, A_{k+1}^{(r-1)}, A_{k+2}^{(r-1)}, \cdots, A_{N+1}^{(r-1)}), \\ &\qquad G[N,t](A_1^{(r-1)}, \cdots, A_{k-1}^{(r-1)}, A_k^{(r-1)}, A_{k+2}^{(r-1)}, \cdots, A_{N+1}^{(r-1)})\Big) \\ &\leq t[N]_k d(A_k^{(r-1)}, A_{k+1}^{(r-1)}) \leq \cdots \leq t[N]_k^{r-1} d(A_k, A_{k+1}), \end{aligned}$$

we have

$$d(A_J^{(r+1)}, A_J^{(r)}) \leq \sum_{k=1}^{N} v_k d(A_k^{(r-1)}, A_{k+1}^{(r-1)}) \leq \sum_{k=1}^{N} v_k t[N]_k^{r-1} d(A_k, A_{k+1}).$$

Putting $\rho = \max\{1-t, t\}$ and $M = \max_k d(A_k, A_{k+1})$, we have $t[N]_k \leq \rho$ and

$$d(A_J^{(r+1)}, A_J^{(r)}) \leq \Big(\sum_{k=1}^{N} v_k\Big) M \rho^{r-1}.$$

Therefore, for $s > r$,

$$\begin{aligned} d(A_J^{(s)}, A_J^{(r)}) &\leq \sum_{j=1}^{s-r} d(A_J^{(s-j+1)}, A_J^{(s-j)}) \leq \Big(\sum_{k=1}^{N} v_k\Big) \sum_{j=1}^{s-r} M \rho^{s-j-1} \\ &= \Big(\sum_{k=1}^{N} v_k\Big) M \frac{\rho^{r-1}(1-\rho^{s-r})}{1-\rho} \leq \Big(\sum_{k=1}^{N} v_k\Big) M \frac{\rho^{r-1}}{1-\rho} \to 0 \quad \text{as } r \to \infty, \end{aligned}$$

which means the sequence $\{A_J^{(r)}\}$ for $J = 1, \cdots, N+1$ is Cauchy. Finally, we show $\{A_J^{(r)}\}$ for $J=1, \cdots, N+1$ have the same limit. It is enough to show that $\lim_{r\to\infty} A_1^{(r)} = \lim_{r\to\infty} A_2^{(r)}$. Let B_1 and B_2 such that $\lim_{r\to\infty} A_1^{(r)} = B_1$ and $\lim_{r\to\infty} A_2^{(r)} = B_2$. Then

$$d(B_1, B_2) \leq d(B_1, A_1^{(r)}) + d(A_1^{(r)}, A_2^{(r)}) + d(A_2^{(r)}, B_2)$$

and

$$0 \leq d(A_1^{(r)}, A_2^{(r)}) \leq t[N]_1^{r-1} d(A_1, A_2) \to 0 \quad \text{as } r \to \infty.$$

We conclude that $B_1 = B_2$. Hence $\{A_J^{(r)}\}$ for $J = 1, \cdots, N+1$ have the same limit by the same way. Thus $G[n,t]$ is defined and the required inequality holds by Lemma 6.8.

□

Remark 6.2 *The $\frac{1}{2}$-weighted geometric mean $G[n, \frac{1}{2}](A_1, \cdots, A_n)$ coincides with the geometric one $G(A_1, \cdots, A_n)$ due to Ando-Li-Mathias.*

We sum up some properties of the weighted geometric mean:

Theorem 6.5 *Let $0 < t < 1$ and any positive integer $n \geq 2$.*

(P1) *Consistency with scalars. If $A_1, \cdots, A_n$ mutually commute for $i = 1, \cdots, n$, then*

$$G[n,t](A_1, \cdots, A_n) = \prod_{i=1}^{n} A_i^{t[n]_i},$$

where $\{t[n]_i\}$ is defined by (6.7) and (6.8).

(P2) *Joint homogeneity. For $\alpha_i > 0$*

$$\begin{aligned} G[n,t](\alpha_1 A_1, \cdots, \alpha_n A_n) &= G[n,t](\alpha_1, \cdots, \alpha_n) G[n,t](A_1, \cdots, A_n) \\ &= \prod_{i=1}^{n} \alpha_i^{t[n]_i} G[n,t](A_1, \cdots, A_n), \end{aligned}$$

where $\{t[n]_i\}$ is defined by (6.7) and (6.8).

(P3) *Monotonicity. The mapping $(A_1, \cdots, A_n) \mapsto G[n,t](A_1, \cdots, A_n)$ is monotone, i.e. if $A_i \geq B_i$ for $i = 1, \cdots, n$, then*

$$G[n,t](A_1, \cdots, A_n) \geq G[n,t](B_1, \cdots, B_n).$$

(P4) *Congruence invariance. For every invertible operator T*

$$G[n,t](T^* A_1 T, \cdots, T^* A_n T) = T^* G[n,t](A_1, \cdots, A_n) T.$$

(P5) *Joint concavity. The mapping $(A_1, \cdots, A_n) \mapsto G[n,t](A_1, \cdots, A_n)$ is jointly concave:*

$$G[n,t]\Big(\sum_{i=1}^{n} \lambda_i A_{1i}, \cdots, \sum_{i=1}^{n} \lambda_i A_{ni}\Big) \geq \sum_{i=1}^{n} \lambda_i G[n,t](A_{1i}, \cdots, A_{ni}),$$

where $\lambda_i \geq 0$ with $\sum_{i=1}^{n} \lambda_i = 1$.

(P6) *Self-duality.*

$$G[n,t](A_1, \cdots, A_n) = G[n,t](A_1^{-1}, \cdots, A_n^{-1})^{-1}.$$

(P7) *The arithmetic-geometric-harmonic mean inequality holds:*

$$H[n,t](A_1, \cdots, A_n) \leq G[n,t](A_1, \cdots, A_n) \leq A[n,t](A_1, \cdots, A_n). \tag{AGH}$$

Proof. The properties (P1)-(P7) can be easily proved by induction and the fact that they are known to be true for $n = 2$. To illustrate that we prove (P7). We know that the result is true for $n = 2$. Now let us assume it is true for n and prove it for $n + 1$.

$$\begin{aligned} A_i^{(r+1)} &= G[n,t]((A_j^{(r)})_{j \neq i}) \leq A[n,t]((A_j^{(r)})_{j \neq i}) \\ &\leq A[n,t]((\widetilde{A_j^{(r)}})_{j \neq i}) = \widetilde{A_i^{(r+1)}} \end{aligned}$$

for $i = 1, \cdots, n+1$. Therefore, as $r \to \infty$, we have $G[n+1,t](A_1, \cdots, A_{n+1}) \leq A[n+1,t](A_1, \cdots, A_{n+1})$. By (P6), we have the left-hand side of (P7). □

6.3 The Kantorovich type inequality

First of all, we show a converse of the weighted arithmetic-geometric mean inequality of n operators, which is an improvement of Theorem 6.1 for $n \geq 3$:

Theorem 6.6 *For any positive integer $n \geq 2$, let $A_1, A_2, \cdots, A_n$ be positive invertible operators on a Hilbert space H such that $0 < mI_H \leq A_i \leq MI_H$ for $i = 1, 2, \cdots, n$ and some scalars $0 < m < M$. Then*

$$A[n,t](A_1, \cdots, A_n) \leq \frac{(M+m)^2}{4Mm} G[n,t](A_1, \cdots, A_n)$$

for $0 < t < 1$.

Remark 6.3 *In the case of $t = \frac{1}{2}$, we have*

$$\frac{A_1 + \cdots + A_n}{n} \leq \frac{(M+m)^2}{4Mm} G(A_1, \cdots, A_n). \tag{6.13}$$

For $n = 3$, the constant in (6.13) coincides with one in Theorem 6.1. For $n \geq 4$, the constant in (6.13) is less than one in Theorem 6.1.

To prove Theorem 6.6, we need the following lemma.

Lemma 6.9 *Let Φ be a positive linear mapping on the algebra $B(H)$ of all bounded linear operators on a Hilbert space H such that $\Phi(I_H) = I_H$. Then*

$$\Phi(A) \leq \frac{(M+m)^2}{4Mm} \Phi(A^{-1})^{-1}$$

for all positive operators A such that $mI_H \leq A \leq MI_H$ for some scalars $M > m > 0$.

Proof. Since $t^{-1} \leq \frac{M+m}{Mm} - \frac{1}{Mm}t$ for all $t \in [m, M]$, we have

$$\begin{aligned}
\Phi(A) &\leq (M+m)I - Mm\Phi(A) \\
&= \frac{(M+m)^2}{4Mm}\Phi(A^{-1})^{-1} - \left(\frac{M+m}{2\sqrt{Mm}}\Phi(A^{-1})^{-\frac{1}{2}} - \sqrt{Mm}\Phi(A)^{\frac{1}{2}}\right)^2 \\
&\leq \frac{(M+m)^2}{4Mm}\Phi(A^{-1})^{-1}.
\end{aligned}$$

□

Proof of Theorem 6.6. Let a mapping $\Psi : B(H) \oplus \cdots \oplus B(H) \mapsto B(H) \oplus \cdots \oplus B(H)$ be defined by

$$\Psi \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix} = \begin{pmatrix} t[n]_1 A_1 + \cdots + t[n]_n A_n & & 0 \\ & \ddots & \\ 0 & & t[n]_1 A_1 + \cdots + t[n]_n A_n \end{pmatrix}$$

where $\{t[n]_i\}$ is defined by (6.7). Then Ψ is a positive linear mapping such that $\Psi(I_H) = I_H$. Since

$$m \begin{pmatrix} I_H & & 0 \\ & \ddots & \\ 0 & & I_H \end{pmatrix} \le \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix} \le M \begin{pmatrix} I_H & & 0 \\ & \ddots & \\ 0 & & I_H \end{pmatrix},$$

it follows from Lemma 6.9 that

$$\Psi \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix} \le \frac{(M+m)^2}{4Mm} \Psi \begin{pmatrix} A_1^{-1} & & 0 \\ & \ddots & \\ 0 & & A_n^{-1} \end{pmatrix}^{-1}$$

and hence

$$A[n,t](A_1, \cdots, A_n) \le \frac{(M+m)^2}{4Mm} H[n,t](A_1, \cdots, A_n).$$

By (P7) in Theorem 6.5 we have the desired inequality

$$A[n,t](A_1, \cdots, A_n) \le \frac{(M+m)^2}{4Mm} G[n,t](A_1, \cdots, A_n).$$

□

By using Theorem 6.6 and the weighted arithmetic-geometric mean inequality, we obtain a weighted version of Greub-Rheinboldt inequality of n operators:

Theorem 6.7 *For any positive integer $n \ge 2$, let $A_1, A_2, \cdots, A_n$ be positive invertible operators on a Hilbert space H such that $mI_H \le A_i \le MI_H$ for $i = 1, 2, \cdots, n$ and some scalars $0 < m < M$. Then for $0 < t < 1$*

$$\langle A_1 x, x\rangle^{t[n]_1} \langle A_2 x, x\rangle^{t[n]_2} \cdots \langle A_n x, x\rangle^{t[n]_n} \le \frac{(M+m)^2}{4Mm} \langle G[n,t](A_1, A_2, \cdots, A_n)x, x\rangle$$

holds for all $x \in H$, where $\{t[n]_i\}$ is defined by (6.7) and (6.8).

Proof. For $0 < t < 1$, we have

$$\begin{aligned} \langle A_1 x, x\rangle^{t[n]_1} \langle A_2 x, x\rangle^{t[n]_2} \cdots \langle A_n x, x\rangle^{t[n]_n} &\le t[n]_1 \langle A_1 x, x\rangle + \cdots + t[n]_n \langle A_n x, x\rangle \\ &= \langle A[n,t](A_1, \cdots, A_n)x, x\rangle \\ &\le \frac{(M+m)^2}{4Mm} \langle G[n,t](A_1, A_2, \cdots, A_n)x, x\rangle \end{aligned}$$

for all $x \in H$. □

Remark 6.4 *If we put $t = \frac{1}{2}$ in Theorem 6.7, then $t[n]_i = \frac{1}{n}$ for all $i = 1, \cdots, n$. Therefore, we have an improvement of Theorem 6.2 for $n \ge 4$.*

6.4 The Specht type inequality

We recall a 2-operators version of the Specht theorem (Theorem 6.3): If A_1 and A_2 are positive invertible operators such that $mI_H \le A_1, A_2 \le MI_H$ for some scalars $0 < m < M$, then

$$(1-t)A_1 + tA_2 \le S(h) A_1 \#_t A_2 \qquad \text{for all } t \in [0,1],$$

where $h = \frac{M}{m}$. Actually, the Specht ratio is the upper bound of the arithmetic mean by the geometric one for positive numbers. We show a noncommutative version of the Specht theorem of n operators. For this, we state the following lemma.

Lemma 6.10 *Let $A_1, A_2, \cdots, A_n$ be positive invertible operators such that $mI_H \le A_i \le MI_H$ for some scalars $0 < m < M$ and $i = 1, 2, \cdots, n$ and $\alpha_1, \alpha_2, \cdots, \alpha_n$ positive numbers with $\sum_{i=1}^n \alpha_i = 1$. Put $h = \frac{M}{m}$. Then*

$$\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_n A_n \le S(h)\ \exp\left(\alpha_1 \log A_1 + \alpha_2 \log A_2 + \cdots + \alpha_n \log A_n\right),$$

where $S(h)$ is the Specht ratio defined by (2.35).

Proof. Put $\mathbb{A} = \mathrm{diag}(A_1, \cdots, A_n)$ and $y = (\sqrt{\alpha_1}x, \cdots, \sqrt{\alpha_n}x)^T$ for every unit vector x. By Theorem 2.14, we have

$$\langle \mathbb{A}y, y\rangle \le S(h)\ \exp\langle \log \mathbb{A}\ y, y\rangle$$

since $mI_H \le \mathbb{A} \le MI_H$. Therefore, it follows from Jensen's inequality that

$$\begin{aligned}\langle(\alpha_1 A_1 + \cdots + \alpha_n A_n)x, x\rangle = \langle \mathbb{A}y, y\rangle &\le S(h)\ \exp\langle \log \mathbb{A}\ y, y\rangle \\ &= S(h)\ \exp\left\langle \sum_{i=1}^n \alpha_i \log A_i x, x\right\rangle \\ &\le S(h)\ \langle \exp(\alpha_1 \log A_1 + \cdots + \alpha_n \log A_n)x, x\rangle\end{aligned}$$

for every unit vector $x \in H$ and hence we have

$$\alpha_1 A_1 + \cdots + \alpha_n A_n \le S(h)\ \exp\left(\alpha_1 \log A_1 + \cdots + \alpha_n \log A_n\right).$$

□

By virtue of Lemma 6.10, we have the following theorem.

Theorem 6.8 *For any positive integer $n \ge 3$, let $A_1, \cdots, A_n$ be positive invertible operators such that $mI_H \le A_i \le MI_H$ for $i = 1, 2, \cdots, n$ and some scalars $0 < m < M$. Put $h = \frac{M}{m}$. Then for $0 < t < 1$*

$$A[n,t](A_1, \cdots, A_n) \le S(h)^2\ G[n,t](A_1, \cdots, A_n).$$

Proof. By Lemma 6.10, it follows that

$$A[n,t](A_1^{-1},\cdots,A_n^{-1}) \le S(h)\ \exp\left(A[n,t](\log A_1^{-1},\cdots,\log A_n^{-1})\right).$$

Taking inverse, we have

$$H[n,t](A_1,\cdots,A_n) \ge S(h)^{-1}\ \exp\left(A[n,t](\log A_1,\cdots,\log A_n)\right)$$

and this implies

$$\begin{aligned} A[n,t](A_1,\cdots,A_n) &\le S(h)\ \exp\left(A[n,t](\log A_1,\cdots,\log A_n)\right) \\ &\le S(h)^2\ H[n,t](A_1,\cdots,A_n). \end{aligned}$$

Therefore, we have

$$\begin{aligned} A[n,t](A_1,\cdots,A_n) &\le S(h)^2\ H[n,t](A_1,\cdots,A_n) \\ &\le S(h)^2\ G[n,t](A_1,\cdots,A_n) \end{aligned}$$

and we have this theorem. □

By using Theorem 6.8 and the weighted arithmetic-geometric mean inequality, we obtain another n operators version of Grueb-Rheinboldt inequality:

Theorem 6.9 *For any positive integer $n \ge 3$, let $A_1,\cdots,A_n$ be positive invertible operators on a Hilbert space H such that $0 < mI_H \le A_i \le MI_H$ for $i = 1,\cdots,n$ and some scalars $0 < m < M$. Put $h = \frac{M}{m}$. Then*

$$(A_1x,x)^{t[n]_1}(A_2x,x)^{t[n]_2}\cdots(A_nx,x)^{t[n]_n} \le S(h)^2(G[n,t](A_1,\cdots,A_n)x,x)$$

for all $x \in H$, where $\{t[n]_i\}$ is defined by (6.7) and (6.8).

6.5 The Golden-Thompson-Segal inequality

For the construction of nonlinear relativistic quantum fields, Segal proved that

$$\|e^{H+K}\| \ \le\ \|e^H e^K\|.$$

Also, motivated by quantum statistical mechanics, Golden, Symanzik and Thompson independently proved that

$$\text{Tr } e^{H+K} \le \text{Tr } e^H e^K$$

holds for Hermitian matrices H and K. This inequality is called Golden-Thompson trace inequality.

In the final section, we discuss the Golden-Thompson-Segal type inequalities for the operator norm. Ando and Hiai gave a lower bound on $\|e^{H+K}\|$ in terms of the geometric mean: For two self-adjoint operators H and K and $\alpha \in [0,1]$,

$$\| \left(e^{pH} \#_\alpha e^{pK}\right)^{\frac{1}{p}} \| \le \|e^{(1-\alpha)H+\alpha K}\| \tag{6.14}$$

holds for all $p > 0$ and the left-hand side of (6.14) converges to the right-hand side as $p \downarrow 0$.

Hiai and Petz showed the following geometric mean version of the Lie-Trotter formula: If A and B are positive invertible and $t \in [0,1]$, then

$$\lim_{p\to 0} (A^p \#_t B^p)^{\frac{1}{p}} = e^{(1-t)\log A + t\log B}.$$

We firstly show an n-variable version of the Lie-Trotter formula for the weighted geometric mean:

Lemma 6.11 *Let $A_1, A_2, \cdots, A_n$ be positive invertible operators such that $mI_H \le A_i \le MI_H$ for $i = 1,\cdots,n$ and some scalars $0 < m \le M$, and let $\lambda_1,\cdots,\lambda_n \in [0,1]$ such that $\sum_{i=1}^n \lambda_i = 1$. Then $G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}$ uniformly converges to the chaotically geometric mean $e^{A[n,t](\log A_1,\cdots,\log A_n)}$ as $p \downarrow 0$.*

Proof. It follows that for each $\lambda_i > 0$ such that $\sum_{i=1}^n \lambda_i = 1$,

$$0 \le \log \sum_{i=1}^n \lambda_i A_i - \sum_{i=1}^n \lambda_i \log A_i \le \log S(h).$$

In particular, we have

$$0 \le \log A[n,t](A_1,\cdots,A_n) - A[n,t](\log A_1,\cdots,\log A_n) \le \log S(h).$$

Replacing A_i by A_i^p for $p > 0$,

$$0 \le \log A[n,t](A_1^p,\cdots,A_n^p) - A[n,t](\log A_1^p,\cdots,\log A_n^p) \le \log S(h^p)$$

and hence

$$0 \le \log A[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}} - A[n,t](\log A_1,\cdots,\log A_n) \le \log S(h^p)^{\frac{1}{p}}.$$

Since $S(h^p)^{\frac{1}{p}} \to 1$ as $p \downarrow 0$, it follows that $A[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}$ uniformly converges to the chaotically geometric mean $e^{A[n,t](\log A_1,\cdots,\log A_n)}$ as $p \downarrow 0$.

On the other hand, since

$$0 \le \log A[n,t](A_1^{-1},\cdots,A_n^{-1}) - A[n,t](\log A_1^{-1},\cdots,\log A_n^{-1}) \le \log S(h^{-1}),$$

it follows from $S(h^{-1}) = S(h)$ that

$$0 \ge \log H[n,t](A_1,\cdots,A_n) - A[n,t](\log A_1,\cdots,\log A_n) \ge -\log S(h)$$

and this implies

$$0 \geq \log H[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}} - A[n,t](\log A_1,\cdots,\log A_n) \geq -\log S(h^p)^{\frac{1}{p}}$$

for all $p > 0$. Hence $H[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}$ uniformly converges to the chaotically geometric mean $e^{A[n,t](\log A_1,\cdots,\log A_n)}$ as $p \downarrow 0$.

By arithmetic-geometric-harmonic mean inequality, we have

$$\log H[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}} \leq \log G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}} \leq \log A[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}$$

for all $p > 0$ and hence we have this lemma. □

For the case of $n = 2$, Ando and Hiai are showed that the norm $\|(A_1^p \ \#_t \ A_2^p)^{\frac{1}{p}}\|$ is monotone increasing for $p > 0$. For $n \geq 3$, we have the following result.

Lemma 6.12 *Let A_i be positive invertible operators such that $mI_H \leq A_i \leq MI_H$ for $i = 1,\cdots,n$ and some scalars $0 < m \leq M$. Put $h = \frac{M}{m}$. Then for each $0 < q < p$*

$$\begin{aligned} &S(h^p)^{-\frac{2}{p}} \ \|G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}\| \\ &\qquad \leq \|G[n,t](A_1^q,\cdots,A_n^q)^{\frac{1}{q}}\| \leq S(h^p)^{\frac{2}{p}} \ \|G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}\|, \end{aligned}$$

where $S(h)$ is defined by (2.35).

Proof. By the arithmetic-geometric mean inequality, it follows that for each $0 < q < p$

$$\begin{aligned} G[n,t](A_1^{\frac{q}{p}},\cdots,A_n^{\frac{q}{p}}) &\leq A[n,t](A_1^{\frac{q}{p}},\cdots,A_n^{\frac{q}{p}}) \\ &\leq A[n,t](A_1,\cdots,A_n)^{\frac{q}{p}} \text{ by concavity of } t^{\frac{q}{p}} \text{ and } 0 < \tfrac{q}{p} < 1 \\ &\leq S(h)^{\frac{2q}{p}} G[n,t](A_1,\cdots,A_n)^{\frac{q}{p}} \end{aligned}$$

The last inequality follows from Theorem 6.8 and the Löwner-Heinz theorem. Replacing A_i by A_i^p, we have

$$G[n,t](A_1^q,\cdots,A_n^q) \leq S(h^p)^{\frac{2q}{p}} G[n,t](A_1^p,\cdots,A_n^p)^{\frac{q}{p}}.$$

Also,

$$G[n,t](A_1^{-q},\cdots,A_n^{-q}) \leq S(h^{-p})^{\frac{2q}{p}} G[n,t](A_1^{-p},\cdots,A_n^{-p})^{\frac{q}{p}}$$

and hence

$$G[n,t](A_1^q,\cdots,A_n^q) \geq S(h^p)^{-\frac{2q}{p}} G[n,t](A_1^p,\cdots,A_n^p)^{\frac{q}{p}}.$$

Therefore we have for all $q > 0$

$$\begin{aligned} &S(h^p)^{-\frac{2}{p}} \ \|G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}\| \\ &\qquad \leq \|G[n,t](A_1^q,\cdots,A_n^q)^{\frac{1}{q}}\| \leq S(h^p)^{\frac{2}{p}} \ \|G[n,t](A_1^p,\cdots,A_n^p)^{\frac{1}{p}}\|. \end{aligned}$$

□

By Lemma 6.12, we show n-variable versions of a complement of the Golden-Thompson-Segal type inequality due to Ando and Hiai:

Theorem 6.10 *Let $H_1, H_2, \cdots, H_n$ be self-adjoint operators such that $mI_H \le H_i \le MI_H$ for $i = 1, \cdots, n$ and some scalars $m \le M$. Then*

$$S\left(e^{p(M-m)}\right)^{-\frac{2}{p}} \|G[n,t](e^{pH_1}, \cdots, e^{pH_n})^{\frac{1}{p}}\|$$

$$\le \|e^{A[n,t](H_1,\cdots,H_n)}\| \le S\left(e^{p(M-m)}\right)^{\frac{2}{p}} \|G[n,t](e^{pH_1}, \cdots, e^{pH_n})^{\frac{1}{p}}\| \quad (6.15)$$

for all $p > 0$ and the both-hand sides of (6.15) *converge to the middle-hand side as $p \downarrow 0$, where the Specht ratio $S(h)$ is defined by* (2.35).

Proof. If we replace A_i by e^{H_i} in Lemma 6.12, then it follows that

$$S\left(e^{p(M-m)}\right)^{-\frac{2}{p}} \|G[n,t](e^{pH_1}, \cdots, e^{pH_n})^{\frac{1}{p}}\|$$

$$\le \|G[n,t](e^{qH_1}, \cdots, e^{qH_n})^{\frac{1}{q}}\| \le S\left(e^{p(M-m)}\right)^{\frac{2}{p}} \|G[n,t](e^{pH_1}, \cdots, e^{pH_n})^{\frac{1}{p}}\|$$

for all $0 < q < p$. Hence we have (6.15) as $q \downarrow 0$ by Lemma 6.11.

The latter part of this theorem follows from $S\left(e^{p(M-m)}\right)^{\frac{2}{p}} \to 1$ as $p \downarrow 0$. □

6.6 Notes

For our exposition we have used Ando-Li-Mathias [13], Yamazaki [292], J.I. Fujii-M. Fujii-Nakamura-Pečarić-Seo [60] and J.I. Fujii-M. Fujii-Seo [63].

Chapter 7

Differential Geometry of Operators

In this chapter, we study some differential-geometrical structure of operators. The space of positive invertible operators of a unital C*-algebra has the natural structure of a reductive homogenous manifold with a Finsler metric. Then a pair of points A and B can be joined by a unique geodesic $A \#_t B$ for $t \in [0,1]$ and we consider estimates of the upper bounds for the difference between the geodesic and extended interpolational paths by terms of the spectra of positive operators.

7.1 Introduction

We recall the Kubo-Ando theory of operator means [165]: A mapping $(A,B) \to A \ \sigma \ B$ in the cone of positive invertible operators is called an operator mean if the following conditions are satisfied:

Monotonicity $A \le C$ and $B \le D$ imply $A \ \sigma \ B \le C \ \sigma \ D$.

Upper continuity $A_n \downarrow A$ and $B_n \downarrow B$ imply $A_n \ \sigma B_n \downarrow A \ \sigma \ B$.

Transformer inequality $T^*(A\sigma B)T \le (T^*AT)\sigma(T^*BT)$ for all operator T.

Normalized condition $A \,\sigma\, A = A.$

In [124, Chapter 5], several inequalities associated with operator means are discussed. For example, the bound β in the inequality

$$\Phi(A\ \sigma_1\ B) \geq \alpha\Phi(A)\ \sigma_2\ \Phi(B) + \beta\Phi(A)$$

is determined, where A and B are positive invertible operators on a Hilbert space H, σ_1, σ_2 are two operator means with not affine representing functions, Φ is a unital positive linear mapping and $\alpha > 0$ is a given real constant.

We observe the weighted arithmetic mean ∇_α and the weighted geometric mean $\#_\alpha$, for $\alpha \in [0,1]$, defined by

$$A\ \nabla_\alpha\ B := (1-\alpha)A + \alpha B \quad \text{and} \quad A\ \#_\alpha\ B := A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha A^{\frac{1}{2}},$$

respectively. Like the numerical case, the arithmetic-geometric mean inequality holds:

$$A\ \#_\alpha\ B \leq A\ \nabla_\alpha\ B \qquad \text{for all } \alpha \in [0,1]. \tag{7.1}$$

In [124, Corollary 5.36] it is obtained the following converse inequality of the arithmetic-geometric mean inequality (7.1): Let A and B be positive invertible operators satisfying $0 < m_1 I_H \leq A \leq M_1 I_H$ and $0 < m_2 I_H \leq B \leq M_2 I_H$. Then

$$A\ \nabla_\alpha\ B - A\ \#_\alpha\ B \leq \max\{1-\alpha+\alpha m - m^\alpha, 1-\alpha+\alpha M - M^\alpha\}A,$$

where $m = \frac{m_2}{M_1}$ and $M = \frac{M_2}{m_1}$.

Tominaga [280] showed the another converse of (7.1) for the arithmetic mean and the geometric one: Let A and B be positive operators on a Hilbert space H satisfying $mI_H \leq A, B \leq I_H$ for some scalars $0 < m < M$. Then (like the numerical case)

$$A\ \nabla_\alpha\ B - A\ \#_\alpha\ B \leq hL(m,M)\log S(h) \qquad \text{for all } \alpha \in [0,1],$$

where $h = \frac{M}{m}$, the logarithmic mean $L(m,M)$ is defined by (2.41) and the Specht ratio $S(h)$ is defined by (2.35).

7.2 Interpolational paths

Let $\mathscr{A}$ be a unital C*-algebra, $\mathscr{A}^+$ (resp. $\mathscr{A}^h$) be the set of all positive invertible (resp. self-adjoint) operators of $\mathscr{A}$. Following an excellent work due to Corach, Porta and Recht [37, 38], $\mathscr{A}^+$ is a real analytic open submanifold of $\mathscr{A}^h$ and its tangent space $(T\mathscr{A}^+)_A$ at any $A \in \mathscr{A}^+$ is naturally identified to $\mathscr{A}^h$. For each $A \in \mathscr{A}^+$, the norm $\|X\|_A = \|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}\|$,

$X \in (T\mathscr{A}^+)_A$ defined a Finslar structure on the tangent bundle $T\mathscr{A}^+$. For every $A, B \in \mathscr{A}^+$, there is a unique geodesic joining A and B:

$$\gamma_{A,B}(t) = A \ \#_t \ B \qquad \text{for } t \in [0,1].$$

As usual, the length of a smooth curve γ in $\mathscr{A}^+$ is defined by

$$l(\gamma) = \int_0^1 \|\dot{\gamma}(t)\|_{\gamma(t)} dt$$

and the geodesic distance between A and B in $\mathscr{A}^+$ is

$$d(A,B) = \inf\{l(\gamma) : \ \gamma \text{ joins } A \text{ and } B\}.$$

Then it follows that

$$d(A,B) = \|\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})\|,$$

also see [14]. It is a general fact that $(\mathscr{A}^+, d)$ is a complete metric space.

J.I.Fujii [55] showed that if the manifold $\mathscr{A}^+$ has a metric $L_a(X) = \|X\|$ (resp. $L_h(X) = \|A^{-1}XA^{-1}\|$) on the tangent space $T\mathscr{A}^+$, the geodecis and the distance from A to B for $A, B \in \mathscr{A}^+$ are given by

$$A \ \nabla_t \ B = (1-t)A + tB \quad \text{and} \quad d_1(A,B) = \|B - A\|$$

$$\text{(resp.} \quad A \ !_t \ B = \left((1-t)A^{-1} + tB^{-1}\right)^{-1} \quad \text{and} \quad d_{-1}(A,B) = \|A^{-1} - B^{-1}\|.\text{)}$$

The paths of means $m_t = \#_t, \nabla_t$ and $!_t$ satisfy the following interpolationality [89]:

$$(A \ m_p \ B) \ m_t \ (A \ m_q \ B) = A \ m_{(1-t)p+tq} \ B$$

for $0 \le p, q, t \le 1$.

We next recall an interpolational path for symmetric operator means. Following after [89, 96], for a symmetric mean σ, a parametrized operator mean σ_t is called an interpolational path for σ if it satisfies

(1) $A \ \sigma_0 \ B = A$, $A \ \sigma_{1/2} \ B = A \ \sigma \ B$ and $A \ \sigma_1 \ B = B$,

(2) $(A \ \sigma_p \ B) \ \sigma \ (A \ \sigma_q \ B) = A \ \sigma_{\frac{p+q}{2}} \ B$,

(3) the mapping $t \mapsto A \ \sigma_t \ B$ is norm continuous for each A and B.

Typical examples of symmetric means are so-called power means:

$$A \ m_r \ B = A^{\frac{1}{2}} \left(\frac{1 + (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r}{2} \right)^{\frac{1}{r}} A^{\frac{1}{2}} \qquad \text{for } r \in [-1,1]$$

and their interpolational paths from A to B via $A \ m_r \ B$ are given as follows: For each $r \in [-1,1]$

$$A \ m_{r,t} \ B = A^{\frac{1}{2}} \left(1 - t + t\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r\right)^{\frac{1}{r}} A^{\frac{1}{2}} \qquad \text{for } t \in [0,1].$$

In particular, $A\ m_{1,t}\ B = A\ \nabla_t\ B$, $A\ m_{0,t}\ B = A\ \#_t\ B$ and $A\ m_{-1,t}\ B = A\ !_t\ B$.

Here we consider them in a general setting: For positive invertible operators A and B, an extended path $A\ m_{r,t}\ B$ is defined as

$$A\ m_{r,t}\ B = A^{\frac{1}{2}}\left(1-t+t(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^r\right)^{\frac{1}{r}}A^{\frac{1}{2}} \qquad \text{for all } r\in\mathbb{R} \text{ and } t\in[0,1].$$

The representing function $f_{r,t}$ for $m_{r,t}$ is given by

$$f_{r,t}(\xi) = 1\ m_{r,t}\ \xi = (1-t+t\xi^r)^{\frac{1}{r}} \qquad \text{for } \xi>0.$$

Notice that $A\ m_{r,t}\ B$ for $r\notin[-1,1]$ is no longer an operator mean, but we list some properties of interpolational paths $m_{r,t}$ and the representing function $f_{r,t}$, also see [62].

Since every function $f_{r,t}(\xi)$ is strictly increasing and strictly convex (resp. strictly concave) for $r>1$ (resp. $r<1$), it follows that an extended path $A\ m_{r,t}\ B$ for each $t\in(0,1)$ is nondecreasing and norm continuous for $r\in\mathbb{R}$: For $r\le s$

$$A\ m_{r,t}\ B \le A\ m_{s,t}\ B.$$

Moreover, it is also interpolational for all $r\in\mathbb{R}$. In particular, the transposition formula holds:

$$B\ m_{r,t}\ A = A\ m_{r,1-t}\ B. \tag{7.2}$$

For the sake of convenience, we prepare the following notation: For $k_2>k_1>0$, $r\in\mathbb{R}$ and $t\in[0,1]$

$$a(r,t) = \frac{f_{r,t}(k_2)-f_{r,t}(k_1)}{k_2-k_1} \quad \text{and} \quad b(r,t) = \frac{k_2f_{r,t}(k_1)-k_1f_{r,t}(k_2)}{k_2-k_1}. \tag{7.3}$$

We investigate estimates of the upper bounds for the difference between extended interpolational paths:

Lemma 7.1 *Let A and B be positive invertible operators such that $k_1A\le B\le k_2A$ for some scalars $0<k_1<k_2$. Then for $r\le s$ and $t\in(0,1)$*

$$0\le A\ m_{s,t}\ B - A\ m_{r,t}\ B \le \beta A \qquad \text{if} \quad r\le 1 \tag{7.4}$$

and

$$0\le A\ m_{s,t}\ B - A\ m_{r,t}\ B \le \beta' A \qquad \text{if} \quad r\ge 1 \tag{7.5}$$

hold for

$$\beta = \beta(r,s,t,k_1,k_2) = \max_{k_1\le\xi\le k_2}\{f_{s,t}(\xi)-a(r,t)\xi-b(r,t)\}$$

and

$$\beta' = \beta'(r,s,t,k_1,k_2) = \max_{k_1\le\xi\le k_2}\{a(s,t)\xi+b(s,t)-f_{r,t}(\xi)\},$$

where a,b are defined by (7.3)*.*

Proof. Suppose that $r \le 1$. If we put $C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$, then we have $k_2 I_H \ge C \ge k_1 I_H > 0$. Since $f_{r,t}(\xi)$ is concave for $r \le 1$, it follows from the definition of β that

$$\beta \ge f_{s,t}(\xi) - a(r,t)\xi - b(r,t) \ge f_{s,t}(\xi) - f_{r,t}(\xi) \qquad \text{for all } \xi \in [k_1, k_2],$$

and hence

$$\beta I_H \ge f_{s,t}(C) - f_{r,t}(C).$$

This fact implies

$$\beta A \ge A^{\frac{1}{2}} f_{s,t}(C) A^{\frac{1}{2}} - A^{\frac{1}{2}} f_{r,t}(C) A^{\frac{1}{2}} = A\ m_{s,t}\ B - A\ m_{r,t}\ B,$$

which gives the desired result (7.4). Conversely, if $r \ge 1$, then $f_{s,t}(\xi)$ is convex for $1 \le r \le s$ and (7.5) follows from the same way. □

Remark 7.1 *The constant $\beta = \beta(s,r,t,k_1,k_2)$ and $\beta' = \beta'(s,r,t,k_1,k_2)$ in Lemma 7.1 can be written explicitly as*

$$\beta = \begin{cases} a(r,t)\left(\dfrac{1-t}{t}\right)^{\frac{1}{s}} \left(t^{\frac{1}{s-1}} a(r,t)^{\frac{s}{1-s}} - 1\right)^{\frac{s-1}{s}} - b(r,t) & \text{if} \quad k_1 \le \xi_0 \le k_2 \\ f_{s,t}(k_1) - f_{r,t}(k_1) & \text{if} \quad \xi_0 \le k_1 \\ f_{s,t}(k_2) - f_{r,t}(k_2) & \text{if} \quad k_2 \le \xi_0 \end{cases}$$

where $\xi_0 = \left(\frac{1}{1-t}\left(\frac{a(r,t)}{t}\right)^{\frac{s}{1-s}} - \frac{t}{1-t}\right)^{-\frac{1}{s}}$ *and*

$$\beta' = \begin{cases} -a(s,t)\left(\dfrac{1-t}{t}\right)^{\frac{1}{r}} \left(t^{\frac{1}{r-1}} a(s,t)^{\frac{r}{1-r}} - 1\right)^{\frac{r-1}{r}} + b(s,t) & \text{if} \quad k_1 \le \xi_1 \le k_2 \\ f_{s,t}(k_1) - f_{r,t}(k_1) & \text{if} \quad \xi_1 \le k_1 \\ f_{s,t}(k_2) - f_{r,t}(k_2) & \text{if} \quad k_2 \le \xi_1 \end{cases}$$

where $\xi_1 = \left(\frac{1}{1-t}\left(\frac{a(s,t)}{t}\right)^{\frac{r}{1-r}} - \frac{t}{1-t}\right)^{-\frac{1}{r}}.$

By Lemma 7.1, we obtain estimates of the upper bounds for the difference between the geodesic $A\ \#_t\ B$ and extended interpolational paths:

Theorem 7.1 *Let A and B be positive invertible operators such that $k_1 A \le B \le k_2 A$ for some scalars $0 < k_1 < k_2$. Then for each $t \in (0,1)$*

$$0 \le A\ m_{s,t}\ B - A\ \#_t\ B \le \beta(0,s,t,k_1,k_2)A \qquad \text{for } s \ge 0 \tag{7.6}$$

and

$$0 \le A\ \#_t\ B - A\ m_{r,t}\ B \le \beta(r,0,t,k_1,k_2)A \qquad \text{for } r \le 0, \tag{7.7}$$

where β is defined by Remark 7.1.

As special cases of Theorem 7.1, we obtain an estimate of the upper bound for the difference between the geodesic $A \#_t B$ and the arithmetic interpolational paths $A \nabla_t B$, the harmonic one $A \,!_t\, B$:

Theorem 7.2 *Let A and B be positive invertible operators such that $k_1 A \le B \le k_2 A$ for some scalars $0 < k_1 < k_2$. Then for each $t \in (0,1)$*

$$0 \le A \nabla_t B - A \#_t B \le \max\{1-t+tk_1-k_1^t, 1-t+tk_2-k_2^t\}A$$

and

$$0 \le A \#_t B - A \,!_t\, B \le \beta(-1,0,t,k_1,k_2)A,$$

where
$\beta(-1,0,t,k_1,k_2) =$

$$= \begin{cases} \dfrac{1-t}{\big((1-t)k_1+t\big)\big((1-t)k_2+t\big)} \left(\big((1-t)k_1+t\big)\big((1-t)k_2+t\big)^{\frac{1}{1-t}} - k_1k_2\right) \\ \qquad\qquad \text{if} \quad k_1^{1-t} \le \big((1-t)k_1+t\big)\big((1-t)k_2+t\big) \le k_2^{1-t} \\ k_2^t - \dfrac{k_2}{(1-t)k_2+t} \qquad \text{if} \quad k_2^{1-t} \le \big((1-t)k_1+t\big)\big((1-t)k_2+t\big) \\ k_1^t - \dfrac{k_1}{(1-t)k_1+t} \qquad \text{if} \quad k_1^{1-t} \ge \big((1-t)k_1+t\big)\big((1-t)k_2+t\big). \end{cases}$$

Proof. If we put $r = 0$ and $s = 1$ in (7.6) of Theorem 7.1 , then $f_{1,t}(\xi) = 1-t+t\xi$ and $f_{0,t}(\xi) = \xi^t$. Since $a(0,t) = \frac{k_2^t-k_1^t}{k_2-k_1}$, the condition $f'_{1,t}(k_2) \le a(0,t) \le f'_{1,t}(k_1)$ is equivalent to $a(0,t) = t$. Therefore we have

$$\beta = \begin{cases} 1-t+tk_2-k_2^t & \text{if} \quad \dfrac{k_2^t-k_1^t}{k_2-k_1} \le t \\ 1-t+tk_1-k_1^t & \text{if} \quad \dfrac{k_2^t-k_1^t}{k_2-k_1} \ge t. \end{cases}$$

Similarly, we have the latter part of this theorem by using (7.7) in Theorem 7.1. □

Next, we show estimates of the lower bounds of the ratio for extended interpolational paths:

Lemma 7.2 *Let A and B be positive invertible operators such that $k_1 A \le B \le k_2 A$ for some scalars $0 < k_1 < k_2$. Then for $r \le s$ and $t \in (0,1)$*

$$A \, m_{r,t} \, B \ge \alpha A \, m_{s,t} \, B \qquad \text{if} \quad r \le 1$$

and

$$A \, m_{r,t} \, B \ge \alpha' A \, m_{s,t} \, B \qquad \text{if} \quad r \ge 1$$

hold for

$$\alpha = \alpha(r,s,t,k_1,k_2) = \min_{k_1 \le \xi \le k_2} \left\{ \frac{a(r,t)\xi + b(r,t)}{f_{s,t}(\xi)} \right\}$$

and

$$\alpha' = \alpha'(r,s,t,k_1,k_2) = \min_{k_1 \le \xi \le k_2} \left\{ \frac{f_{r,t}(\xi)}{a(s,t)\xi + b(s,t)} \right\},$$

where a, b are defined by (7.3).

Proof. Suppose that $r < 1$. Since $f_{r,t}(\xi)$ is concave for $r < 1$, it follows that

$$\frac{f_{r,t}(\xi)}{f_{s,t}(\xi)} \ge \frac{a(r,t)\xi + b(r,t)}{f_{s,t}(\xi)} \ge \alpha$$

and hence $f_{r,t}(\xi) \ge \alpha f_{s,t}(\xi)$ on $[k_1, k_2]$. Therefore we have

$$A\ m_{r,t}\ B = A^{\frac{1}{2}} f_{r,t}(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}} \ge \alpha A^{\frac{1}{2}} f_{s,t}(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}} = \alpha A\ m_{s,t}\ B.$$

Similarly, since $f_{s,t}(\xi)$ is convex for $1 \le r \le s$, the latter part follows from the same way. □

Remark 7.2 *The constant $\alpha = \alpha(r,s,t,k_1,k_2)$ and $\alpha' = \alpha'(r,s,t,k_1,k_2)$ in Lemma 7.2 can be written explicitly as follows: In the case of $s \ge 1$,*

$$\alpha = \alpha' = \min \left\{ \frac{f_{r,t}(k_1)}{f_{s,t}(k_1)}, \frac{f_{r,t}(k_2)}{f_{s,t}(k_2)} \right\}.$$

In the case of $s \le 1$,

$$\alpha = \begin{cases} \dfrac{a(r,t)\xi_0 + b(r,t)}{(1 - t + t\xi_0^s)^{\frac{1}{s}}} & \text{if} \quad k_1 \le \xi_0 \le k_2 \\ \dfrac{f_{r,t}(k_2)}{f_{s,t}(k_2)} & \text{if} \quad k_2 \le \xi_0 \\ \dfrac{f_{r,t}(k_1)}{f_{s,t}(k_1)} & \text{if} \quad k_1 \ge \xi_0, \end{cases}$$

where $\xi_0 = \left(\dfrac{1-t}{t} \dfrac{a(r,t)}{b(r,t)} \right)^{\frac{1}{s-1}}$ *and*

$$\alpha' = \begin{cases} \dfrac{(1 - t + t\xi_1^r)^{\frac{1}{r}}}{a(s,t)\xi_1 + b(s,t)} & \text{if} \quad k_1 \le \xi_1 \le k_2 \\ \dfrac{f_{r,t}(k_2)}{f_{s,t}(k_2)} & \text{if} \quad k_2 \le \xi_1 \\ \dfrac{f_{r,t}(k_1)}{f_{s,t}(k_1)} & \text{if} \quad k_1 \ge \xi_1, \end{cases}$$

where $\xi_1 = \left(\dfrac{1-t}{t} \dfrac{a(s,t)}{b(s,t)} \right)^{\frac{1}{r-1}}.$

By Lemma 7.2, we have the following theorem.

Theorem 7.3 *Let A and B be positive invertible operators such that $k_1A \le B \le k_2A$ for some scalars $0 < k_1 < k_2$. Then for each $t \in (0,1)$*

$$A \#_t B \ge \min\left\{\frac{k_1^t}{1-t+tk_1}, \frac{k_2^t}{1-t+tk_2}\right\} A \nabla_t B$$

and

$$A \,!_t\, B \ge \alpha(-1,0,t,k_1,k_2)\, A \#_t B$$

holds for

$$\alpha(-1,0,t,k_1,k_2) = \begin{cases} \dfrac{(k_1k_2)^{1-t}}{\big((1-t)k_1+t\big)\big((1-t)k_2+t\big)} & \text{if} \quad k_1 \le 1 \le k_2 \\ \dfrac{k_2^{1-t}}{(1-t)k_2+t} & \text{if} \quad 1 \le k_1 \\ \dfrac{k_1^{1-t}}{(1-t)k_1+t} & \text{if} \quad k_2 \le 1. \end{cases}$$

7.3 Velocity vector of extended paths

Kamei and Fujii [67, 68] defined the relative operator entropy $S(A|B)$, for positive invertible operators A and B, by

$$S(A|B) = A^{\frac{1}{2}}\left(\log A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)A^{\frac{1}{2}},$$

which is a relative version of the operator entropy $-A\log A$ considered by Nakamura-Umegaki [238]. The relative operator entropy $S(A|B)$ is exactly the velocity vector $\dot{\gamma}_{A,B}(0)$ of the geodesic $A \#_t B$ at $t = 0$:

$$S(A|B) = \lim_{t\to 0}\frac{A \#_t B - A \#_0 B}{t} = \dot{\gamma}_{A,B}(0).$$

In [153], Kamei analogously generalizes the relative operator entropy: For each $r \in \mathbb{R}$

$$S_r(A|B) = \lim_{t\to 0}\frac{A\, m_{r,t}\, B - A\, m_{r,0}\, B}{t},$$

which is considered as the right differential coefficient at $t = 0$ of the extended path $A\, m_{r,t}\, B$. By the fact that

$$\lim_{t\to 0}\frac{(1-t+t\xi^r)^{\frac{1}{r}}-1}{t} = \frac{\xi^r-1}{r},$$

it follows that

$$S_r(A|B) = \frac{A^{\frac{1}{2}}\left(A^{-\frac{1}{2}}BA^{-\frac{1}{2}}\right)^r A^{\frac{1}{2}} - A}{r} \qquad \text{for } r \in \mathbb{R}$$

and the representing function is

$$f_r(\xi) = (\xi^r - 1)/r.$$

In particular,

$$S_1(A|B) = \lim_{t\to 0}\frac{A \nabla_t B - A}{t} = B - A$$
$$S_0(A|B) = S(A|B)$$
$$S_{-1}(A|B) = \lim_{t\to 0}\frac{A\, !_t\, B - A}{t} = A - AB^{-1}A.$$

Since $f_r(\xi)$ is monotone increasing on $r \in \mathbb{R}$, the velocity vectors $S_r(A|B)$ is monotone increasing on $r \in \mathbb{R}$:

$$r \le s \qquad \text{implies} \qquad S_r(A|B) \le S_s(A|B).$$

The left differentiable coefficient of $A\ m_{r,t}\ B$ at $t = 1$ is $-S_r(B|A)$:

$$\lim_{t\to 1}\frac{A\ m_{r,t}\ B - A\ m_{r,1}\ B}{t-1} = -S_r(B|A).$$

If $B \ge A$, then the velocity vectors of extended paths at $t = 0, 1$ are positive:

$$S_r(A|B) \ge 0 \qquad \text{and} \qquad -S_r(B|A) \ge 0.$$

For the sake of convenience, we prepare the following notation:

$$a(r) = \frac{f_r(k_2) - f_r(k_1)}{k_2 - k_1} \qquad \text{and} \qquad b(r) = \frac{k_2 f_r(k_1) - k_1 f_r(k_2)}{k_2 - k_1} \tag{7.8}$$

for $0 < k_1 < k_2$ and $r \in \mathbb{R}$.

We investigate estimates of the upper bounds for the difference between velocity vectors of extended interpolational paths.

Lemma 7.3 *Let A and B be positive invertible operators such that $k_1A \le B \le k_2A$ for some scalars $0 < k_1 < k_2$. Then for $r \le s$*

$$S_s(A|B) - S_r(A|B) \le \gamma A \qquad \text{if } r \le s \le 1, \tag{7.9}$$

$$S_s(A|B) - S_r(A|B) \le \max\left\{\frac{k_1^s - 1}{s} - \frac{k_1^r - 1}{r}, \frac{k_2^s - 1}{s} - \frac{k_2^r - 1}{r}\right\}A \quad \text{if } r \le 1 \le s \tag{7.10}$$

and

$$S_s(A|B) - S_r(A|B) \le \gamma' A \qquad \text{if } 1 \le r \le s$$

hold for

$$\gamma = \gamma(r,s,k_1,k_2) = \max_{k_1 \le \xi \le k_2} \{f_s(\xi) - a(r)\xi - b(r)\}$$

and

$$\gamma' = \gamma'(r,s,k_1,k_2) = \max_{k_1 \le \xi \le k_2} \{a(s)\xi + b(s) - f_r(\xi)\},$$

where a, b *are defined by* (7.8).

Proof. Suppose that $r \le 1$. If we put $C = A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$, then we have $0 < k_1 I_H \le C \le k_2 I_H$. Since $f_r(\xi)$ is concave for $r \le 1$, it follows that

$$f_s(\xi) - f_r(\xi) \le f_s(\xi) - a(r)\xi - b(r) \le \gamma$$

and hence we have the desired result (7.9) and (7.10). The remainder parts follow from the same way. □

Remark 7.3 *The constant* $\gamma = \gamma(r,s,k_1,k_2)$ *and* $\gamma' = \gamma'(r,s,k_1,k_2)$ *in Lemma 7.3 can be written explicitly as*

$$\gamma = \begin{cases} \dfrac{1-s}{s} a(r)^{\frac{s}{s-1}} - b(r) - \dfrac{1}{s} & \text{if} \quad k_1 \le a(r)^{\frac{1}{s-1}} \le k_2 \\ \dfrac{k_2^s - 1}{s} - \dfrac{k_2^r - 1}{r} & \text{if} \quad k_2 \le a(r)^{\frac{1}{s-1}} \\ \dfrac{k_1^s - 1}{s} - \dfrac{k_1^r - 1}{r} & \text{if} \quad k_1 \ge a(r)^{\frac{1}{s-1}} \end{cases}$$

and

$$\gamma' = \begin{cases} \dfrac{r-1}{r} a(s)^{\frac{r}{r-1}} + b(s) + \dfrac{1}{r} & \text{if} \quad k_1 \le a(s)^{\frac{1}{r-1}} \le k_2 \\ \dfrac{k_2^s - 1}{s} - \dfrac{k_2^r - 1}{r} & \text{if} \quad k_2 \le a(s)^{\frac{1}{r-1}} \\ \dfrac{k_1^s - 1}{s} - \dfrac{k_1^r - 1}{r} & \text{if} \quad k_1 \ge a(s)^{\frac{1}{r-1}}. \end{cases}$$

By Lemma 7.3, we obtain estimates of the upper bound for the difference between the velocity vectors $S(A|B)$ and $S_r(A|B)$ of the extended interpolational paths $A\ m_{r,t}\ B$ at $t = 0$:

Theorem 7.4 *Let* A *and* B *be positive invertible operators such that* $k_1 A \le B \le k_2 A$ *for some scalars* $0 < k_1 < k_2$. *Then*

$$S_s(A|B) - S(A|B) \le \gamma A \qquad \text{for} \quad 0 \le s \le 1$$

and

$$S_s(A|B) - S(A|B) \le \max\left\{\frac{k_1^s - 1}{s} - \log k_1, \frac{k_2^s - 1}{s} - \log k_2\right\} A \qquad \text{for} \quad 1 \le s,$$

where

$$\gamma = \begin{cases} \dfrac{1-s}{s}\left(\dfrac{\log k_2 - \log k_1}{k_2 - k_1}\right)^{\frac{s}{s-1}} - \dfrac{k_2 \log k_1 - k_1 \log k_2}{k_2 - k_1} - \dfrac{1}{s} & \\ \qquad\qquad \textit{if} \quad k_1 \le \left(\dfrac{\log k_2 - \log k_1}{k_2 - k_1}\right)^{\frac{1}{s-1}} \le k_2 \\ \dfrac{k_2^s - 1}{s} - \log k_2 & \textit{if} \quad k_2 \le \left(\dfrac{\log k_2 - \log k_1}{k_2 - k_1}\right)^{\frac{1}{s-1}} \\ \dfrac{k_1^s - 1}{s} - \log k_1 & \textit{if} \quad k_1 \ge \left(\dfrac{\log k_2 - \log k_1}{k_2 - k_1}\right)^{\frac{1}{s-1}} . \end{cases}$$

7.4 α-operator divergence

The concept of α-divergence plays an important role in the information geometry.

Let $(\mathbf{X}, \mathscr{A}, \mu)$ be a measure space, where μ is a finite or a σ-finite measure on $(\mathbf{X}, \mathscr{A})$ and let assume that $\mathbf{P}$ and $\mathbf{Q}$ are two (probability) measures on $(\mathbf{X}, \mathscr{A})$ such that $\mathbf{P} \ll \mu$, $\mathbf{Q} \ll \mu$ are absolutely continuous with respect to a measure μ, e.g. $\mu = \mathbf{P} + \mathbf{Q}$ and that $p = \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mu}$ and $q = \frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mu}$ the (densities) Radon-Nikodym derivative of $\mathbf{P}$ and $\mathbf{Q}$ with respect to μ. Following [5], the basic asymmetric α-divergence is defined as follows: For positive valued measurable functions p and q, and $\alpha \in \mathbb{R}$

$$D_\alpha(p\|q) := \frac{4}{1-\alpha^2}\int\left\{\frac{1-\alpha}{2}p(x) + \frac{1+\alpha}{2}q(x) - p(x)^{\frac{1-\alpha}{2}}q(x)^{\frac{1+\alpha}{2}}\right\}\mathrm{d}\mu(x) \quad (\alpha \ne \pm 1), \tag{7.11}$$

$$D_{-1}(p\|q) \equiv D_1(q\|p) := \int\left\{q(x) - p(x) + p(x)\log\frac{p(x)}{q(x)}\right\}\mathrm{d}\mu(x).$$

If we put $t = \frac{1+\alpha}{2}$ in (7.11), then

$$D_t(p\|q) := \frac{1}{t(1-t)}\int\{(1-t)p(x) + tq(x) - p(x)^{1-t}q(x)^t\}\mathrm{d}\mu(x) \qquad (t \ne 0, 1).$$

From the viewpoint of this, Fujii [53] defined the following operator version of α-divergence in the differential geometry: For positive invertible operators A and B,

$$D_\alpha(A,B) := \frac{1}{\alpha(1-\alpha)}\big(A \nabla_\alpha B - A \#_\alpha B\big) \qquad (0 < \alpha < 1).$$

In particular,

$$\begin{aligned}
D_1(A,B) &:= \lim_{\alpha\uparrow 1} D_\alpha(A,B) = \lim_{\alpha\uparrow 1}\left(\frac{A-B}{\alpha} - \frac{B\#_{1-\alpha}A - B}{\alpha(1-\alpha)}\right)\\
&= A - B - S(B|A)\\
D_0(A,B) &:= \lim_{\alpha\downarrow 0} D_\alpha(A,B) = \lim_{\alpha\downarrow 0}\left(\frac{B-A}{1-\alpha} - \frac{A\#_\alpha B - A}{\alpha(1-\alpha)}\right)\\
&= B - A - S(A|B).
\end{aligned}$$

By definition, α-operator divergence is considered as the difference between the arithmetic and the geometric interpolational paths. In particular, for the case $\alpha = 1/2$, it follows that α-operator divergence coincides with by four times the difference of the geometric mean and the arithmetic mean. For the case of density operators, it coincides with a relative entropy introduced by Beravkin and Staszewski [20] in C^*-algebra setting.

Also we have the following different interpretation of α-operator divergence.

Theorem 7.5 *The α-operator divergence is the difference between two velocity vectors $S_1(A|B)$ and $S_\alpha(A|B)$: For each $\alpha \in (0,1)$*

$$\begin{aligned}
D_\alpha(A,B) &= \frac{1}{1-\alpha}\Big(S_1(A|B) - S_\alpha(A|B)\Big)\\
&= \frac{1}{\alpha}\Big(S_1(B|A) - S_{1-\alpha}(B|A)\Big).
\end{aligned}$$

We investigate estimates of the upper bounds for α-operator divergence:

Theorem 7.6 *Let A and B be positive invertible operators such that $k_1A \le B \le k_2A$ for some scalars $0 < k_1 < k_2$. Then α-operator divergence is positive and for every operator mean ρ and $\alpha \in (0,1)$*

$$(\beta A)\ \rho\ (\overline{\beta} B) \ge D_\alpha(A,B) \ge 0$$

holds for

$$\begin{aligned}
\beta &= \max\left\{\frac{1-\alpha+\alpha k_1 - k_1^\alpha}{\alpha(1-\alpha)}, \frac{1-\alpha+\alpha k_2 - k_2^\alpha}{\alpha(1-\alpha)}\right\}\\
\overline{\beta} &= \max\left\{\frac{\alpha+(1-\alpha)k_2^{-1} - k_2^{\alpha-1}}{\alpha(1-\alpha)}, \frac{\alpha+(1-\alpha)k_1^{-1} - k_1^{\alpha-1}}{\alpha(1-\alpha)}\right\}.
\end{aligned}$$

Proof. Since $A\ \nabla_\alpha\ B \ge A\ \#_\alpha\ B$ $(0 \le \alpha \le 1)$, it follows that α-operator divergence is positive, that is, $D_\alpha(A,B) \ge 0$. On the other hand, it follows from Theorem 7.2 that $\beta A \ge D_\alpha(A,B) \ge 0$. Since $A\ \nabla_\alpha\ B - A\ \#_\alpha\ B = B\ \nabla_{1-\alpha}\ A - B\ \#_{1-\alpha}\ A$ by (7.2), we applied B, A and $1-\alpha$ in Theorem 7.2 to obtain the constant $\overline{\beta} = \beta(0,1,1-\alpha,k_2^{-1},k_1^{-1})$ such that $\overline{\beta}B \ge D_\alpha(A,B) \ge 0$ because $k_2^{-1}B \le A \le k_1^{-1}B$. Therefore we have for every operator mean ρ

$$(\beta A)\ \rho\ (\overline{\beta} B) \ge D_\alpha(A,B)\ \rho\ D_\alpha(A,B) = D_\alpha(A,B) \ge 0.$$

□

If we put $\alpha \to 0,1$ in Theorem 7.6, then we have the following corollary.

Corollary 7.1 *Let A and B be positive invertible operators such that $k_1A \le B \le k_2A$ for some scalars $0 < k_1 < k_2$. Then for every operator mean ρ*

$$(\beta A)\ \rho\ (\overline{\beta}B) \ge D_0(A,B) = S_1(A|B) - S_0(A|B)$$

holds for $\beta = \max\{k_1 - 1 - \log k_1, k_2 - 1 - \log k_2\}$ *and* $\overline{\beta} = \max\{1 - k_2^{-1} - k_2^{-1}\log k_2, 1 - k_1^{-1} - k_1^{-1}\log k_1\}$ *and*

$$(\beta A)\ \rho\ (\overline{\beta}B) \ge D_1(A,B) = S_1(B|A) - S_0(B|A)$$

holds for $\beta = \max\{1 - k_2^{-1} - k_2^{-1}\log k_2, 1 - k_1^{-1} + k_1^{-1}\log k_1\}$ *and* $\overline{\beta} = \max\{k_1 - 1 + \log k_1, k_2 - 1 - \log k_2\}$.

7.5 Notes

For our exposition we have used J.I. Fujii-Mićić-Pečarić-Seo [71], Kamei-J.I. Fujii [67, 68] and J.I. Fujii [53]. Further study may be seen in [55, 56].

Chapter 8

Mercer's Type Inequality

This chapter devotes some properties of Mercer's type inequalities. A variant of Jensen's operator inequality for convex functions, which is a generalization of Mercer's result, is proved. We show a monotonicity property for Mercer's power means for operators and a comparison theorem for quasi-arithmetic means for operators.

8.1 Classical version

Let $a \le x_1 \le x_2 \le \cdots \le x_n \le b$ and let $w_i, 1 \le i \le n$, be nonnegative weights such that $\sum_{i=1}^n w_i = 1$. Then Jensen's inequality asserts:

Theorem 8.1 *If f is convex on $[a,b]$, then*

$$f\left(\sum_{i=1}^n w_i x_i\right) \le \sum_{i=1}^n w_i f(x_i). \tag{8.1}$$

Proof. Refer to [124, Theorem 1.1] for the proof. □

The following theorem is a variant of Jensen's inequality (8.1).

Theorem 8.2 *If f is convex on $[a,b]$, then*

$$f\left(a+b-\sum_{i=1}^{n} w_i x_i\right) \leq f(a)+f(b)-\sum_{i=1}^{n} w_i f(x_i).$$

Proof. If we put $y_i = a+b-x_i$, then $a+b = x_i+y_i$, so that the pairs a,b and x_i,y_i possess the same mid-point. Since there exists $\lambda \in [0,1]$ that

$$x_i = \lambda a+(1-\lambda)b, \quad y_i = (1-\lambda)a+\lambda b \quad \text{for } 1 \leq i \leq n,$$

it follows from (8.1) twice that

$$\begin{aligned} f(y_i) &\leq (1-\lambda)f(a)+\lambda f(b) \\ &= f(a)+f(b)-\big[\lambda f(a)+(1-\lambda)f(b)\big] \\ &\leq f(a)+f(b)-f\big(\lambda a+(1-\lambda)b\big) \\ &= f(a)+f(b)-f(x_i) \end{aligned}$$

and hence we have

$$f(a+b-x_i) \leq f(a)+f(b)-f(x_i) \quad \text{for } 1 \leq i \leq n. \tag{8.2}$$

Therefore it follows that

$$\begin{aligned} f\left(a+b-\sum_{i=1}^{n} w_i x_i\right) &= f\left(\sum_{i=1}^{n} w_i(a+b-x_i)\right) \\ &\leq \sum_{i=1}^{n} w_i f(a+b-x_i) \quad \text{by (8.1)} \\ &\leq \sum_{i=1}^{n} w_i\big[f(a)+f(b)-f(x_i)\big] \quad \text{by (8.2)} \\ &= f(a)+f(b)-\sum_{i=1}^{n} w_i f(x_i). \end{aligned}$$

□

Let A, G and H be the arithmetic, geometric and harmonic means of the positive numbers $0 < x_1 \leq x_2 \leq \cdots \leq x_n$ formed with the positive weights w_i whose sum is unity. Since $(b-t)(t-a)$ is non-negative for $0 < a \leq t \leq b$, division by t gives

$$a+b-t \geq \frac{ab}{t} \qquad \text{(with equality if and only if } t=a \text{ or } t=b).$$

Put $t = x_i$ for $i = 1,2,\ldots,n$. Forming the arithmetic mean on the left and geometric mean on the right derives the following inequality:

$$a+b-A \geq \frac{ab}{G}. \tag{8.3}$$

Making the substitutions $a \to a^{-1}$, $b \to b^{-1}$, $x_i \to x_i^{-1}$ in it and taking inverses extends (8.3) to

$$a+b-A \geq \frac{ab}{G} \geq \left(a^{-1}+b^{-1}-H^{-1}\right)^{-1}.$$

With $r > 0$, we substitute $a \to a^r$, $b \to b^r$, $x_i \to x_i^r$ in this and then raise all three members to the power $\frac{1}{r}$. We get

$$\left(a^r+b^r-\sum_{i=1}^{n} w_i x_i^r\right)^{\frac{1}{r}} > \frac{ab}{G} > \left(a^{-r}+b^{-r}-\sum_{i=1}^{n} w_i x_i^{-1}\right)^{-\frac{1}{r}}.$$

Now introducing the notation

$$Q_r(a,b,x) = \left(a^r+b^r-\sum_{i=1}^{n} w_i x_i^r\right)^{\frac{1}{r}} \qquad \text{for all real } r \neq 0,$$

these last inequalities read

$$Q_r(a,b,x) > Q_0(a,b,x) > Q_{-r}(a,b,x) \qquad \text{for } r > 0, \tag{8.4}$$

where

$$Q_0(a,b,x) = \lim_{r\to 0} Q_r(a,b,x) = \frac{ab}{G}$$

This consideration leads us to formulate the following theorem.

Theorem 8.3 *Let* $+\infty > r > s > -\infty$. *Then*

$$b > Q_r(a,b,x) > Q_s(a,b,x) > a. \tag{8.5}$$

Proof. There are three cases which remain to be considered:
(a) $r > s > 0$, (b) $0 > r > s$, and (c) $r > 0 > s$.
Once these are proved it is a simple matter to verify that

$$\lim_{r\to+\infty} Q_r(a,b,x) = b \qquad \text{and} \qquad \lim_{r\to-\infty} Q_r(a,b,x) = a,$$

giving the upper and lower bounds in the theorem.

The cases (b) and (c) follow easily from (a) and (8.4) above. So let us suppose the truth of case (a) for the moment and dispose of these other cases first.

(a) reads

$$\left(a^r+b^r-\sum_{i=1}^{n} w_i x_i^r\right)^{\frac{1}{r}} > \left(a^s+b^s-\sum_{i=1}^{n} w_i x_i^s\right)^{\frac{1}{s}} \qquad \text{for} \quad r > s > 0.$$

If we make the substitutions $a \to a^{-1}$, $b \to b^{-1}$, $x_i \to x_i^{-1}$ in this and then invert both sides it reads

$$\left(a^{-r}+b^{-r}-\sum_{i=1}^{n} w_i x_i^{-r}\right)^{-\frac{1}{r}} < \left(a^{-s}+b^{-s}-\sum_{i=1}^{n} w_i x_i^{-s}\right)^{-\frac{1}{s}} \qquad \text{for} \quad -r < -s < 0.$$

Writing $r=-p$ and $s=-q$ this reads

$$Q_q(a,b,x) > Q_p(a,b,x) \qquad \text{for} \quad 0>q>p$$

which is case (b).

The case (c) where $r>0>s$ has two subcases namely $|r|>|s|$ and $|s|>|r|$.

The former follows by noting that $Q_r(a,b,x) > Q_{-s}(a,b,x) > Q_s(a,b,x)$ by virtue of (a) and (8.4). The latter follows since $Q_r(a,b,x) > Q_{-r}(a,b,x) > Q_{-s}(a,b,x)$ by virtue of (8.4) and (b). So the cases (b) and (c) have been dealt with.

It now remains to give the proof of case (a). If we put $f(t)=t^{\alpha}$ for $\alpha>1$ in Theorem 8.2, then we have

$$\left(a^{\alpha}+b^{\alpha}-\sum_{i=1}^{n} w_i x_i^{\alpha}\right)^{\frac{1}{\alpha}} > \left(a+b-\sum_{i=1}^{n} w_i x_i\right) \qquad \text{for} \quad \alpha>1.$$

Putting $\alpha=\frac{r}{s}$, making the substitutions $a\to a^s, b\to b^s, x_i\to x_i^s$ and then raising each side to the power $\frac{1}{s}$, we get (a). □

8.2 Operator version

In this section, we show an extension of Theorem 8.2 to self-adjoint operators on a Hilbert space. We use this result to prove a monotonicity property of power means of Mercer's type. Moreover, we consider quasi-arithmetic means in the same way.

First of all, we recall that an operator version of Theorem 8.1 (Jensen's inequality) is true [124, Theorem 1.3]:

Theorem 8.4 *Let $A_1,\dots,A_n \in B(H)$ be self-adjoint operators with $mI_H \le A_j \le MI_H$ for some scalars $m<M$ and let $x_1,\dots x_n \in H$ satisfy $\sum_{i=1}^{n}\|x_i\|^2=1$. If $f\in C([m,M])$ is convex, then*

$$f\left(\sum_{i=1}^{n}\langle A_i x_i, x_i\rangle\right) \le \sum_{i=1}^{n}\langle f(A_i)x_i, x_i\rangle.$$

The following theorem stands for a geometrical property of convexity and is frequently useful.

Theorem 8.5 *Let $A_1,\dots,A_n \in B(H)$ be self-adjoint operators with $mI_H \le A_i \le MI_H$ for some scalars $m<M$ and let $x_1,\dots x_n \in H$ satisfy $\sum_{i=1}^{n}\|x_i\|^2=1$. If $f\in C([m,M])$ is convex, then*

$$\sum_{i=1}^{n}\langle f(A_i)x_i, x_i\rangle \le \frac{M-\sum_{i=1}^{n}\langle A_i x_i, x_i\rangle}{M-m} f(m) + \frac{\sum_{i=1}^{n}\langle A_i x_i, x_i\rangle - m}{M-m} f(M).$$

Proof. Since f is convex on $[m,M]$, we have

$$f(t) \leq \frac{M-t}{M-m} f(m) + \frac{t-m}{M-m} f(M) \qquad \text{for all } t \in [m,M].$$

Since $mI_H \leq A_i \leq MI_H$ for $i = 1,\ldots,n$ and $\sum_{i=1}^n \langle x_i, x_i \rangle = 1$, it follows that $m \leq \sum_{i=1}^n \langle A_i x_i, x_i \rangle \leq M$. Using the functional calculus, we have this theorem. □

The following theorem is an operator vertion of Mercer's inequality.

Theorem 8.6 *Let $A_1,\ldots,A_n \in B(H)$ be self-adjoint operators with $mI_H \leq A_i \leq MI_H$ for some scalars $m < M$ and let $x_1,\ldots x_n \in H$ satisfy $\sum_{i=1}^n \|x_i\|^2 = 1$. If $f \in C([m,M])$ is convex, then we have the following variant of Jensen's inequality*

$$f\left(m+M-\sum_{i=1}^n \langle A_i x_i, x_i \rangle\right) \leq f(m) + f(M) - \sum_{i=1}^n \big\langle f\big(A_i\big) x_i, x_i \big\rangle. \tag{8.6}$$

In fact, to be more specific, we have the following series of inequalities

$$\begin{aligned} & f\left(m+M-\sum_{i=1}^n \langle A_i x_i, x_i \rangle\right) \\ & \leq \sum_{i=1}^n \big\langle f(mI_H + MI_H - A_i) x_i, x_i \big\rangle \\ & \leq \frac{M - \sum_{i=1}^n \langle A_i x_i, x_i \rangle}{M-m} f(M) + \frac{\sum_{i=1}^n \langle A_i x_i, x_i \rangle - m}{M-m} f(m) \\ & \leq f(m) + f(M) - \sum_{i=1}^n \big\langle f(A_i) x_i, x_i \big\rangle. \end{aligned} \tag{8.7}$$

If a function f is concave, then the inequalities (8.6) *and* (8.7) *are reversed.*

Proof. From the conditions $m\langle x_i, x_i \rangle \leq \langle A_i x_i, x_i \rangle \leq M\langle x_i, x_i \rangle$ for all $i = 1,\ldots,n$ and $\sum_{i=1}^n \langle x_i, x_i \rangle = 1$, by summing it follows that $m \leq \sum_{i=1}^n \langle A_i x_i, x_i \rangle \leq M$ and hence, $m \leq m + M - \sum_{i=1}^n \langle A_i x_i, x_i \rangle \leq M$.

Since f is continuous and convex, the same is also true for the function $g : [m,M] \to \mathbb{R}$ defined by $g(t) = f(m+M-t)$, $t \in [m,M]$. By Theorem 8.4,

$$g\left(\sum_{i=1}^n \langle A_i x_i, x_i \rangle\right) \leq \sum_{i=1}^n \big\langle g(A_i) x_i, x_i \big\rangle,$$

i.e.

$$f\left(m+M-\sum_{i=1}^n \langle A_i x_i, x_i \rangle\right) \leq \sum_{i=1}^n \big\langle f(mI_H + MI_H - A_i) x_i, x_i \big\rangle.$$

Applying Theorem 8.5 to g and then to f, we have

$$\begin{aligned}
&\sum_{i=1}^{n}\left\langle f\left(mI_H+MI_H-A_i\right)x_i,x_i\right\rangle \\
&\leq \frac{M-\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle}{M-m}g(m)+\frac{\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle-m}{M-m}g(M) \\
&= \frac{M-\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle}{M-m}f(M)+\frac{\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle-m}{M-m}f(m) \\
&= f(m)+f(M)-\left[\frac{M-\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle}{M-m}f(m)+\frac{\sum_{i=1}^{n}\langle A_ix_i,x_i\rangle-m}{M-m}f(M)\right] \\
&\leq f(m)+f(M)-\sum_{i=1}^{n}\left\langle f(A_i)x_i,x_i\right\rangle.
\end{aligned}$$

The last statement follows immediately from the fact that if f is concave then $-f$ is convex. □

Next, we consider an operator version of power means of Mercer's type.

Let $\mathbf{A}=(A_1,\ldots,A_n)$, where $A_i\in B(H)$ are self-adjoint operators with $mI_H\leq A_i\leq MI_H$ for some scalars $0<m<M$, and $\mathbf{x}=(x_1,\ldots,x_n)$, where $x_i\in H$ satisfy $\sum_{i=1}^{n}\langle x_i,x_i\rangle=1$. We define, for any $r\in\mathbb{R}$

$$\widetilde{M}_r(\mathbf{A},\mathbf{x}):=\begin{cases}\left[m^r+M^r-\sum_{i=1}^{n}\langle A_i^rx_i,x_i\rangle\right]^{\frac{1}{r}}, & r\neq 0,\\ \exp\left(\log m+\log M-\sum_{i=1}^{n}\left\langle(\log A_i)x_i,x_i\right\rangle\right), & r=0\,.\end{cases}$$

Observe that, since $0<m\langle x_i,x_i\rangle\leq\langle A_ix_i,x_i\rangle\leq M\langle x_i,x_i\rangle$ and $\sum_{i=1}^{n}\langle x_i,x_i\rangle=1$, then

- $0<m^r\leq\sum_{i=1}^{n}\langle A_i^rx_i,x_i\rangle\leq M^r$ for all $r>0$,
- $0<M^r\leq\sum_{i=1}^{n}\langle A_i^rx_i,x_i\rangle\leq m^r$ for all $r<0$,
- $\log m\leq\sum_{i=1}^{n}\langle(\log A_i)x_i,x_i\rangle\leq\log M$.

Hence, $\widetilde{M}_r(\mathbf{A},\mathbf{x})$ is well defined.

Furthermore, we define, for any $r,s\in\mathbb{R}$

$$R(r,s,\mathbf{A},\mathbf{x}):=\begin{cases}\left[\sum_{i=1}^{n}\left\langle(m^rI_H+M^rI_H-A_i^r)^{\frac{s}{r}}x_i,x_i\right\rangle\right]^{\frac{1}{s}}, & r\neq 0,s\neq 0,\\ \exp\left(\sum_{i=1}^{n}\left\langle\log(m^rI_H+M^rI_H-A_i^r)^{\frac{1}{r}}x_i,x_i\right\rangle\right), & r\neq 0,s=0,\\ \left[\sum_{i=1}^{n}\left\langle\exp\left(s(\log m)I_H+(\log M)I_H-\log A_i\right)x_i,x_i\right\rangle\right]^{\frac{1}{s}}, & r=0,s\neq 0,\end{cases}$$

$$S(r,s,\mathbf{A},\mathbf{x}) := \begin{cases} \left[\dfrac{M^r - S_r}{M^r - m^r} \cdot M^s + \dfrac{S_r - m^r}{M^r - m^r} \cdot m^s\right]^{\frac{1}{s}}, & r \neq 0, s \neq 0, \\ \exp\Big(\dfrac{M^r - S_r}{M^r - m^r} \cdot \log M + \dfrac{S_r - m^r}{M^r - m^r} \cdot \log m\Big), & r \neq 0, s = 0, \\ \left[\dfrac{(\log M) - S_0}{\log M - \log m} \cdot M^s + \dfrac{S_0 - (\log m)}{\log M - \log m} \cdot m^s\right]^{\frac{1}{s}}, & r = 0, s \neq 0, \end{cases}$$

where $S_r = \sum_{i=1}^n \langle A_i^r x_i, x_i\rangle$ and $S_0 = \sum_{i=1}^n \langle (\log A_i) x_i, x_i\rangle$. It is easy to see that $R(r,s,\mathbf{A},\mathbf{x})$ and $S(r,s,\mathbf{A},\mathbf{x})$ are also well defined.

Theorem 8.7 *If $r,s \in \mathbb{R}$, $r < s$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{x}) \leq \widetilde{M}_s(\mathbf{A},\mathbf{x}).$$

Furthermore,

$$\widetilde{M}_r(\mathbf{A},\mathbf{x}) \leq R(r,s,\mathbf{A},\mathbf{x}) \leq S(r,s,\mathbf{A},\mathbf{x}) \leq \widetilde{M}_s(\mathbf{A},\mathbf{x}). \tag{8.8}$$

Proof. STEP 1: Assume $0 < r < s$.
In this case we have $0 < m^r I_H \leq A_i^r \leq M^r I_H$ $(i = 1,\dots,n)$. Applying Theorem 8.6 to the continuous convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} > 1$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we have

$$\begin{aligned} &\left[m^r + M^r - \sum_{i=1}^n \langle A_i^r x_i, x_i\rangle\right]^{\frac{s}{r}} \\ &\leq \sum_{i=1}^n \Big\langle (m^r I_H + M^r I_H - A_i^r)^{\frac{s}{r}} x_i, x_i\Big\rangle \\ &\leq \frac{M^r - \sum_{i=1}^n \langle A_i^r x_i, x_i\rangle}{M^r - m^r} M^s + \frac{\sum_{i=1}^n \langle A_i^r x_i, x_i\rangle - m^r}{M^r - m^r} m^s \\ &\leq m^s + M^s - \sum_{i=1}^n \langle A_i^s x_i, x_i\rangle, \end{aligned}$$

or

$$\Big[\widetilde{M}_r(\mathbf{A},\mathbf{x})\Big]^s \leq \Big[R(r,s,\mathbf{A},\mathbf{x})\Big]^s \leq \Big[S(r,s,\mathbf{A},\mathbf{x})\Big]^s \leq \Big[\widetilde{M}_s(\mathbf{A},\mathbf{x})\Big]^s.$$

Since $s > 0$, this gives (8.8).

STEP 2: Assume $r < s < 0$.
In this case we have $0 < M^r I_H \leq A_i^r \leq m^r I_H$ $(i = 1,\dots,n)$. Applying Theorem 8.6 to the continuous concave function $f(t) = t^{\frac{s}{r}}$ (note that $0 < \frac{s}{r} < 1$ here) and replacing A_i, m and M with A_i^r, M^r and m^r, respectively, we have

$$\begin{aligned}
&\left[M^r + m^r - \sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle\right]^{\frac{s}{r}} \\
&\geq \sum_{i=1}^{n} \left\langle (M^r I_H + m^r I_H - A_i^r)^{\frac{s}{r}} x_i, x_i \right\rangle \\
&\geq \frac{m^r - \sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle}{m^r - M^r} m^s + \frac{\sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle - M^r}{m^r - M^r} M^s \\
&\geq M^s + m^s - \sum_{i=1}^{n} \langle A_i^s x_i, x_i\rangle
\end{aligned}$$

or

$$\left[\widetilde{M}_r(\mathbf{A},\mathbf{x})\right]^s \geq \left[R(r,s,\mathbf{A},\mathbf{x})\right]^s \geq \left[S(r,s,\mathbf{A},\mathbf{x})\right]^s \geq \left[\widetilde{M}_s(\mathbf{A},\mathbf{x})\right]^s.$$

Since $s < 0$, this gives (8.8).

STEP 3: Assume $r < 0 < s$.

In this case we have $0 < M^r I_H \leq A_i^r \leq m^r I_H \ (i = 1, \ldots, n)$. Applying Theorem 8.6 to the continuous convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} < 0$ here) and proceeding in the same way as in STEP 1, we obtain (8.8).

STEP 4: Assume $r < 0, s = 0$.

In this case we have $0 < M^r I_H \leq A_i^r \leq m^r I_H \ (i = 1, \ldots, n)$. Applying Theorem 8.6 to the continuous convex function $f(t) = \frac{1}{r} \log t$ (note that $\frac{1}{r} < 0$ here) and replacing A_i, m and M with A_i^r, M^r and m^r, respectively, we have

$$\begin{aligned}
&\frac{1}{r} \log \left(M^r + m^r - \sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle\right) \\
&\leq \sum_{i=1}^{n} \left\langle \frac{1}{r} \log (M^r I_H + m^r I_H - A_i^r) x_i, x_i \right\rangle \\
&\leq \frac{m^r - \sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle}{m^r - M^r} \cdot \log m + \frac{\sum_{i=1}^{n} \langle A_i^r x_i, x_i\rangle - M^r}{m^r - M^r} \cdot \log M \\
&\leq \log M + \log m - \sum_{i=1}^{n} \langle (\log A_i) x_i, x_i\rangle
\end{aligned}$$

or

$$\log \widetilde{M}_r(\mathbf{A},\mathbf{x}) \leq \log R(r,0,\mathbf{A},\mathbf{x}) \leq \log S(r,0,\mathbf{A},\mathbf{x}) \leq \log \widetilde{M}_0(\mathbf{A},\mathbf{x}).$$

This gives (8.8) for $s = 0$.

STEP 5: Assume $r = 0, s > 0$.

We have $(\log m) I_H \leq \log A_i \leq (\log M) I_H \ (i = 1, \ldots, n)$. Applying Theorem 8.6 to the continuous convex function $f(t) = \exp(st)$ and replacing A_i, m and M with $\log A_i$, $\log m$ and $\log M$, respectively, we have

$$\begin{aligned}
&\exp\left(s\Big(\log m+\log M-\sum_{i=1}^{n}\langle(\log A_i)x_i,x_i\rangle\Big)\right)\\
&\leq \sum_{i=1}^{n}\langle\exp\left(s\left((\log m)I_H+(\log M)I_H-\log A_i\right)\right)x_i,x_i\rangle\\
&\leq \frac{\log M-\sum_{i=1}^{n}\langle(\log A_i)x_i,x_i\rangle}{\log M-\log m}M^s+\frac{\sum_{i=1}^{n}\langle(\log A_i)x_i,x_i\rangle-\log m}{\log M-\log m}m^s\\
&\leq m^s+M^s-\sum_{i=1}^{n}\langle A_i^s x_i,x_i\rangle
\end{aligned}$$

or

$$\left[\widetilde{M}_0(\mathbf{A},\mathbf{x})\right]^s \leq \left[R(0,s,\mathbf{A},\mathbf{x})\right]^s \leq \left[S(0,s,\mathbf{A},\mathbf{x})\right]^s \leq \left[\widetilde{M}_s(\mathbf{A},\mathbf{x})\right]^s.$$

Since $s>0$, this gives (8.8) for $r=0$. □

Next, we consider quasi-arithmetic means of Mercer's type.

Let $\mathbf{A}=(A_1,\dots,A_n)$, where $A_i\in B(H)$ are self-adjoint operators with $mI_H\leq A_i\leq MI_H$ for some scalars $m<M$, and $\mathbf{x}=(x_1,\dots,x_n)$, where $x_i\in H$ satisfy $\sum_{i=1}^{n}\langle x_i,x_i\rangle=1$. Let $\varphi,\psi\in C([m,M])$ be strictly monotonic functions on an interval $[m,M]$. We define

$$\widetilde{M}_\varphi(\mathbf{A},\mathbf{x}):=\varphi^{-1}\Big(\varphi(m)+\varphi(M)-\sum_{i=1}^{n}\big\langle\varphi(A_i)x_i,x_i\big\rangle\Big).$$

Observe that, since $mI_H\leq A_i\leq MI_H$ and $\sum_{i=1}^{n}\langle x_i,x_i\rangle=1$, then

- $\varphi(m)\leq\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle\leq\varphi(M)$ if φ is increasing,
- $\varphi(M)\leq\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle\leq\varphi(m)$ if φ is decreasing.

Hence, $\widetilde{M}_\varphi(\mathbf{A},\mathbf{x})$ is well defined.

Theorem 8.8 *Under the above hypotheses,*

(i) *if either $\psi\circ\varphi^{-1}$ is convex and ψ is strictly increasing, or $\psi\circ\varphi^{-1}$ is concave and ψ is strictly decreasing, then*

$$\widetilde{M}_\varphi(\mathbf{A},\mathbf{x})\leq\widetilde{M}_\psi(\mathbf{A},\mathbf{x}). \tag{8.9}$$

In fact, to be more specific, we have the following series of inequalities

$$\begin{aligned}
&\widetilde{M}_\varphi(\mathbf{A},\mathbf{x})\\
&\leq\psi^{-1}\left(\sum_{i=1}^{n}\Big\langle\big(\psi\varphi^{-1}\big)\left(\varphi(m)I_H+\varphi(M)I_H-\varphi(A_i)\right)x_i,x_i\Big\rangle\right) \qquad (8.10)\\
&\leq\psi^{-1}\left(\frac{\varphi(M)-\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle}{\varphi(M)-\varphi(m)}\psi(M)+\frac{\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle-\varphi(m)}{\varphi(M)-\varphi(m)}\psi(m)\right)\\
&\leq\widetilde{M}_\psi(\mathbf{A},\mathbf{x}),
\end{aligned}$$

(ii) *if either* $\psi\circ\varphi^{-1}$ *is concave and* ψ *is strictly increasing, or* $\psi\circ\varphi^{-1}$ *is convex and* ψ *is strictly decreasing, then the reverse inequalities of* (8.9) *and* (8.10) *hold.*

Proof. Suppose that $\psi\circ\varphi^{-1}$ is convex. If in Theorem 8.6 we let $f=\psi\circ\varphi^{-1}$ and replace A_i, m and M with $\varphi(A_i)$, $\varphi(m)$ and $\varphi(M)$, respectively, then we obtain

$$\begin{aligned}
&\left(\psi\circ\varphi^{-1}\right)\left(\varphi(m)+\varphi(M)-\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle\right)\\
&\le\sum_{i=1}^{n}\left\langle\left(\psi\circ\varphi^{-1}\right)\left(\varphi(m)I_H+\varphi(M)I_H-\varphi(A_i)\right)x_i,x_i\right\rangle \qquad (8.11)\\
&\le\frac{\varphi(M)-\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle}{\varphi(M)-\varphi(m)}\psi(M)+\frac{\sum_{i=1}^{n}\langle\varphi(A_i)x_i,x_i\rangle-\varphi(m)}{\varphi(M)-\varphi(m)}\psi(m)\\
&\le\psi(m)+\psi(M)-\sum_{i=1}^{n}\langle\psi(A_i)x_i,x_i\rangle .
\end{aligned}$$

If $\psi\circ\varphi^{-1}$ is concave then we obtain the reverse of inequalities (8.11).

If ψ is strictly increasing, then the inverse function ψ^{-1} is also strictly increasing, so that (8.11) implies (8.10). If ψ is strictly decreasing, then the inverse function ψ^{-1} is also strictly decreasing, so that in this case the reverse of (8.11) implies (8.10). Analogously, we get the reverse of (8.10) in the cases when $\psi\circ\varphi^{-1}$ is convex and ψ is strictly decreasing, or $\psi\circ\varphi^{-1}$ is concave and ψ is strictly increasing. □

8.3 Operator version with mappings

Assume that $(\Phi_1,\dots,\Phi_n)$ is an $n-$tuple of positive linear mappings $\Phi_i : B(H)\to B(K)$, $i=1,\dots,n$. If $\sum_{i=1}^{n}\Phi_i(I_H)=I_K$, we say that $(\Phi_1,\dots,\Phi_n)$ is *unital*.

We have the following generalization of discrete Jensen's operator inequality.

Theorem 8.9 *Let* $(A_1,\dots,A_n)$ *be an* $n-$*tuple of self-adjoint operators in* $B(H)$ *with spectra in* $[m,M]$ *for some scalars* $m<M$, *and let* $(\Phi_1,\dots,\Phi_n)$ *be a unital* $n-$*tuple positive linear mappings* $\Phi_i : B(H)\to B(K)$, $i=1,\dots,n$. *If* f *is an operator convex function on* $[m,M]$, *then*

$$f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right)\le\sum_{i=1}^{n}\Phi_i(f(A_i)). \qquad (8.12)$$

Proof. Using continuity of f, Φ_i and uniform approximation of self-adjoint operators by simple operators using decomposition of unit we can assume that $A_i=\sum_{j\in I_i}t_{i,j}e_{i,j}$ where

I_i are finite sets and $\{e_{i,j}\}_{j\in I_i}$ are decompositions of unit, $i = 1,\ldots,n$. We have

$$\begin{aligned} f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right) &= f\left(\sum_{i=1}^{n}\Phi_i(\sum_{j\in I_i}t_{i,j}e_{i,j})\right) = f\left(\sum_{i=1}^{n}\sum_{j\in I_i}t_{i,j}\Phi_i(e_{i,j})\right) \\ &= f\left(\sum_{i=1}^{n}\sum_{j\in I_i}\sqrt{\Phi_i(e_{i,j})}t_{i,j}\sqrt{\Phi_i(e_{i,j})}\right) \\ &\le \sum_{i=1}^{n}\sum_{j\in I_i}\sqrt{\Phi_i(e_{i,j})}f(t_{i,j})\sqrt{\Phi_i(e_{i,j})} \\ &= \sum_{i=1}^{n}\sum_{j\in I_i}f(t_{i,j})\Phi_i(e_{i,j}) = \sum_{i=1}^{n}\Phi_i(f(A_i)). \end{aligned}$$

The second proof: We use the idea from [81] (also compare to [221]). If f is operator convex in $I = [0,1]$ and $f(0) \le 0$, we can suppose, with no loss of generality, that it is non-positive. Then there is a connection σ such that $-f(t) = t\,\sigma\,(1-t)$. We use the following properties of a connection σ:

(i) $\Phi(A\,\sigma\,B) \le \Phi(A)\,\sigma\,\Phi(B)$ for a positive linear mapping Φ and positive operators A and B ([15]).

(ii) (subadditivity) $\sum_{i=1}^{n}A_i\,\sigma\,B_i \le (\sum_{i=1}^{n}A_i)\ \sigma\ (\sum_{i=1}^{n}B_i)$ for positive $n-$tuples $(A_1,\ldots,A_n)$ and $(B_1,\ldots,B_n)$ ([81]).

We obtain

$$\begin{aligned} -\sum_{i=1}^{n}\Phi_i(f(A_i)) &= \sum_{i=1}^{n}\Phi_i\left(A_i\,\sigma\,(I_H - A_i)\right) \\ &\le \sum_{i=1}^{n}\Phi_i(A_i)\,\sigma\,\Phi_i(I_H - A_i) \le \left(\sum_{i=1}^{n}\Phi_i(A_i)\right)\ \sigma\ \left(\sum_{i=1}^{n}\Phi_i(I_H - A_i)\right) \\ &= \left(\sum_{i=1}^{n}\Phi_i(A_i)\right)\ \sigma\ \left(I_K - \sum_{i=1}^{n}\Phi_i(A_i)\right) = -f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right). \end{aligned}$$

Consider now an arbitrary operator convex function f defined on $[0,1]$. The function $\tilde{f}(x) = f(x) - f(0)$ satisfies the previous conditions, so (8.12) becomes

$$f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right) \le \sum_{i=1}^{n}\Phi_i(f(A_i)) + f(0)\left(I_K - \sum_{i=1}^{n}\Phi_i(I_H)\right). \tag{8.13}$$

By setting $g(x) = f\left((\beta-\alpha)x+\alpha\right)$ one may reduce the statement for operator convex functions defined on an arbitrary interval $[\alpha,\beta]$ to operator convex functions defined on the interval $[0,1]$. □

We show a variant of Jensen's operator inequality which is an extension of Theorem 8.2 and Theorem 8.6 to self-adjoint operators and positive linear mappings.

Theorem 8.10 *Let $(A_1,\dots,A_n)$ be $n-$tuple of self-adjoint operators in $B(H)$ with spectra in $[m,M]$ for some scalars $m<M$, and let $(\Phi_1,\dots,\Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i : B(H)\to B(K)$, $i=1,\dots,n$. If $f\in C([m,M])$ is convex on $[m,M]$, then*

$$f\left(mI_K+MI_K-\sum_{i=1}^{n}\Phi_i(A_i)\right)\le f(m)I_K+f(M)I_K-\sum_{i=1}^{n}\Phi_i(f(A_i)). \tag{8.14}$$

In fact, to be more specific, the following series of inequalities holds

$$\begin{aligned} &f\left(mI_K+MI_K-\sum_{i=1}^{n}\Phi_i(A_i)\right)\\ &\le \frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}f(M)+\frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}f(m)\\ &\le f(m)I_K+f(M)I_K-\sum_{i=1}^{n}\Phi_i(f(A_i)). \end{aligned}\tag{8.15}$$

If a function f is concave, then inequalities (8.14) *and* (8.15) *are reversed.*

Proof. Since f is continuous and convex, the same is also true for the function $g:[m,M]\to\mathbb{R}$ defined by $g(t)=f(m+M-t)$, $t\in[m,M]$. Hence, the following inequalities

$$f(t)\le\frac{t-m}{M-m}f(M)+\frac{M-t}{M-m}f(m)\quad\text{and}\quad g(t)\le\frac{t-m}{M-m}g(M)+\frac{M-t}{M-m}g(m)$$

hold for every $t\in[m,M]$ (see e.g. [249, p. 2]).

Since $mI_H\le A_i\le MI_H$ for $i=1,\dots,n$ and $\sum_{i=1}^{n}\Phi_i(I_H)=I_K$, it follows that $mI_K\le\sum_{i=1}^{n}\Phi_i(A_i)\le MI_K$. Now, using the functional calculus we have

$$g\left(\sum_{i=1}^{n}\Phi_i(A_i)\right)\le\frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}g(M)+\frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}g(m)$$

or

$$\begin{aligned} &f\left(mI_K+MI_K-\sum_{i=1}^{n}\Phi_i(A_i)\right)\\ &\le\frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}f(m)+\frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}f(M)\\ &=f(m)I_K+f(M)I_K-\left[\frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}f(m)+\frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}f(M)\right]. \end{aligned}\tag{8.16}$$

On the other hand, we also have

$$f(A_i)\le\frac{A_i-mI_H}{M-m}f(M)+\frac{MI_H-A_i}{M-m}f(m).$$

Applying positive linear mappings Φ_i and summing, it follows that

$$\sum_{i=1}^{n}\Phi_i(f(A_i)) \leq \frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}f(M)+\frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}f(m). \tag{8.17}$$

Using inequalities (8.16) and (8.17), we obtain desired inequalities (8.14) and (8.15).

The last statement follows immediately from the fact that if φ is concave then $-\varphi$ is convex. □

We consider Mercer's power means for positive linear mappings.

Let $\mathbf{A} = (A_1,\ldots,A_n)$ be an $n-$tuple of positive invertible operators in $B(H)$ with $Sp(A_i) \subseteq [m,M]$ for some scalars $0 < m < M$, and let $\mathbf{\Phi} = (\Phi_1,\ldots,\Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i : B(H) \to B(K)$, $i = 1,\ldots,n$. We define, for any $r \in \mathbb{R}$

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) := \begin{cases} \left[m^r I_K + M^r I_K - \sum_{i=1}^{n}\Phi_i(A_i^r)\right]^{\frac{1}{r}}, & r \neq 0, \\ \exp\Big((\log m)I_K + (\log M)I_K - \sum_{i=1}^{n}\Phi_i\big(\log(A_i)\big)\Big), & r = 0. \end{cases} \tag{8.18}$$

Observe that, since $0 < mI_H \leq A_i \leq MI_H$ and $\sum_{i=1}^{n}\Phi_i(I_H) = I_K$, then

- $0 < m^r I_K \leq \sum_{i=1}^{n}\Phi_i(A_i^r) \leq M^r I_K$ for all $r > 0$,
- $0 < M^r I_K \leq \sum_{i=1}^{n}\Phi_i(A_i^r) \leq m^r I_K$ for all $r < 0$,
- $(\log m)I_K \leq \sum_{i=1}^{n}\Phi_i(\log(A_i)) \leq (\log M)I_K$.

Hence, $\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})$ is well defined.

Furthermore, we define a constant $\Delta(m,M,p)$ for $0 < m < M$ and $p \in \mathbb{R}$ as follows:

$$\Delta(m,M,p) := \begin{cases} K\left(m^p, M^p, \frac{1}{p}\right) = \frac{p(m^pM-M^pm)}{(1-p)(M^p-m^p)}\left(\frac{(1-p)(M-m)}{m^pM-M^pm}\right)^{\frac{1}{p}}, & p \neq 0, \\ S\left(\frac{M}{m}\right) = \frac{M-m}{\log M-\log m}\exp\left(\frac{m(1+\log M)-M(1+\log m)}{M-m}\right), & p = 0. \end{cases}$$

We remark that $\Delta(m,M,0) = \lim_{p\to 0}\Delta(m,M,p)$ by using Theorem 2.17.

We show a monotonicity property of Mercer's power means for positive linear mappings and investigate a complementary domain to Mercer's power means.

Theorem 8.11 *Let $r,s \in \mathbb{R}$, $r < s$.*
(i) *If either $r \leq -1$ or $s \geq 1$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \leq \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.19}$$

(ii) *If $-1 < r$ and $s < 1$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \leq \Delta(m,M,s)\cdot\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.20}$$

Proof. (i) STEP 1: Suppose that $0 < r < s$ and $s \geq 1$.

Applying the inequality (8.14) to the convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} > 1$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we have

$$\left[m^r I_K + M^r I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right]^{\frac{s}{r}} \leq m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right). \tag{8.21}$$

Raising both sides to the power $\frac{1}{s}$ $\left(0 < \frac{1}{s} \leq 1\right)$, it follows from the Löwner-Heinz theorem (Theorem 3.1) that (8.19) holds.

STEP 2: Suppose that $r < 0$ and $s \geq 1$.

Applying the inequality (8.14) to the convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} < 0$ here) and proceeding in the same way as in STEP 1, we have that (8.19) holds.

STEP 3: Suppose that $r = 0$ and $s \geq 1$.

Applying the inequality (8.14) to the convex function $f(t) = \exp(s \cdot t)$ and replacing A_i, m and M with $\log(A_i)$, $\log m$ and $\log M$, respectively, we have

$$\begin{aligned}
&\exp\left(s\Big((\log m)\, I_K + (\log M)\, I_K - \sum_{i=1}^{n} \Phi_i\big(\log(A_i)\big)\Big)\right) \\
&\leq \exp(s \log m)\, I_K + \exp(s \log M)\, I_K - \sum_{i=1}^{n} \Phi_i\Big(\exp\big(s \log(A_i)\big)\Big) \\
&= m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right)
\end{aligned} \tag{8.22}$$

or

$$\left[\widetilde{M}_0(\mathbf{A}, \mathbf{\Phi})\right]^s \leq \left[\widetilde{M}_s(\mathbf{A}, \mathbf{\Phi})\right]^s.$$

Raising both sides to the power $\frac{1}{s}$ $\left(0 < \frac{1}{s} \leq 1\right)$, it follows from the Löwner-Heinz theorem that (8.19) holds for $r = 0$.

STEP 4: Suppose that $r \leq -1$ and $s > r$.

The inequality (8.19) follows from the above cases replacing A_i, r and s by A_i^{-1}, $-s$ and $-r$, respectively, and using the equality $\widetilde{M}_{-s}\left(\mathbf{A}^{-1}, \mathbf{\Phi}\right) = \widetilde{M}_s(\mathbf{A}, \mathbf{\Phi})^{-1}$, where $\mathbf{A}^{-1} = \left(A_1^{-1}, \ldots, A_n^{-1}\right)$.

(ii) STEP 1: Suppose that $0 < r < s < 1$.

In the same way as in (i) STEP 1 we obtain inequality (8.21). Observe that, since $m^s I_K \leq \sum_{i=1}^{n} \Phi_i\left(A_i^s\right) \leq M^s I_K$, it follows that $m^s I_K \leq m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right) \leq M^s I_K$. Raising both sides of (8.21) to the power $\frac{1}{s}$ $\left(\frac{1}{s} > 1\right)$, it follows from Theorem 4.3 (i) that

$$\widetilde{M}_r(\mathbf{A}, \mathbf{\Phi}) \leq K\left(m^s, M^s, \frac{1}{s}\right) \widetilde{M}_s(\mathbf{A}, \mathbf{\Phi}).$$

STEP 2: Suppose that $0 = r < s < 1$.

In the same way as in (i) STEP 3 we obtain inequality (8.22). With the same observation as in (ii) STEP 1 and raising both sides of (8.22) to the power $\frac{1}{s}$ $\left(\frac{1}{s} > 1\right)$, it follows from Theorem 4.3 (i) that

$$\widetilde{M}_0(\mathbf{A},\mathbf{\Phi}) \le K\left(m^s, M^s, \frac{1}{s}\right)\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).$$

STEP 3: Suppose that $-1<r<s<0$.

The proof follows from (ii) STEP 1 replacing A_i, r and s by A_i^{-1}, $-s$ and $-r$, respectively, and using the equality $K(M,m,p)=K(m,M,p)$ (see [96, p. 77]).

STEP 4: Suppose that $-1<r<s=0$.

Applying the inequality (8.14) to the convex function $f(t)=\frac{1}{r}\log t$ and replacing A_i, m and M with A_i^r, M^r and m^r, respectively, we obtain

$$\frac{1}{r}\log\left(m^r I_K + M^r I_K - \sum_{i=1}^n \Phi_i(A_i^r)\right) \le (\log m)I_K + (\log M)I_K - \sum_{i=1}^n \Phi_i(\log(A_i).$$

Observing that both sides have spectra in $[\log m, \log M]$, it follows from Theorem 4.7 that (8.20) holds for $s=0$.

STEP 5: Suppose that $-1<r<0<s<1$.

In the same way as in (i) STEP 2 we obtain inequality (8.21). With the same observation as in (ii) STEP 1 it follows from Theorem 4.3 (i) that

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \le K\left(m^s, M^s, \frac{1}{s}\right)\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).$$

□

Furthermore, we define $S(r,s,\mathbf{A},\mathbf{\Phi})$ for $\mathbf{A}$, $\mathbf{\Phi}$ as in (8.18) and $r,s\in\mathbb{R}$ as follows:

$$S(r,s,\mathbf{A},\mathbf{\Phi}) := \begin{cases} \left[\dfrac{M^r I_K - S_r}{M^r-m^r}M^s + \dfrac{S_r - m^r I_K}{M^r-m^r}m^s\right]^{\frac{1}{s}}, & r\neq 0, s\neq 0,\\ \exp\left(\dfrac{M^r I_K - S_r}{M^r-m^r}\log M + \dfrac{S_r - m^r I_K}{M^r-m^r}\log m\right), & r\neq 0, s=0,\\ \left[\dfrac{(\log M)I_K - S_0}{\log M-\log m}M^s + \dfrac{S_0-(\log m)I_K}{\log M-\log m}m^s\right]^{\frac{1}{s}}, & r=0, s\neq 0, \end{cases} \tag{8.23}$$

where $S_r=\sum_{i=1}^n \Phi_i(A_i^r)$ and $S_0=\sum_{i=1}^n \Phi_i(\log(A_i))$. It is easy to see that $S(r,s,\mathbf{A},\mathbf{\Phi})$ is well defined.

If we use inequalities (8.15) instead of the inequality (8.14), then we have the following results.

Theorem 8.12 *Let $r,s\in\mathbb{R}$, $r<s$.*

(i) *If $s\ge 1$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \le S(r,s,\mathbf{A},\mathbf{\Phi}) \le \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.24}$$

If $r\le -1$, then

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \le S(s,r,\mathbf{A},\mathbf{\Phi}) \le \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.25}$$

(ii) *If $-1<r$ and $s<1$, then*

$$\frac{1}{\Delta(m,M,s)}\cdot\widetilde{M}_r(\mathbf{A},\Phi) \le S(r,s,\mathbf{A},\mathbf{\Phi}) \le \Delta(m,M,s)\cdot\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.26}$$

Proof. (i) STEP 1: Suppose that $0 < r < s$ and $s \geq 1$.

Applying inequalities (8.15) to the convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} \geq 1$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we have

$$\left[m^r I_K + M^r I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right]^{\frac{s}{r}} \leq \frac{M^r I_K - S_r}{M^r - m^r} M^s + \frac{S_r - m^r I_K}{M^r - m^r} m^s$$
$$\leq m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right). \tag{8.27}$$

Raising these inequalities to the power $\frac{1}{s}$ $\left(0 < \frac{1}{s} \leq 1\right)$, it follows from the Löwner-Heinz theorem that the desired inequality (8.24) holds.

STEP 2: Suppose that $r < 0$ and $s \geq 1$.

Applying inequalities (8.15) to the convex function $f(t) = t^{\frac{s}{r}}$ (note that $\frac{s}{r} < 0$ here) and proceeding in the same way as in STEP 1, we obtain the desired inequality (8.24).

STEP 3: Suppose that $r = 0$ and $s \geq 1$.

Applying inequalities (8.15) to the convex function $f(t) = \exp(s \cdot t)$ and replacing A_i, m and M with $\log A_i$, $\log m$ and $\log M$, respectively, we have

$$\exp\left(s\Big((\log m) I_K + (\log M) I_K - \sum_{i=1}^{n} \Phi_i\big(\log(A_i)\big)\Big)\right)$$
$$\leq \frac{(\log M) I_K - S_0}{\log M - \log m} \cdot \exp(s \log M) + \frac{S_0 - (\log m) I_K}{\log M - \log m} \cdot \exp(s \log m)$$
$$\leq \exp(s \log m) I_K + \exp(s \log M) I_K - \sum_{i=1}^{n} \Phi_i\Big(\exp\big(s \log(A_i)\big)\Big)$$
$$= m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right) \tag{8.28}$$

or

$$\left[\widetilde{M}_0(\mathbf{A}, \mathbf{\Phi})\right]^s \leq [S(0, s, \mathbf{A}, \mathbf{\Phi})]^s \leq \left[\widetilde{M}_s(\mathbf{A}, \mathbf{\Phi})\right]^s.$$

Raising these inequalities to the power $\frac{1}{s}$ $\left(0 < \frac{1}{s} \leq 1\right)$, it follows from the Löwner-Heinz theorem that (8.24) holds for $r = 0$.

STEP 4: Suppose that $r \leq -1$ and $s > r$.

The proof of (8.25) follows using the same way as in the above cases.

(ii) STEP 1: Suppose that $0 < r < s < 1$.

In the same way as in (i) STEP 1 we obtain inequalities (8.27). Observe that, since $m^r I_K \leq \sum_{i=1}^{n} \Phi_i\left(A_i^r\right) \leq M^r I_K$ and $m^s I_K \leq \sum_{i=1}^{n} \Phi_i\left(A_i^s\right) \leq M^s I_K$, it follows that $m^s I_K \leq \left[m^r I_K + M^r I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right]^{\frac{s}{r}} \leq M^s I_K$ and $m^s I_K \leq m^s I_K + M^s I_K - \sum_{i=1}^{n} \Phi_i\left(A_i^s\right) \leq M^s I_K$.

Raising inequalities (8.27) to the power $\frac{1}{s}$ $\left(\frac{1}{s}>1\right)$, it follows from Theorem 4.3 (i) that

$$\begin{aligned}
&K\left(m^s,M^s,\frac{1}{s}\right)^{-1}\left[m^rI_K+M^rI_K-\sum_{i=1}^n\Phi_i\left(A_i^r\right)\right]^{\frac{1}{r}}\\
&\le\left[\frac{M^rI_K-S_r}{M^r-m^r}\cdot M^s+\frac{S_r-m^rI_K}{M^r-m^r}m^s\right]^{\frac{1}{s}}\\
&\le K\left(m^s,M^s,\frac{1}{s}\right)\left[m^sI_K+M^sI_K-\sum_{i=1}^n\Phi_i\left(A_i^s\right)\right]^{\frac{1}{s}},
\end{aligned}$$

which gives the desired inequality (8.26).

STEP 2: Suppose that $0=r<s<1$.

In the same way as in (i) STEP 3 we obtain inequalities (8.28). Observe that, since $(\log m)I_K\le(\log m)I_K+(\log M)I_K-\sum_{i=1}^n\Phi_i(\log(A_i))\le(\log M)I_K$ and $m^sI_K\le\sum_{i=1}^n\Phi_i\left(A_i^s\right)$ $\le M^sI_K$, it follows that

$$m^sI_K\le\exp\left(s\left((\log m)I_K+(\log M)I_K-\sum_{i=1}^n\Phi_i(\log(A_i))\right)\right)\le M^sI_K$$

and $m^sI_K\le m^sI_K+M^sI_K-\sum_{i=1}^n\Phi_i\left(A_i^s\right)\le M^sI_K$. Raising inequalities (8.28) to the power $\frac{1}{s}$ $\left(\frac{1}{s}>1\right)$, it follows from Theorem 4.3 (i) that (8.26) holds for $r=0$.

STEP 3: Suppose that $-1<r<s<0$.

Applying reversed inequalities (8.15) to the concave function $f(t)=t^{\frac{s}{r}}$ (note that $0<\frac{s}{r}<1$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we obtain reversed (8.27). With the same observation as in STEP 1 it follows that (8.26) holds.

STEP 4: Suppose that $-1<r<s=0$.

Applying inequalities (8.15) to the convex function $f(t)=\frac{1}{r}\log t$ (note that $\frac{1}{r}<0$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we obtain

$$\begin{aligned}
&\frac{1}{r}\log\left(m^rI_K+M^rI_K-\sum_{i=1}^n\Phi_i\left(A_i^r\right)\right)\\
&\le\frac{M^rI_K-S_r}{M^r-m^r}\cdot\log M+\frac{S_r-m^r}{M^r-m^r}\cdot\log m\\
&\le(\log m)I_K+(\log M)I_K-\sum_{i=1}^n\Phi_i(\log(A_i)).
\end{aligned}$$

Now, it follows from Theorem 4.7 that

$$S\left(e^{\log M-\log m}\right)^{-1}\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\le S(r,0,\mathbf{A},\mathbf{\Phi})\le S\left(e^{\log M-\log m}\right)\widetilde{M}_0(\mathbf{A},\mathbf{\Phi}),$$

which gives (8.26) holds for $s=0$.

STEP 5: Suppose that $-1<r<0<s<1$.

Applying inequalities (8.15) to the convex function $f(t)=t^{\frac{s}{r}}$ (note that $\frac{s}{r}<0$ here) and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we obtain inequalities (8.27). Proceeding in the same way as in STEP 1, we obtain (8.26). □

Remark 8.1 *Since obviously $S(r,r,\mathbf{A},\mathbf{\Phi}) = \widetilde{M}_r(\mathbf{A},\mathbf{\Phi})$, inequalities in Theorem 8.12* (i) *give us*

$$S(r,r,\mathbf{A},\mathbf{\Phi}) \leq S(r,s,\mathbf{A},\mathbf{\Phi}) \leq S(s,s,\mathbf{A},\mathbf{\Phi}), \;\; r < s, \; s \geq 1$$

and

$$S(r,r,\mathbf{A},\mathbf{\Phi}) \leq S(s,r,\mathbf{A},\mathbf{\Phi}) \leq S(s,s,\mathbf{A},\mathbf{\Phi}), \;\; r < s, \; r \leq -1.$$

An open problem is to give list of inequalities comparing "mixed means" $S(r,s,\mathbf{A},\mathbf{\Phi})$ in remaining cases.

Finally, we consider quasi-arithmetic means of Mercer's type for positive linear mappings.

Let $\mathbf{A}$ and $\mathbf{\Phi}$ be as in the previous context and $m < M$. Let $\varphi, \psi \in C([m,M])$ be strictly monotonic functions on an interval $[m,M]$. We define

$$\widetilde{M}_\varphi(\mathbf{A},\mathbf{\Phi}) := \varphi^{-1}\left(\varphi(m) I_K + \varphi(M) I_K - \sum_{i=1}^{n} \Phi_i(\varphi(A_i))\right). \tag{8.29}$$

It is easy to see that $\widetilde{M}_\varphi(\mathbf{A},\mathbf{\Phi})$ is well defined.

Theorem 8.13 *Under the above hypotheses,*

(i) *if either $\psi \circ \varphi^{-1}$ is convex and ψ^{-1} is operator monotone, or $\psi \circ \varphi^{-1}$ is concave and $-\psi^{-1}$ is operator monotone, then*

$$\widetilde{M}_\varphi(\mathbf{A},\mathbf{\Phi}) \leq \widetilde{M}_\psi(\mathbf{A},\mathbf{\Phi}). \tag{8.30}$$

In fact, to be more specific, we have the following series of inequalities

$$\begin{aligned} &\widetilde{M}_\varphi(\mathbf{A},\mathbf{\Phi}) \\ &\leq \psi^{-1}\left(\frac{\varphi(M) I_K - \sum_{i=1}^{n} \Phi_i(\varphi(A_i))}{\varphi(M) - \varphi(m)} \cdot \psi(M)\right. \\ &\left. + \frac{\sum_{i=1}^{n} \Phi_i(\varphi(A_i)) - \varphi(m) I_K}{\varphi(M) - \varphi(m)} \cdot \psi(m)\right) \\ &\leq \widetilde{M}_\psi(\mathbf{A},\mathbf{\Phi}). \end{aligned} \tag{8.31}$$

(ii) *if either $\psi \circ \varphi^{-1}$ is concave and ψ^{-1} is operator monotone, or $\psi \circ \varphi^{-1}$ is convex and $-\psi^{-1}$ is operator monotone, then inequalities* (8.30) *and* (8.31) *are reversed.*

Proof. Suppose that $\psi \circ \varphi^{-1}$ is convex. If in Theorem 8.10 we let $f = \psi \circ \varphi^{-1}$ and replace A_i, m and M with $\varphi(A_i)$, $\varphi(m)$ and $\varphi(M)$, respectively, then we obtain

$$
\begin{aligned}
&\left(\psi\circ\varphi^{-1}\right)\left(\varphi(m)I_K+\varphi(M)I_K-\sum_{i=1}^n\Phi_i(\varphi(A_i))\right)\\
&\le \frac{\varphi(M)I_K-\sum_{i=1}^n\Phi_i(\varphi(A_i))}{\varphi(M)-\varphi(m)}\cdot\left(\psi\circ\varphi^{-1}\right)(\varphi(M))\\
&+\frac{\sum_{i=1}^n\Phi_i(\varphi(A_i))-\varphi(m)I_K}{\varphi(M)-\varphi(m)}\cdot\left(\psi\circ\varphi^{-1}\right)(\varphi(m))\\
&\le\left(\psi\circ\varphi^{-1}\right)(\varphi(m))I_K+\left(\psi\circ\varphi^{-1}\right)(\varphi(M))I_K-\sum_{i=1}^n\Phi_i\left(\left(\psi\circ\varphi^{-1}\right)(\varphi(A_i))\right)
\end{aligned}
$$

or

$$
\begin{aligned}
&\psi\left(\varphi^{-1}\left(\varphi(m)I_K+\varphi(M)I_K-\sum_{i=1}^n\Phi_i(\varphi(A_i))\right)\right)\\
&\le\frac{\varphi(M)I_K-\sum_{i=1}^n\Phi_i(\varphi(A_i))}{\varphi(M)-\varphi(m)}\cdot\psi(M)+\frac{\sum_{i=1}^n\Phi_i(\varphi(A_i))-\varphi(m)I_K}{\varphi(M)-\varphi(m)}\cdot\psi(m)\\
&\le\psi(m)I_K+\psi(M)I_K-\sum_{i=1}^n\Phi_i(\psi(A_i)). \qquad (8.32)
\end{aligned}
$$

If $\psi\circ\varphi^{-1}$ is concave then we obtain the reverse of inequalities (8.32).

If ψ^{-1} is operator monotone, then (8.32) implies (8.31). If $-\psi^{-1}$ is operator monotone, then the reverse of (8.32) implies (8.31). Analogously, we get the reverse of (8.31) in the cases when $\psi\circ\varphi^{-1}$ is convex and $-\psi^{-1}$ is operator monotone, or $\psi\circ\varphi^{-1}$ is concave and ψ^{-1} is operator monotone. □

8.4 Chaotic order version

Let $\mathbf{A}=(A_1,\ldots,A_n)$ be an $n-$tuple of positive invertible operators in $B(H)$ with $Sp(A_i)\subseteq[m,M]$ for some scalars $0<m<M$, and let $\mathbf{\Phi}=(\Phi_1,\ldots,\Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i: B(H)\to B(K)$, $i=1,\ldots,n$. We recall that we define the r-th power operator mean for $r\in\mathbb{R}$ as

$$
M_r(\mathbf{A},\mathbf{\Phi}):=\begin{cases}\left(\sum_{i=1}^n\Phi_i(A_i^r)\right)^{\frac{1}{r}}, & r\neq 0,\\ \exp\left(\sum_{i=1}^n\Phi_i(\log(A_i))\right), & r=0.\end{cases} \qquad (8.33)
$$

The ordering among these means is given in Chapter 9. Here we discuss the chaotic ordering among them: the chaotic order $A \gg B$ for $A, B > 0$ means $\log A \geq \log B$, also see 3.4.

The following theorems are generalizations of the theorems in [124, p.135, 136].

Theorem 8.14 *If $r, s \in \mathbb{R}$, $r < s$, then*

$$M_r(\mathbf{A}, \mathbf{\Phi}) \ll M_s(\mathbf{A}, \mathbf{\Phi}).$$

Proof. STEP 1: Assume $0 < r < s$. Applying Theorem 8.9 to the operator concave function $f(t) = t^{\frac{r}{s}}$ (note that $0 < \frac{r}{s} < 1$ here) and replacing A_i with A_i^s we have

$$\left(\sum_{i=1}^{n} \Phi_i\left(A_i^s\right)\right)^{\frac{r}{s}} \geq \sum_{i=1}^{n} \Phi_i\left(A_i^r\right).$$

Since the function $f(t) = \log t$ is operator monotone and $r > 0$, it follows that

$$\frac{1}{s} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^s\right)\right) \geq \frac{1}{r} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right),$$

i.e.
$$\log M_r(\mathbf{A}, \mathbf{\Phi}) \leq \log M_s(\mathbf{A}, \mathbf{\Phi}).$$

STEP 2: Assume $r < s < 0$. Applying Theorem 8.9 to the operator concave function $f(t) = t^{\frac{s}{r}}$ (note that $0 < \frac{s}{r} < 1$ here) and replacing A_i with A_i^r we have

$$\left(\sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right)^{\frac{s}{r}} \geq \sum_{i=1}^{n} \Phi_i\left(A_i^s\right).$$

Since $s < 0$, it follows that

$$\frac{1}{r} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right) \leq \frac{1}{s} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^s\right)\right),$$

i.e.
$$\log M_r(\mathbf{A}, \mathbf{\Phi}) \leq \log M_s(\mathbf{A}, \mathbf{\Phi}).$$

STEP 3: Assume $r < 0 = s$. Applying Theorem 8.9 to the operator convex function $f(t) = \frac{1}{r} \log t$ (note that $\frac{1}{r} < 0$ here) and replacing A_i with A_i^r we have

$$\frac{1}{r} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^r\right)\right) \leq \sum_{i=1}^{n} \Phi_i \log\left(A_i\right),$$

i.e.
$$\log M_r(\mathbf{A}, \mathbf{\Phi}) \leq \log M_0(\mathbf{A}, \mathbf{\Phi}).$$

STEP 4: Assume $r = 0 < s$. Applying Theorem 8.9 to the operator concave function $f(t) = \frac{1}{s} \log t$ and replacing A_i with A_i^s we have

$$\frac{1}{s} \log\left(\sum_{i=1}^{n} \Phi_i\left(A_i^s\right)\right) \geq \sum_{i=1}^{n} \Phi_i\left(\log\left(A_i\right)\right),$$

i.e. $$\log M_0(\mathbf{A},\mathbf{\Phi}) \le \log M_s(\mathbf{A},\mathbf{\Phi}).$$

STEP 5: Assume $r<0<s$. From STEP 3 and STEP 4 it follows that
$$\log M_r(\mathbf{A},\mathbf{\Phi}) \le \log M_0(\mathbf{A},\mathbf{\Phi}) \le \log M_s(\mathbf{A},\mathbf{\Phi}).$$ □

To prove the following theorem, we need the following lemma.

Lemma 8.1 *Let $(A_1,\ldots,A_n)$ be an $n-$tuple of self-adjoint operators in $B(H)$ with spectra in $[m,M]$ for some scalars $0<m<M$, and $(\Phi_1,\ldots,\Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i : B(H)\to B(K)$, $i=1,\ldots,n$. Denote $h=\frac{M}{m}$. Then*

$$\alpha_2\left(\sum_{i=1}^n \Phi_i(A_i)\right)^p \le \sum_{i=1}^n \Phi_i(A_i^p) \le \alpha_1\left(\sum_{i=1}^n \Phi_i(A_i)\right)^p$$

for

$$\alpha_1 = \begin{cases} K(h,p) & \text{if} \quad p<0 \text{ or } 1<p, \\ 1 & \text{if} \quad 0<p<1, \end{cases}$$
$$\alpha_2 = \begin{cases} K(h,p)^{-1} & \text{if} \quad p<-1 \text{ or } 2<p, \\ 1 & \text{if} \quad -1\ge p<0 \text{ or } 1\le p\le 2, \\ K(h,p) & \text{if} \quad 0<p<1, \end{cases}$$

where the generalized Kantorovich constant $K(m,M,p)$ is defined by (2.29).

Proof. This lemma is proved in a similar way as [124, Lemma 4.13] using converses of Jensen's inequality. □

Theorem 8.15 *If $r,s\in\mathbb{R}$, $r<s$, then*

$$\Delta(h,r,s)^{-1}M_s(\mathbf{A},\mathbf{\Phi}) \ll M_r(\mathbf{A},\mathbf{\Phi}), \tag{8.34}$$

where the generalized Specht ratio $\Delta(h,r,s)$ cf. [124, eq. (2.97)] for $h>0$ is defined as

$$\Delta(h,r,s) = \begin{cases} K\left(h^r,\frac{s}{r}\right)^{\frac{1}{s}} & \text{if} \quad r<s,\ r,s\neq 0, \\ \left(\dfrac{e\log h^{\frac{p}{h^p-1}}}{h^{\frac{p}{h^p-1}}}\right)^{\frac{sgn(p)}{p}} & \text{if} \quad r=0<s=p, \text{ or } r=p<s=0. \end{cases} \tag{8.35}$$

Proof. STEP 1: Assume $0<r<s$. Then $0<m^sI_K\le\sum_{i=1}^n\Phi_i(A_i^s)\le M^sI_K$. Applying Lemma 8.1 with $p=\frac{r}{s}$ $(0<p<1)$ and replacing A_i with A_i^s we obtain

$$K\left(h^s,\frac{r}{s}\right)\left(\sum_{i=1}^n\Phi_i(A_i^s)\right)^{\frac{r}{s}} \le \sum_{i=1}^n\Phi_i(A_i^r).$$

Since the function $f(t)=\log t$ is operator monotone and $r>0$, we have

$$\log\left(K\left(h^s,\frac{r}{s}\right)^{\frac{1}{r}}\left(\sum_{i=1}^n\Phi_i(A_i^s)\right)^{\frac{1}{s}}\right) \le \log\left(\sum_{i=1}^n\Phi_i(A_i^r)\right)^{\frac{1}{r}},$$

i.e.

$$\log\left(\Delta(h,r,s)^{-1}M_s(\mathbf{A},\mathbf{\Phi})\right)\leq\log M_r(\mathbf{A},\mathbf{\Phi}). \tag{8.36}$$

Since $\Delta(h,s,r)=\Delta(h,r,s)^{-1}$ (see [124, p. 87]), (8.34) follows from (8.36).
STEP 2: Assume $r<s<0$. Then, $0<M^r I_K\leq\sum_{i=1}^n\Phi_i(A_i^r)\leq m^r I_K$. Applying Lemma 8.1 with $p=\frac{s}{r}$ $(0<p<1)$ and replacing A_i with A_i^r we obtain

$$K\left(h^r,\frac{s}{r}\right)\left(\sum_{i=1}^n\Phi_i(A_i^r)\right)^{\frac{s}{r}}\leq\sum_{i=1}^n\Phi_i(A_i^s).$$

Since $s<0$, we have

$$\log\left(K\left(h^r,\frac{s}{r}\right)^{\frac{1}{s}}\left(\sum_{i=1}^n\Phi_i(A_i^r)\right)^{\frac{1}{r}}\right)\geq\log\left(\sum_{i=1}^n\Phi_i(A_i^s)\right)^{\frac{1}{s}},$$

i.e.

$$\log\left(\Delta(h,r,s)M_r(\mathbf{A},\mathbf{\Phi})\right)\geq\log M_s(\mathbf{A},\mathbf{\Phi}). \tag{8.37}$$

Now, (8.34) follows from (8.37).
STEP 3: Assume $r<0<s$. If $0<-r<s$ or $0<s<-r$, we let $p=\frac{r}{s}$ or $p=\frac{s}{r}$ in Lemma 8.1 $(-1<p<0)$, respectively. Then we obtain

$$\sum_{i=1}^n\Phi_i(A_i^r)\leq K\left(h^s,\frac{r}{s}\right)\left(\sum_{i=1}^n\Phi_i(A_i^s)\right)^{\frac{r}{s}}$$

or

$$\sum_{i=1}^n\Phi_i(A_i^s)\leq K\left(h^r,\frac{s}{r}\right)\left(\sum_{i=1}^n\Phi_i(A_i^r)\right)^{\frac{s}{r}}.$$

So we have

$$\log M_r(\mathbf{A},\mathbf{\Phi})\geq\log\left(\Delta(h,r,s)^{-1}M_s(\mathbf{A},\mathbf{\Phi})\right)$$

or

$$\log M_s(\mathbf{A},\mathbf{\Phi})\leq\log\left(\Delta(h,r,s)M_r(\mathbf{A},\mathbf{\Phi})\right).$$

STEP 4: Assume $r=0<s$. If $r\to0$ in (8.36), then

$$\log\left(\Delta(h,0,s)^{-1}M_s(\mathbf{A},\mathbf{\Phi})\right)\leq\log M_0(\mathbf{A},\mathbf{\Phi}).$$

STEP 5: Assume $r<s=0$. If $s\to0$ in (8.37), then

$$\log M_0(\mathbf{A},\mathbf{\Phi})\leq\log\left(\Delta(h,r,0)M_r(\mathbf{A},\mathbf{\Phi})\right).$$

□

Next, we consider the chaotic ordering among Mercer's power operator means defined by (8.18).

Theorem 8.16 *If $r,s \in \mathbb{R}$, $r < s$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \ll \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).$$

Proof. Analogously to the proof of Theorem 8.14, but using Theorem 8.10 instead of Theorem 8.9. □

Now, we define, for any $r,s \in \mathbb{R}$

$$R(r,s,\mathbf{A},\mathbf{\Phi}) := \begin{cases} \left[\sum_{i=1}^{n} \Phi_i\left([(m^r+M^r)I_H - A_i^r]^{\frac{s}{r}}\right)\right]^{\frac{1}{s}}, & r \neq 0, s \neq 0, \\ \exp\left(\sum_{i=1}^{n} \Phi_i\left(\log\,[(m^r+M^r)I_H - A_i^r]^{\frac{1}{r}}\right)\right), & r \neq 0, s = 0, \\ \left[\sum_{i=1}^{n} \Phi_i\left(\exp s\,[(\log mM)I_H - \log A_i]\right)\right]^{\frac{1}{s}}, & r = 0, s \neq 0, \end{cases} \tag{8.38}$$

and $S(r,s,\mathbf{A},\mathbf{\Phi})$ by (8.23). It is easy to see that $R(r,s,\mathbf{A},\mathbf{\Phi})$ is well defined and also notice that $R(r,r,\mathbf{A},\mathbf{\Phi}) = S(r,r,\mathbf{A},\mathbf{\Phi}) = \widetilde{M}_r(\mathbf{A},\mathbf{\Phi})$ (including $r = 0$).

Theorem 8.17 *Let $r,s \in \mathbb{R}$, $r < s$.*
(i) *If $r \geq 0$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \ll S(s,r,\mathbf{A},\mathbf{\Phi}) \ll R(s,r,\mathbf{A},\mathbf{\Phi}) \ll \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.39}$$

(ii) *If $s \leq 0$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \ll R(r,s,\mathbf{A},\mathbf{\Phi}) \ll S(r,s,\mathbf{A},\mathbf{\Phi}) \ll \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \tag{8.40}$$

(iii) *If $r < 0 < s$, then*

$$\begin{aligned} \widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) &\ll R(r,0,\mathbf{A},\mathbf{\Phi}) \ll S(r,0,\mathbf{A},\mathbf{\Phi}) \ll \widetilde{M}_0(\mathbf{A},\mathbf{\Phi}) \\ &\ll S(s,0,\mathbf{A},\mathbf{\Phi}) \ll R(s,0,\mathbf{A},\mathbf{\Phi}) \ll \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \end{aligned} \tag{8.41}$$

Proof. (i) STEP 1: Assume $0 < r < s$. Applying Theorem 8.10 to the operator concave function $f(t) = t^{\frac{r}{s}}$ (note that $0 < \frac{r}{s} < 1$ here) and replacing A_i, m and M with A_i^s, m^s and M^s we have

$$\begin{aligned} &\left((m^s+M^s)I_K - \sum_{i=1}^{n} \Phi_i(A_i^s)\right)^{\frac{r}{s}} \geq \sum_{i=1}^{n} \Phi_i\left(((m^s+M^s)I_H - A_i^s)^{\frac{r}{s}}\right) \\ &\geq \frac{M^sI_K - S_s}{M^s-m^s}M^r + \frac{S_s - m^sI_K}{M^s-m^s}m^r \geq (m^r+M^r)I_K - \sum_{i=1}^{n} \Phi_i(A_i^r). \end{aligned}$$

Since the function $f(t) = \log t$ is operator monotone and $r > 0$, it follows that (8.39) holds.

STEP 2: Assume $r = 0 < s$. Applying Theorem 8.10 to the operator concave function $f(t) = \frac{1}{s}\log t$ and replacing A_i, m and M with A_i^s, m^s and M^s we have

$$\frac{1}{s}\log\left((m^s+M^s)I_K - \sum_{i=1}^{n}\Phi_i(A_i^s)\right) \geq \sum_{i=1}^{n}\Phi_i\left(\frac{1}{s}\log((m^s+M^s)I_H - A_i^s)\right)$$
$$\geq \frac{M^sI_K - S_s}{M^s - m^s}\log M + \frac{S_s - m^sI_K}{M^s - m^s}\log m \geq (\log mM)I_K - \sum_{i=1}^{n}\Phi_i(\log(A_i)),$$

which gives (8.39) for $r = 0$.

(ii) STEP 1: Assume $r < s < 0$. Applying Theorem 8.10 to the operator concave function $f(t) = t^{\frac{s}{r}}$ (note that $0 < \frac{s}{r} < 1$ here) and replacing A_i, m and M with A_i^r, m^r and M^r we have

$$\left((m^r+M^r)I_K - \sum_{i=1}^{n}\Phi_i(A_i^r)\right)^{\frac{s}{r}} \geq \sum_{i=1}^{n}\Phi_i\left(((m^r+M^r)I_H - A_i^r)^{\frac{s}{r}}\right)$$
$$\geq \frac{M^rI_K - S_r}{M^r - m^r}M^s + \frac{S_r - m^rI_K}{M^r - m^r}m^s \geq (m^s+M^s)I_K - \sum_{i=1}^{n}\Phi_i(A_i^s).$$

Since $s < 0$, it follows that (8.40) holds.

STEP 2: Assume $r < 0 = s$. Applying Theorem 8.10 to the operator convex function $f(t) = \frac{1}{r}\log t$ (note that $\frac{1}{r} < 0$ here) and replacing A_i, m and M with A_i^r, m^r and M^r we have

$$\frac{1}{r}\log\left((m^r+M^r)I_K - \sum_{i=1}^{n}\Phi_i(A_i^r)\right) \leq \sum_{i=1}^{n}\Phi_i\left(\frac{1}{r}\log((m^r+M^r)I_H - A_i^r)\right)$$
$$\leq \frac{M^rI_K - S_r}{M^r - m^r}\log M + \frac{S_r - m^rI_K}{M^r - m^r}\log m \leq (\log mM)I_K - \sum_{i=1}^{n}\Phi_i(\log(A_i)),$$

which gives (8.40) for $s = 0$.

(iii) Assume $r < 0 < s$. The desired inequality (8.41) follows set $s = 0$ in (ii) and $r = 0$ in (i). □

Remark 8.2 *If we define by $\widetilde{M}_r(B) = (m^r1 + M^r1 - B^r)^{\frac{1}{r}}$ (Mercer's mean for positive invertible operator B with $Sp(B)\subset[m,M]$, $0<m<M$) and by $\widetilde{M}_r(\mathbf{A}) = (\widetilde{M}_r(A_1),\ldots,\widetilde{M}_r(A_n))$ (for $n-$tuple $\mathbf{A}$ of positive invertible operators), we can write:*

$$\widetilde{M}_r(\mathbf{A},\boldsymbol{\Phi}) = \widetilde{M}_r(M_r(\mathbf{A},\boldsymbol{\Phi}))$$

$$R(r,s,\mathbf{A},\boldsymbol{\Phi}) = M_s(\widetilde{M}_r(\mathbf{A}),\boldsymbol{\Phi}),$$

so we can describe inequalities in Theorem 8.17 as mixed mean inequalities. One can also ask the question: What is the complete set of inequalities among mixed means $M_r(\widetilde{M}_s(\mathbf{A}),\boldsymbol{\Phi})$, $\widetilde{M}_s(M_r(\mathbf{A},\boldsymbol{\Phi}))$, $\widetilde{M}_r(M_s(\mathbf{A},\boldsymbol{\Phi}))$ and $M_s(\widetilde{M}_r(\mathbf{A}),\boldsymbol{\Phi})$ under the chaotic order? One part of the answer is in Theorem 8.16 and Theorem 8.17.

8.5 Refinements

In this section, we give a refinement of Mercer's inequality for operator convex functions. We use that result to refine monotonicity properties of power means of Mercer's type for operators. Finally, we consider related quasi-arithmetic means for operators.

Theorem 8.18 *Let $(A_1,\dots,A_n)$ be an $n-$tuple of self-adjoint operators in $B(H)$ with spectra in $[m,M]$ for some scalars $0<m<M$, and $(\Phi_1,\dots,\Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i : B(H)\to B(K)$, $i=1,\dots,n$. If $f\in C([m,M])$ is operator convex, then we have the following series of inequalities*

$$\begin{aligned}
&f\left(mI_K+MI_K-\sum_{i=1}^{n}\Phi_i(A_i)\right)\leq\sum_{i=1}^{n}\Phi_i\left(f\left(mI_H+MI_H-A_i\right)\right)\\
&\leq\frac{MI_K-\sum_{i=1}^{n}\Phi_i(A_i)}{M-m}\cdot f(M)+\frac{\sum_{i=1}^{n}\Phi_i(A_i)-mI_K}{M-m}\cdot f(m) \qquad (8.42)\\
&\leq f(m)I_K+f(M)I_K-\sum_{i=1}^{n}\Phi_i\left(f(A_i)\right).
\end{aligned}$$

If a function f is operator concave, then the inequalities (8.42) *are reversed.*

Proof. The proof of this theorem is quite similar to the proof of Theorem 8.6. We omit the details. □

We give applications to the ordering among Mercer's power operator means defined by (8.18).

Let $R(r,s,\mathbf{A},\mathbf{\Phi})$ and $S(r,s,\mathbf{A},\mathbf{\Phi})$ are defined by (8.38) and (8.23), respectively. To simplify notations, in what follows we will write $\widetilde{M}_r$, $R(r,s)$, $S(r,s)$ instead of $\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})$, $R(r,s,\mathbf{A},\mathbf{\Phi})$, $S(r,s,\mathbf{A},\mathbf{\Phi})$, respectively.

Figure 8.1 illustrates regions $(i)-(vii)$ which determine the seven cases occurring in Theorem 8.19.

Theorem 8.19 *Let $r,s\in\mathbb{R}$, $r<s$.*

(i) *If $1\leq r$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\leq S(s,r,\mathbf{A},\mathbf{\Phi})\leq R(s,r,\mathbf{A},\mathbf{\Phi})\leq\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}). \qquad (8.43)$$

(ii) *If $s\leq -1$, then*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\leq R(r,s,\mathbf{A},\mathbf{\Phi})\leq S(r,s,\mathbf{A},\mathbf{\Phi})\leq\widetilde{M}(\mathbf{A},\mathbf{\Phi})_s. \qquad (8.44)$$

(iii) *If $r\leq -1, s\geq 1$, then*

$$\begin{aligned}
&\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\leq R(r,-1,\mathbf{A},\mathbf{\Phi})\leq S(r,-1,\mathbf{A},\mathbf{\Phi})\\
&\leq\widetilde{M}_{-1}(\mathbf{A},\mathbf{\Phi})\leq S(1,-1,\mathbf{A},\mathbf{\Phi})\leq R(1,-1,\mathbf{A},\mathbf{\Phi})\\
&\leq\widetilde{M}_1(\mathbf{A},\mathbf{\Phi})\leq S(s,1,\mathbf{A},\mathbf{\Phi})\leq R(s,1,\mathbf{A},\mathbf{\Phi}) \qquad (8.45)\\
&\leq\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).
\end{aligned}$$

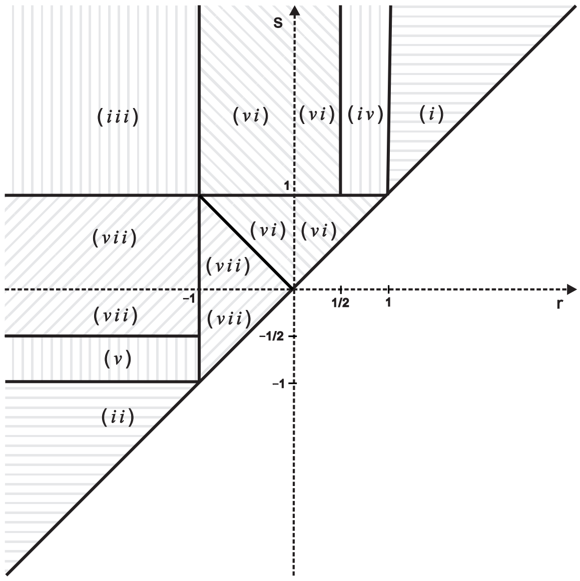


Figure 8.1: Regions $(i)-(vii)$

(iv) *If* $\frac{1}{2}<r<1<s$, *then*

$$\begin{aligned}&\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\le R(r,1,\mathbf{A},\mathbf{\Phi})\le S(r,1,\mathbf{A},\mathbf{\Phi})\\ \le\ &\widetilde{M}_1(\mathbf{A},\mathbf{\Phi})\le S(s,1,\mathbf{A},\mathbf{\Phi})\le R(s,1,\mathbf{A},\mathbf{\Phi})\\ \le\ &\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).\end{aligned}\tag{8.46}$$

(v) *If* $r<-1<s<-\frac{1}{2}$, *then*

$$\begin{aligned}&\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\le R(r,-1,\mathbf{A},\mathbf{\Phi})\le S(r,-1,\mathbf{A},\mathbf{\Phi})\\ \le\ &\widetilde{M}_{-1}(\mathbf{A},\mathbf{\Phi})\le S(s,-1,\mathbf{A},\mathbf{\Phi})\le R(s,-1,\mathbf{A},\mathbf{\Phi})\\ \le\ &\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).\end{aligned}\tag{8.47}$$

(vi) *If* $-1<r\le\frac{1}{2}, s\ge 1$; *or* $-s\le r<s\le 1$, *then*

$$\begin{aligned}\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\ &\le\ C(m,M,r)\,S(s,r,\mathbf{A},\mathbf{\Phi})\le C(m,M,r)^2\,R(s,r,\mathbf{A},\mathbf{\Phi})\\ &\le\ C(m,M,r)^3\,\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).\end{aligned}\tag{8.48}$$

(vii) *If* $r\le -1, -\frac{1}{2}\le s<1$; *or* $-1\le r<s\le -r$, *then*

$$\begin{aligned}\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\ &\le\ C(m,M,s)\,R(r,s,\mathbf{A},\mathbf{\Phi})\le C(m,M,s)^2\,S(r,s,\mathbf{A},\mathbf{\Phi})\\ &\le\ C(m,M,s)^3\,\widetilde{M}_s(\mathbf{A},\mathbf{\Phi}).\end{aligned}\tag{8.49}$$

Proof. To simplify notations, in this proof we will write $\widetilde{M}_r$, $R(r,s)$, $S(r,s)$ instead of $\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})$, $R(r,s,\mathbf{A},\mathbf{\Phi})$, $S(r,s,\mathbf{A},\mathbf{\Phi})$, respectively.

(i) Suppose that $1 \le r < s$.

Applying inequalities (8.42) to the operator concave function $f(t) = t^{\frac{r}{s}}$ (note that $0 < \frac{r}{s} \le 1$ here) and replacing A_i, m and M with A_i^s, m^s and M^s, respectively, we have

$$\left[\widetilde{M}_s\right]^r \ge [R(s,r)]^r \ge [S(s,r)]^r \ge \left[\widetilde{M}_r\right]^r. \tag{8.50}$$

Raising these inequalities to the power $\frac{1}{r}$, by the Löwner-Heinz theorem it follows that (8.43) holds.

(ii) Suppose that $r < s \le -1$.

Applying inequalities (8.43) to $\mathbf{A}^{-1} = \left(A_1^{-1},\ldots,A_n^{-1}\right)$ and observing that $\widetilde{M}_{-r}(\mathbf{A}^{-1},\mathbf{\Phi}) = \left[\widetilde{M}_r(\mathbf{A},\mathbf{\Phi})\right]^{-1}$, $\widetilde{M}_{-s}(\mathbf{A}^{-1},\mathbf{\Phi}) = \left[\widetilde{M}_s(\mathbf{A},\mathbf{\Phi})\right]^{-1}$, $S\left(-r,-s,\mathbf{A}^{-1},\mathbf{\Phi}\right) = [S(r,s,\mathbf{A},\mathbf{\Phi})]^{-1}$, $R\left(-r,-s,\mathbf{A}^{-1},\mathbf{\Phi}\right) = [R(r,s,\mathbf{A},\mathbf{\Phi})]^{-1}$, we have

$$\left[\widetilde{M}_s\right]^{-1} \le [S(r,s)]^{-1} \le [R(r,s)]^{-1} \le \left[\widetilde{M}_r\right]^{-1}.$$

Hence, (8.44) holds.

(iii) Suppose that $r \le -1$ and $s \ge 1$.

Applying inequalities (8.42) to the operator convex function $f(t) = t^{-1}$ we have

$$\left[\widetilde{M}_1\right]^{-1} \le [R(1,-1)]^{-1} \le [S(1,-1)]^{-1} \le \left[\widetilde{M}_{-1}\right]^{-1}.$$

Hence,

$$\widetilde{M}_{-1} \le S(1,-1) \le R(1,-1) \le \widetilde{M}_1.$$

If we let $r = 1$ in (8.43) and $s = -1$ in (8.43) then it follows that $\widetilde{M}_1 \le S(s,1) \le R(s,1) \le \widetilde{M}_s$ and $\widetilde{M}_r \le R(r,-1) \le S(r,-1) \le \widetilde{M}_{-1}$ holds. Hence, (8.45) holds.

(iv) Suppose that $\frac{1}{2} < r < 1 < s$.

Applying inequalities (8.42) to the operator convex function $f(t) = t^{\frac{1}{r}}$ and replacing A_i, m and M with A_i^r, m^r and M^r, respectively, we have

$$\widetilde{M}_r \le R(r,1) \le S(r,1) \le \widetilde{M}_1.$$

If we let $r = 1$ in (8.43) then it follows that (8.46) holds.

(v) Suppose that $r < -1 < s < -\frac{1}{2}$.

Applying inequalities (8.46) to $\mathbf{A}^{-1} = \left(A_1^{-1},\ldots,A_n^{-1}\right)$ and following analogous arguing as in (ii), we obtain (8.47).

(vi)

STEP 1: Suppose that $0 < r \le \frac{1}{2}, 1 \le s$.

In the same way as in (i) we obtain that (8.50) holds in this case. Raising (8.50) to the power $\frac{1}{r}$, by Theorem 4.3 (i) it follows that (8.48) holds.

STEP 2: Suppose that $-1 < r < 0$, $1 \le s$.

Applying inequalities (8.42) to the operator convex function $f(t) = t^{\frac{r}{s}}$ and replacing A_i, m and M with A_i^s, m^s and M^s, respectively, we have

$$\left[\widetilde{M}_s\right]^r \le [R(s,r)]^r \le [S(s,r)]^r \le \left[\widetilde{M}_r\right]^r.$$

Raising these inequalities to the power $\frac{1}{r}$, by Theorem 4.3 (ii) it follows that (8.48) holds (since $K(M,m,p) = K(m,M,p)$ by [124, p. 77]).

STEP 3: Suppose that $0 < r < s \le 1$ and $-1 \le -s \le r < 0$.

In the same way as in STEP 1 and STEP 2, we have (8.48).

STEP 4: Suppose that $0 = r < s$.

Applying inequalities (8.42) to the operator concave function $f(t) = \frac{1}{s}\log t$ and replacing A_i, m and M with A_i^s, m^s and M^s, respectively, we obtain

$$\log \widetilde{M}_s \ge \log R(s,0) \ge \log S(s,0) \ge \log \widetilde{M}_0.$$

By using Theorem 4.7, it follows that (8.48) holds for $r = 0$.

(vii) Suppose that $r \le -1, -\frac{1}{2} \le s < 1$; or $-1 \le r < s \le -r$.

Applying inequalities (8.48) to $\mathbf{A}^{-1} = \left(A_1^{-1}, \ldots, A_n^{-1}\right)$ and following analogous arguing as in (ii), we obtain (8.49). □

Remark 8.3 *Besides these results in Theorem 8.19, one can prove in the same way that for $r < s < 2r, s > 1$*

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \le R(r,s,\mathbf{A},\mathbf{\Phi}) \le S(r,s,\mathbf{A},\mathbf{\Phi}) \le \widetilde{M}_s(\mathbf{A},\mathbf{\Phi}),$$

and for $r < s < \frac{1}{2}r, r < -1$

$$\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) \le S(s,r,\mathbf{A},\mathbf{\Phi}) \le R(s,r,\mathbf{A},\mathbf{\Phi}) \le \widetilde{M}_s(\mathbf{A},\mathbf{\Phi})$$

also hold, but to include these cases in the figure we should compare sequences of inequalities in common regions (see Remark 8.4).

Remark 8.4 *If we define by $M_r(\mathbf{A},\mathbf{\Phi}) = \left(\sum_{i=1}^n \Phi_i(A_i^r)\right)^{\frac{1}{r}}$ (the weighted power mean), by $\widetilde{M}_r(B) = (m^r 1 + M^r 1 - B^r)^{\frac{1}{r}}$ (Mercer's mean for positive invertible operator B with $Sp(B) \subset [m,M]$, $0 < m < M$) and by $\widetilde{M}_r(\mathbf{A}) = (\widetilde{M}_r(A_1), \ldots, \widetilde{M}_r(A_n))$ (for an $n-$tuple $\mathbf{A}$ of positive invertible operators), we can write:*

$$\begin{aligned}\widetilde{M}_r(\mathbf{A},\mathbf{\Phi}) &= \widetilde{M}_r(M_r(\mathbf{A},\mathbf{\Phi})),\\ R(r,s,\mathbf{A},\mathbf{\Phi}) &= M_s(\widetilde{M}_r(\mathbf{A}),\mathbf{\Phi}),\end{aligned}$$

so one can describe inequalities in Theorem 8.19 as mixed mean inequalities. We can also state the following open problem: What is the complete set of inequalities among mixed means $M_r(\widetilde{M}_s(\mathbf{A}),\mathbf{\Phi})$, $\widetilde{M}_s(M_r(\mathbf{A},\mathbf{\Phi}))$, $\widetilde{M}_r(M_s(\mathbf{A},\mathbf{\Phi}))$ and $M_s(\widetilde{M}_r(\mathbf{A}),\mathbf{\Phi})$? Some special cases are given in Theorem 8.19 and Remark 8.3. Also, it is easy to see that

$$\widetilde{M}_r(M_r(\mathbf{A},\mathbf{\Phi})) \le \widetilde{M}_s(M_r(\mathbf{A},\mathbf{\Phi}))$$

reduces to monotonicity property of Mercer's means, and that in some cases,

$$\widetilde{M}_s(M_s(\mathbf{A},\mathbf{\Phi})) \leq \widetilde{M}_s(M_r(\mathbf{A},\mathbf{\Phi}))$$

reduces to inequalities between $(\sum_{i=1}^n \Phi_i(A_i^r))^{s/r}$ *and* $\sum_{i=1}^n \Phi(A_i^s)$.

Finally, we consider quasi-arithmetic means of Mercer's type defined by (8.29).

Theorem 8.20 *Let* $\mathbf{A}$ *and* $\mathbf{\Phi}$ *be as in the previous context and* $m < M$. *Let* $\varphi, \psi \in C([m,M])$ *be strictly monotonic functions on an interval* $[m,M]$.

(i) *If either* $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is operator monotone, or* $\psi \circ \varphi^{-1}$ *is operator concave and* $-\psi^{-1}$ *is operator monotone, then*

$$\begin{aligned}
&\widetilde{M}_\varphi(\mathbf{A},\mathbf{\Phi}) \\
&\leq \psi^{-1}\left(\sum_{i=1}^n \Phi_i\left(\left(\psi\circ\varphi^{-1}\right)(\varphi(m)I_H + \varphi(M)I_H - \varphi(A_i))\right)\right) \qquad (8.51)\\
&\leq \psi^{-1}\left(\frac{\varphi(M)I_K - \sum_{i=1}^n \Phi_i(\varphi(A_i))}{\varphi(M)-\varphi(m)}\cdot\psi(M) + \frac{\sum_{i=1}^n \Phi_i(\varphi(A_i)) - \varphi(m)I_K}{\varphi(M)-\varphi(m)}\cdot\psi(m)\right)\\
&\leq \widetilde{M}_\psi(\mathbf{A},\mathbf{\Phi}).
\end{aligned}$$

(ii) *If either* $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is operator monotone, or* $\psi \circ \varphi^{-1}$ *is operator convex and* $-\psi^{-1}$ *is operator monotone, then the reverse inequalities* (8.51) *hold.*

Proof. The proof is quite similar to the proof of Theorem 8.8 and we omit the details.
□

Theorem 8.21 *Under the hypotheses of Theorem 8.20, we have*

(i) *if either* φ *is operator concave and* φ^{-1} *is operator monotone or* φ *is operator convex and* $-\varphi^{-1}$ *is operator monotone, and either* ψ *is operator convex and* ψ^{-1} *is operator monotone or* ψ *is operator concave and* $-\psi^{-1}$ *is operator monotone,* *then*

$$\begin{aligned}
&\widetilde{M}_{\varphi}(\mathbf{A},\mathbf{\Phi}) \\
&\le \varphi^{-1}\left(\frac{MI_K-\sum_{i=1}^n \Phi_i(A_i)}{M-m}\cdot\varphi(M)+\frac{\sum_{i=1}^n \Phi_i(A_i)-mI_K}{M-m}\cdot\varphi(m)\right) \\
&\le \varphi^{-1}\left(\sum_{i=1}^n \Phi_i(\varphi(mI_H+MI_H-A_i))\right) \\
&\le \widetilde{M}_1(\mathbf{A},\mathbf{\Phi}) \qquad (8.52)\\
&\le \psi^{-1}\left(\sum_{i=1}^n \Phi_i(\psi(mI_H+MI_H-A_i))\right) \\
&\le \psi^{-1}\left(\frac{MI_K-\sum_{i=1}^n \Phi_i(A_i)}{M-m}\cdot\psi(M)+\frac{\sum_{i=1}^n \Phi_i(A_i)-mI_K}{M-m}\cdot\psi(m)\right) \\
&\le \widetilde{M}_{\psi}(\mathbf{A},\mathbf{\Phi}).
\end{aligned}$$

(ii) *if either φ is operator convex and φ^{-1} is operator monotone or φ is operator concave and $-\varphi^{-1}$ is operator monotone, and either ψ is operator concave and ψ^{-1} is operator monotone or ψ is operator convex and $-\psi^{-1}$ is operator monotone, then the reverse inequalities* (8.52) *hold.*

Proof. Suppose that φ is operator concave and φ^{-1} is operator monotone, and ψ is operator convex and ψ^{-1} is operator monotone. By Theorem 8.18, we have

$$\begin{aligned}
&\varphi\left(mI_K+MI_K-\sum_{i=1}^n \Phi_i(A_i)\right) \\
&\ge \sum_{i=1}^n \Phi_i(\varphi(mI_H+MI_H-A_i)) \\
&\ge \frac{MI_K-\sum_{i=1}^n \Phi_i(A_i)}{M-m}\cdot\varphi(M)+\frac{\sum_{i=1}^n \Phi_i(A_i)-mI_K}{M-m}\cdot\varphi(m) \\
&\ge \varphi(m)I_K+\varphi(M)I_K-\sum_{i=1}^n \Phi_i(\varphi(A_i)).
\end{aligned}$$

Since φ^{-1} is operator monotone, it follows that

$$\begin{aligned}
&\widetilde{M}_{\varphi}(\mathbf{A},\mathbf{\Phi}) \\
&\le \varphi^{-1}\left(\frac{MI_K-\sum_{i=1}^n \Phi_i(A_i)}{M-m}\cdot\varphi(M)+\frac{\sum_{i=1}^n \Phi_i(A_i)-mI_K}{M-m}\cdot\varphi(m)\right) \\
&\le \varphi^{-1}\left(\sum_{i=1}^n \Phi_i(\varphi(mI_H+MI_H-A_i))\right) \\
&\le \widetilde{M}_1(\mathbf{A},\mathbf{\Phi}).
\end{aligned}$$

Also, by Theorem 8.18, we have

$$\begin{aligned}
&\psi\left(mI_K + MI_K - \sum_{i=1}^{n}\Phi_i(A_i)\right)\\
&\leq \sum_{i=1}^{n}\Phi_i\left(\psi\left(mI_H + MI_H - A_i\right)\right)\\
&\leq \frac{MI_K - \sum_{i=1}^{n}\Phi_i(A_i)}{M-m}\cdot\psi(M) + \frac{\sum_{i=1}^{n}\Phi_i(A_i) - mI_K}{M-m}\cdot\psi(m)\\
&\leq \psi(m)\,I_K + \psi(M)\,I_K - \sum_{i=1}^{n}\Phi_i\left(\psi(A_i)\right).
\end{aligned}$$

Since ψ^{-1} is operator monotone, it follows that

$$\begin{aligned}
&\widetilde{M}_1(\mathbf{A},\mathbf{\Phi})\\
&\leq \psi^{-1}\left(\sum_{i=1}^{n}\Phi_i\left(\psi\left(mI_H + MI_H - A_i\right)\right)\right)\\
&\leq \psi^{-1}\left(\frac{MI_K - \sum_{i=1}^{n}\Phi_i(A_i)}{M-m}\cdot\psi(M) + \frac{\sum_{i=1}^{n}\Phi_i(A_i) - mI_K}{M-m}\cdot\psi(m)\right)\\
&\leq \widetilde{M}_\psi(\mathbf{A},\mathbf{\Phi}).
\end{aligned}$$

Hence, we have inequalities (8.52). In remaining cases the proof is analogous. □

Remark 8.5 *Results given in this chapter we can generalize for continuous fields of operators, similarly to how it was done for Jensen's inequality in Chapter 9.*

8.6 Notes

For our exposition we have used Mercer [183, 184, 185], Matković-Pečarić [176, 177] and Matković-Pečarić-I. Perić [178, 179].

Chapter 9

Jensen's Operator Inequality

In this chapter, we give a general formulation of Jensen's operator inequality for some non-unital fields of positive linear mappings, and we consider different types of converse inequalities. We discuss the ordering among power functions in a general setting. As an application we get the order among power means and some comparison theorems for quasi-arithmetic means. We also give a refined calculation of bounds in converses of Jensen's operator inequality.

9.1 Continuous fields of operators

Let T be a locally compact Hausdorff space, and let $\mathscr{A}$ be a C^*-algebra of operators on a Hilbert space H. We say that a field $(x_t)_{t\in T}$ of operators in $\mathscr{A}$ is continuous if the function $t \mapsto x_t$ is norm continuous on T. If in addition μ is a bounded Radon measure on T and the function $t \mapsto \|x_t\|$ is integrable, then we can form the Bochner integral $\int_T x_t \, d\mu(t)$, which is the unique element in $\mathscr{A}$ such that

$$\varphi\left(\int_T x_t \,\mathrm{d}\mu(t)\right) = \int_T \varphi(x_t)\,\mathrm{d}\mu(t)$$

for every linear functional φ in the norm dual $\mathscr{A}^*$, cf. [137, Section 4.1].

Assume furthermore that there is a field $(\Phi_t)_{t\in T}$ of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another C^*-algebra $\mathscr{B}$ of operators on a Hilbert space K. We say that such a field is continuous if the function $t \mapsto \Phi_t(x)$ is continuous for every $x \in \mathscr{A}$. If the C^*-algebras are unital and the field $t \to \Phi_t(\mathbf{1})$ is integrable with integral equals $\mathbf{1}$, we say that $(\Phi_t)_{t\in T}$ is *unital*.

Theorem 9.1 *Let $f : J \to \mathbb{R}$ be an operator convex function defined on an interval J, and let $\mathscr{A}$ and $\mathscr{B}$ be unital C^*-algebras. If $(\Phi_t)_{t\in T}$ is a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ defined on a locally compact Hausdorff space T with a bounded Radon measure μ, then the inequality*

$$f\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \tag{9.1}$$

holds for every bounded continuous field $(x_t)_{t\in T}$ of self-adjoint elements in $\mathscr{A}$ with spectra contained in J.

Proof. We first note that the function $t \mapsto \Phi_t(x_t) \in \mathscr{B}$ is continuous and bounded, hence integrable with respect to the bounded Radon measure μ. We may organize the set $CB(T,\mathscr{A})$ of bounded continuous functions on T with values in $\mathscr{A}$ as a normed involutive algebra by applying the point-wise operations and setting

$$\|(y_t)_{t\in T}\| = \sup_{t\in T}\|y_t\| \qquad (y_t)_{t\in T} \in CB(T,\mathscr{A}),$$

and it is not difficult to verify that the norm is already complete and satisfy the C^*-identity. In fact, this is a standard construction in C^*-algebra theory. It follows that $f((x_t)_{t\in T}) = (f(x_t))_{t\in T}$. We then consider the mapping

$$\pi : CB(T,\mathscr{A}) \to M(\mathscr{B}) \subseteq B(K)$$

defined by setting

$$\pi\left((x_t)_{t\in T}\right) = \int_T \Phi_t(x_t)\,\mathrm{d}\mu(t),$$

and note that it is a unital positive linear mapping. Setting $x = (x_t)_{t\in T} \in CB(T,\mathscr{A})$, we use the Davis-Choi-Jensen inequality to obtain

$$f\left(\pi\left((x_t)_{t\in T}\right)\right) = f(\pi(x)) \le \pi(f(x)) = \pi\left(f\left((x_t)_{t\in T}\right)\right) = \pi\left(\left(f(x_t)\right)_{t\in T}\right),$$

but this is just the statement of the theorem. □

In the following theorem we give a converse of Jensen's inequality (9.1). For a function $f : [m,M] \to \mathbb{R}$ we use the standard notation:

$$\alpha_f = \frac{f(M)-f(m)}{M-m}, \qquad \beta_f = \frac{Mf(m)-mf(M)}{M-m}. \tag{9.2}$$

Theorem 9.2 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let $f,g : [m,M] \to \mathbb{R}$*

and $F : U \times V \to \mathbb{R}$ be functions such that $f([m,M]) \subset U$, $g([m,M]) \subset V$ and F is bounded. If F is operator monotone in the first variable and f is convex in the interval $[m,M]$, then

$$F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right] \le \sup_{m\le z\le M} F\left[\alpha_f z+\beta_f, g(z)\right]\mathbf{1}. \tag{9.3}$$

In the dual case (when f is concave) the opposite inequality holds in (9.3) *with* inf *instead of* sup.

Proof. For convex f the inequality $f(z) \le \alpha_f z + \beta_f$ holds for every $z \in [m,M]$. Thus, by using functional calculus, $f(x_t) \le \alpha_f x_t + \beta_f \mathbf{1}$ for every $t \in T$. Applying the positive linear mappings Φ_t and integrating, we obtain

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \le \alpha_f \int_T \Phi(x_t)\mathrm{d}\mu(t) + \beta_f \mathbf{1}.$$

Now, using operator monotonicity of $F(\cdot, v)$, we obtain

$$\begin{aligned}
&F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right]\\
&\le F\left[\alpha_f \int_T \Phi_t(x_t)\mathrm{d}\mu(t) + \beta_f \mathbf{1}, g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right]\\
&\le \sup_{m\le z\le M} F\left[\alpha_f z+\beta_f, g(z)\right]\mathbf{1}.
\end{aligned}$$

□

Numerous applications of the previous theorem can be given. For example, we give generalizations of some results from [281].

Theorem 9.3 *Let $(A_t)_{t\in T}$ be a continuous field of positive operators on a Hilbert space H defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. We assume the spectra are in $[m,M]$ for some $0 < m < M$. Let furthermore $(x_t)_{t\in T}$ be a continuous field of vectors in H such that $\int_T \|x_t\|^2 \mathrm{d}\mu(t) = 1$. Then for any $\lambda \ge 0$, $p \ge 1$ and $q \ge 1$ we have*

$$\left(\int_T \langle A_t^p x_t, x_t\rangle \mathrm{d}\mu(t)\right)^{1/q} - \lambda \int_T \langle A_t x_t, x_t\rangle \mathrm{d}\mu(t) \le C(\lambda, m, M, p, q), \tag{9.4}$$

where the constant

$$C(\lambda,m,M,p,q) = \begin{cases} M\left(M^{\frac{p}{q}-1} - \lambda\right), & 0 < \lambda \le \dfrac{\alpha_p}{q} M^{p\left(\frac{1}{q}-1\right)} \\ \dfrac{q-1}{q}\left(\dfrac{q}{\alpha_p}\lambda\right)^{\frac{1}{1-q}} + \dfrac{\beta_p}{\alpha_p}\lambda, & \dfrac{\alpha_p}{q} M^{p\left(\frac{1}{q}-1\right)} \le \lambda \le \dfrac{\alpha_p}{q} m^{p\left(\frac{1}{q}-1\right)} \\ m\left(m^{\frac{p}{q}-1} - \lambda\right), & \dfrac{\alpha_p}{q} m^{p\left(\frac{1}{q}-1\right)} \le \lambda \end{cases}$$

and α_p and β_p are the constants α_f and β_f associated with the function $f(z) = z^p$.

Proof. Applying Theorem 9.2 for the functions

$$f(z) = z^p, \qquad F(u,v) = u^{1/q} - \lambda v,$$

and unital fields of positive linear mappings $\Phi_t : B(H) \to \mathbb{C}$ defined by setting $\Phi_t(A_t) = \langle A_t x_t, x_t \rangle$ for $t \in T$, the problem is reduced to determine $\sup_{m \le z \le M} H(z)$ where $H(z) = (\alpha_p z + \beta_p)^{1/q} - \lambda z$. □

Applying Theorem 9.3 we obtain the following result with the $r-$geometric mean $A\#_r B$.

Corollary 9.1 *Let $(A_t)_{t\in T}$ and $(B_t)_{t\in T}$ be continuous fields of positive invertible operators on a Hilbert space H defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ such that*

$$m_1 I_H \le A_t \le M_1 I_H \quad \text{and} \quad m_2 I_H \le B_t \le M_2 I_H$$

for all $t \in T$ for some $0 < m_1 < M_1$ and $0 < m_2 < M_2$. Then for any $\lambda \ge 0$, $s \ge 1$, $p \ge 1$ and any continuous field $(x_t)_{t\in T}$ of vectors in H such that $\int_T \|x_t\|^2 \mathrm{d}\mu(t) = 1$ we have

$$\begin{aligned} &\left(\int_T \langle A_t^p x_t, x_t \rangle \mathrm{d}\mu(t) \right)^{1/p} \left(\int_T \langle B_t^q x_t, x_t \rangle \mathrm{d}\mu(t) \right)^{1/q} - \lambda \int_T \langle B_t^q \#_{1/s} A_t^p x_t, x_t \rangle \mathrm{d}\mu(t) \\ &\qquad \le C\left(\lambda, \frac{m_1^{p/s}}{M_2^{q/s}}, \frac{M_1^{p/s}}{m_2^{q/s}}, s, p \right) M_2^q, \end{aligned} \tag{9.5}$$

where the constant C is defined in Theorem 9.3 and $1/p + 1/q = 1$.

Proof. By using Theorem 9.3 we obtain for any $\lambda \ge 0$, for any continuous field $(C_t)_{t\in T}$ of positive operators with $mI_H \le C_t \le MI_H$ and a square integrable continuous field $(y_t)_{t\in T}$ of vectors in H the inequality

$$\begin{aligned} &\left(\int_T \langle C_t^s y_t, y_t \rangle \,\mathrm{d}\mu(t) \right)^{1/p} \left(\int_T \langle y_t, y_t \rangle \,\mathrm{d}\mu(t) \right)^{1/q} - \lambda \int_T \langle C_t y_t, y_t \rangle \,\mathrm{d}\mu(t) \\ &\qquad \le C(\lambda, m, M, s, p) \int_T \langle y_t, y_t \rangle \,\mathrm{d}\mu(t). \end{aligned} \tag{9.6}$$

Set now $C_t = \left(B_t^{-q/2} A_t^p B_t^{-q/2} \right)^{1/s}$ and $y_t = B_t^{q/2} x_t$ for $t \in T$ in (9.6) and observe that

$$\frac{m_1^{p/s}}{M_2^{q/s}} I_H \le \left(B_t^{-q/2} A_t^p B_t^{-q/2} \right)^{1/s} \le \frac{M_1^{p/s}}{m_2^{q/s}} I_H.$$

By using the definition of the $1/s-$geometric mean and rearranging (9.6) we obtain

$$\begin{aligned} &\left(\int_T \langle A_t^p x_t, x_t \rangle \,\mathrm{d}\mu(t) \right)^{1/p} \left(\int_T \langle B_t^q x_t, x_t \rangle \,\mathrm{d}\mu(t) \right)^{1/q} - \lambda \int_T \langle B_t^q \#_{1/s} A_t^p x_t, x_t \rangle \,\mathrm{d}\mu(t) \\ &\le C\left(\lambda, \tfrac{m_1^{p/s}}{M_2^{q/s}}, \tfrac{M_1^{p/s}}{m_2^{q/s}}, s, p \right) \int_T \langle B_t^q x_t, x_t \rangle \,\mathrm{d}\mu(t) \le C\left(\lambda, \frac{m_1^{p/s}}{M_2^{q/s}}, \frac{M_1^{p/s}}{m_2^{q/s}}, s, p \right) M_2^q, \end{aligned}$$

which gives (9.5). □

In the present context we may obtain results of the Li-Mathias type by using Theorem 9.2 and the following result which is a simple consequence of Theorem 9.1.

Theorem 9.4 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. We assume the spectra are in $[m,M]$. Let furthermore $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let $f,g : [m,M] \to \mathbb{R}$ and $F : U\times V \to \mathbb{R}$ be functions such that $f([m,M]) \subset U$, $g([m,M]) \subset V$ and F is bounded. If F is operator monotone in the first variable and f is operator convex in the interval $[m,M]$, then*

$$F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right] \geq \inf_{m\le z\le M} F[f(z),g(z)]\,\mathbf{1}. \tag{9.7}$$

In the dual case (when f is operator concave) the opposite inequality holds in (9.7) *with* sup *instead of* inf.

We also give generalizations of some results from [46].

Theorem 9.5 *Let f be a convex function on $[0,\infty)$ and let $\|\cdot\|$ be a normalized unitarily invariant norm on $B(H)$ for some finite dimensional Hilbert space H. Let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : B(H) \to B(K)$, where K is a Hilbert space, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. Then for every continuous field of positive operators $(A_t)_{t\in T}$ we have*

$$\int_T \Phi_t(f(A_t))\,\mathrm{d}\mu(t) \le f(0)I_K + \int_T \frac{f(\|A_t\|)-f(0)}{\|A_t\|}\Phi_t(A_t)\,\mathrm{d}\mu(t).$$

Especially, for $f(0)\le 0$, the inequality

$$\int_T \Phi_t(f(A_t))\,\mathrm{d}\mu(t) \le \int_T \frac{f(\|A_t\|)}{\|A_t\|}\Phi_t(A_t)\,\mathrm{d}\mu(t). \tag{9.8}$$

is valid.

Proof. Since f is a convex function, $f(x) \le \frac{M-x}{M-m}f(m) + \frac{x-m}{M-m}f(M)$ for every $x\in [m,M]$, $m\le M$. Since $\|\cdot\|$ is normalized and unitarily invariant, we have $0 < A_t \le \|A_t\|I_H$ and thus

$$f(A_t) \le \frac{\|A_t\|I_H - A_t}{\|A_t\|}f(0) + \frac{A_t}{\|A_t\|}f(\|A_t\|)$$

for every $t\in T$. Applying positive linear mappings and integrating we obtain

$$\int_T \Phi_t(f(A_t))\,\mathrm{d}\mu(t) \le f(0)\left[I_K - \int_T \frac{\Phi_t(A_t)}{\|A_t\|}\,\mathrm{d}\mu(t)\right] + \int_T \frac{f(\|A_t\|)}{\|A_t\|}\Phi_t(A_t)\,\mathrm{d}\mu(t) \tag{9.9}$$

or

$$\int_T \Phi_t(f(A_t))\,\mathrm{d}\mu(t) \le f(0)I_H + \int_T \frac{f(\|A_t\|)-f(0)}{\|A_t\|}\Phi_t(A_t)\,\mathrm{d}\mu(t).$$

Note that since $\int_T \frac{\Phi_t(A_t)}{\|A_t\|}\mathrm{d}\mu(t) \le \int_T \frac{\|A_t\|\Phi_t(I_H)}{\|A_t\|}\mathrm{d}\mu(t) = I_K$, we obtain, for $f(0)\le 0$, inequality (9.8) from (9.9). □

Remark 9.1 *Setting $T=\{1\}$ the inequality* (9.8) *gives*

$$\Phi(f(A)) \leq \frac{f(\|A\|)}{\|A\|}\Phi(A).$$

Furthermore, setting that Φ is the identical mapping, we get the inequality $f(\|A\|) \geq \|f(A)\|$ obtained in [46] under the assumption that f is a non-negative convex function with $f(0)=0$.

Related inequalities may be obtained by using subdifferentials. If $f:\mathbb{R}\to\mathbb{R}$ is a convex function and $[m,M]$ is a closed bounded real interval, then a subdifferential function of f on $[m,M]$ is any function $l:[m,M]\to\mathbb{R}$ such that

$$l(x) \in [f'_-(x), f'_+(x)], \qquad x \in (m,M),$$

where f'_- and f'_+ are the one-sides derivatives of f and $l(m)=f'_+(m)$ and $l(M)=f'_-(M)$. Since this functions are Borel measurable, we may use the Borel functional calculus. Subdifferential function for concave functions is defined in analogous way.

Theorem 9.6 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t:\mathscr{A}\to\mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. If $f:[m,M]\to\mathbb{R}$ is a convex function then*

$$\begin{aligned} f(y)\mathbf{1}+l(y)\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)-y\mathbf{1}\right) &\leq \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \\ \leq f(x)\mathbf{1}-x\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t)&+\int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t) \end{aligned} \tag{9.10}$$

for every $x,y\in[m,M]$, where l is the subdifferential of f on $[m,M]$. In the dual case (f is concave) the opposite inequality holds in (9.10)*.*

Proof. Since f is convex we have $f(x)\geq f(y)+l(y)(x-y)$ for every $x,y\in[m,M]$. By using the functional calculus it then follows that $f(x_t)\geq f(y)\mathbf{1}+l(y)(x_t-y\mathbf{1})$ for $t\in T$. Applying the positive linear mappings Φ_t and integrating, LHS of (9.10) follows. The RHS of (9.10) follows similarly by using the functional calculus in the variable y. □

Numerous inequalities can be obtained from (9.10). For example, LHS of (9.10) may be used to obtain an estimation from below in the sense of Theorem 9.2. Namely, the following theorem holds.

Theorem 9.7 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t:\mathscr{A}\to\mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let $f,g:[m,M]\to\mathbb{R}$ and $F:U\times V\to\mathbb{R}$ be functions such that $f([m,M])\subset U$, $g([m,M])\subset V$, F is bounded, f*

is convex and $f(y)+l(y)(t-y)\in U$ *for every* $y,t\in[m,M]$ *where* l *is the subdifferential of* f *on* $[m,M]$. *If* F *is operator monotone in the first variable, then*

$$F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t),g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right]\geq \inf_{m\leq z\leq M} F\left[f(y)+l(y)(z-y),g(z)\right]\mathbf{1} \tag{9.11}$$

for every $y\in[m,M]$. *In the dual case (when* f *is concave) the opposite inequality holds in* (9.11) *with* sup *instead of* inf.

Using LHS of (9.10) we can give generalizations of some dual results from [46].

Theorem 9.8 *Let* $(x_t)_{t\in T}$ *be a bounded continuous field of positive elements in a unital* C^**-algebra* $\mathscr{A}$ *defined on a locally compact Hausdorff space* T *equipped with a bounded Radon measure* μ, *and let* $(\Phi_t)_{t\in T}$ *be a unital field of positive linear mappings* $\Phi_t\colon \mathscr{A}\to\mathscr{B}$ *from* $\mathscr{A}$ *to another unital* C^**-algebra* $\mathscr{B}$ *acting on a finite dimensional Hilbert space* K. *Let* $\|\cdot\|$ *be unitarily invariant norm on* $B(K)$ *and let* $f\colon[0,\infty)\to\mathbb{R}$ *be an increasing function.*

(1) *If* $\|\mathbf{1}\|=1$ *and* f *is convex with* $f(0)\leq 0$ *then*

$$f\left(\|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|\right)\leq \|\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\|. \tag{9.12}$$

(2) *If* $\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\leq \|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|\mathbf{1}$ *and* f *is concave then*

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\leq f\left(\|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|\right)\mathbf{1}. \tag{9.13}$$

Proof. Since $f(0)\leq 0$ and f is increasing we have $l(y)y-f(y)\geq 0$ and $l(y)\geq 0$. From (9.10) and the triangle inequality we have

$$l(y)\|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|\leq \|\int_T \Phi_t(f(x_t))\|+(l(y)y-f(y)).$$

Now (9.12) follows by setting $y=\|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|$. Inequality (9.13) follows immediately from the assumptions and from the dual case of LHS in (9.10) by setting $y=\|\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\|$. □

Finally, to illustrate how RHS of (9.10) works, we set

$$x=\frac{\|\int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t)\|}{\|\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t)\|}$$

and obtain a Slater type inequality

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\leq f\left(\frac{\|\int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t)\|}{\|\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t)\|}\right)\mathbf{1}$$

under the condition

$$\frac{\int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t)}{\|\int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t)\|}\leq\frac{\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t)}{\|\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t)\|}.$$

9.2 Non-unital fields of positive linear mappings

In this section we observe one type of non-unital fields of positive linear mappings, which is a generalization of the results obtained in previous section.

Let T be a locally compact Hausdorff space with a bounded Radon measure μ. For convenience, we use the abbreviation $P_k[\mathscr{A},\mathscr{B}]$ for the set of all fields $(\Phi_t)_{t\in T}$ of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from a unital C^*-algebra $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$, such that the field $t \to \Phi_t(\mathbf{1})$ is integrable with $\int_T \Phi_t(\mathbf{1})\,\mathrm{d}\mu(t) = k\mathbf{1}$ for some positive scalar k.

Let Φ be a normalized positive linear mapping on $B(H)$ and f an operator convex function on an interval J. We recall that Jensen's inequality asserts that $f(\Phi(A)) \le \Phi(f(A))$ holds for every self-adjoint operator A on a Hilbert space H whose the spectrum is contained in J. But if $\Phi(\mathbf{1}) = k\mathbf{1}$, for some positive scalar k, then $f(\Phi(A)) \not\le \Phi(f(A))$. Really, let $\Phi : \mathbf{M}_2(\mathbf{M}_2(\mathbb{C})) \to \mathbf{M}_2(\mathbf{M}_2(\mathbb{C}))$ be a positive linear mapping defined by

$$\Phi\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A+B & 0 \\ 0 & A+B \end{pmatrix}$$

for $A, B \in \mathbf{M}_2(\mathbb{C})$. Then $\Phi(I) = 2I$. Let $f(t) = t^2$. Then f is the operator convex function. Put

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

We have

$$\begin{aligned} & f\left(\Phi\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) - \Phi\left(f\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) \\ &= \begin{pmatrix} 10 & 5 & 0 & 0 \\ 5 & 5 & 0 & 0 \\ 0 & 0 & 10 & 5 \\ 0 & 0 & 5 & 5 \end{pmatrix} - \begin{pmatrix} 6 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 6 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 3 & 2 \end{pmatrix} \not\ge 0. \end{aligned}$$

But, the following theorem is equivalent to Theorem 9.1.

Theorem 9.9 *Let $f : J \to \mathbb{R}$ be an operator convex function defined on an interval J, and let $\mathscr{A}$ and $\mathscr{B}$ be unital C^*-algebras. Let T be a locally compact Hausdorff space with a bounded Radon measure μ. If a field $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$, then the inequality*

$$f\left(\frac{1}{k}\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le \frac{1}{k}\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \tag{9.14}$$

holds for every bounded continuous field $(x_t)_{t\in T}$ of self-adjoint elements in $\mathscr{A}$ with spectra contained in J. In the dual case (when f is operator concave) the opposite inequality holds in (9.14).

In the present context we may obtain results of the Li-Mathias type, which is a generalization of Theorem 9.2.

Theorem 9.10 *Let $\mathscr{A}$ and $\mathscr{B}$ be unital C^*-algebras. Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in $\mathscr{A}$ with spectra in $[m,M]$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. Furthermore, let $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$ and $f:[m,M]\to\mathbb{R}$, $g:[km,kM]\to\mathbb{R}$ and $F:U\times V\to\mathbb{R}$ be functions such that $(kf)([m,M])\subset U$, $g([km,kM])\subset V$ and F is bounded. Let $\{conx.\}$ (resp. $\{conc.\}$) denotes the set of operator convex (resp. operator concave) functions defined on $[m,M]$. If F is operator monotone in the first variable, then*

$$\begin{aligned}\inf_{km\le z\le kM} F\left[k\cdot h_1\left(\frac{1}{k}z\right),g(z)\right]\mathbf{1} &\le F\left[\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right] \\ &\le \sup_{km\le z\le kM} F\left[k\cdot h_2\left(\frac{1}{k}z\right),g(z)\right]\mathbf{1}\end{aligned}\tag{9.15}$$

holds for every $h_1\in\{conx.\}$, $h_1\le f$ and $h_2\in\{conc.\}$, $h_2\ge f$.

Proof. We prove only RHS of (9.15). Let h_2 be operator concave function on $[m,M]$ such that $f(z)\le h_2(z)$ for every $z\in[m,M]$. Thus, by using the functional calculus, $f(x_t)\le h_2(x_t)$ for every $t\in T$. Applying the positive linear mappings Φ_t and integrating, we obtain

$$\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t) \le \int_T \Phi_t\left(h_2(x_t)\right)\mathrm{d}\mu(t).$$

Furthermore, by using Theorem 9.9, we have

$$\int_T \Phi_t\left(h_2(x_t)\right)\mathrm{d}\mu(t) \le k\cdot h_2\left(\frac{1}{k}\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)$$

and it follows that $\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t)\le k\cdot h_2\left(\frac{1}{k}\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)$. Since $m\Phi_t(\mathbf{1})\le\Phi_t(x_t)$ $\le M\Phi_t(\mathbf{1})$, it follows that $km\mathbf{1}\le\int_T\Phi_t(x_t)\mathrm{d}\mu(t)\le kM\mathbf{1}$. Using operator monotonicity of $F(\cdot,v)$, we obtain

$$\begin{aligned}&F\left[\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right] \\ &\le F\left[k\cdot h_2\left(\frac{1}{k}\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right] \\ &\le \sup_{km\le z\le kM} F\left[k\cdot h_2\left(\frac{1}{k}z\right),g(z)\right]\mathbf{1}.\end{aligned}$$

□

Remark 9.2 *Putting $F(u,v) = u - v$ and $F(u,v) = u^{-1/2}vu^{-1/2}$ in Theorem 9.10, we obtain that*

$$\inf_{km\le z\le kM}\left\{k\cdot h_1\left(\frac{1}{k}z\right)-g(z)\right\}\mathbf{1}\le \int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t)-g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)$$
$$\le \sup_{km\le z\le kM}\left\{k\cdot h_2\left(\frac{1}{k}z\right)-g(z)\right\}\mathbf{1}$$

holds and if in addition $g(t) > 0$ for all $t \in [m,M]$ then

$$\inf_{km\le z\le kM}\frac{k\cdot h_1\left(\frac{1}{k}z\right)}{g(z)}\,g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\le \int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t)$$
$$\le \sup_{km\le z\le kM}\frac{k\cdot h_2\left(\frac{1}{k}z\right)}{g(z)}\,g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)$$

holds for every $h_1 \in \{conx.\}$, $h_1 \le f$ and $h_2 \in \{conc.\}$, $h_2 \ge f$.

Applying RHS of (9.15) for a convex function f (or LHS of (9.15) for a concave function f) we obtain the following theorem (compare with Theorem 9.2).

Theorem 9.11 *Let $(x_t)_{t\in T}$ and $(\Phi_t)_{t\in T}$ be as in Theorem 9.10. Let $f : [m,M] \to \mathbb{R}$, $g : [km,kM] \to \mathbb{R}$ and $F : U\times V \to \mathbb{R}$ be functions such that $(kf)([m,M]) \subset U$, $g([km,kM]) \subset V$ and F is bounded. If F is operator monotone in the first variable and f is convex in the interval $[m,M]$, then*

$$F\left[\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t), g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)\right]\le \sup_{km\le z\le kM} F\left[\alpha_f z+\beta_f k, g(z)\right]\mathbf{1}. \qquad (9.16)$$

In the dual case (when f is concave) the opposite inequality holds in (9.16) *with* inf *instead of* sup.

Proof. We prove only the convex case. For convex f the inequality $f(z) \le \alpha_f z + \beta_f$ holds for every $z \in [m,M]$. Thus, by putting $h_2(z) = \alpha_f z + \beta_f$ in RHS of (9.15) we obtain (9.16). □

Numerous applications of the previous theorem can be given. Namely, applying Theorem 9.11 for the function $F(u,v) = u - \lambda v$, we obtain the following result.

Corollary 9.2 *Let $(x_t)_{t\in T}$ and $(\Phi_t)_{t\in T}$ be as in Theorem 9.10. If $f : [m,M] \to \mathbb{R}$ is convex in the interval $[m,M]$ and $g : [km,kM] \to \mathbb{R}$, then for any $\lambda \in \mathbb{R}$*

$$\int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t)\le \lambda\, g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)+C\mathbf{1}, \qquad (9.17)$$

where

$$C=\sup_{km\le z\le kM}\left\{\alpha_f z+\beta_f k-\lambda g(z)\right\}.$$

If furthermore λg is strictly convex differentiable, then the constant $C \equiv C(m,M,f,g,k,\lambda)$ can be written more precisely as

$$C = \alpha_f z_0 + \beta_f k - \lambda g(z_0),$$

where

$$z_0 = \begin{cases} g'^{-1}(\alpha_f/\lambda) & \text{if } \lambda g'(km) \le \alpha_f \le \lambda g'(kM), \\ km & \text{if } \lambda g'(km) \ge \alpha_f, \\ kM & \text{if } \lambda g'(kM) \le \alpha_f. \end{cases}$$

In the dual case (when f is concave and λg is strictly concave differentiable) the opposite inequality holds in (9.17) *with* min *instead of* max *with the opposite condition while determining z_0.*

Remark 9.3 *We assume that $(x_t)_{t\in T}$ and $(\Phi_t)_{t\in T}$ are as in Theorem 9.10. If $f : [m,M] \to \mathbb{R}$ is convex and $\lambda g : [km,kM] \to \mathbb{R}$ is strictly concave differentiable, then the constant $C \equiv C(m,M,f,g,k,\lambda)$ in* (9.17) *can be written more precisely as*

$$C = \begin{cases} \alpha_f kM + \beta_f k - \lambda g(kM) & \text{if } \alpha_f - \lambda \alpha_{g,k} \ge 0, \\ \alpha_f km + \beta_f k - \lambda g(km) & \text{if } \alpha_f - \lambda \alpha_{g,k} \le 0, \end{cases}$$

where

$$\alpha_{g,k} = \frac{g(kM) - g(km)}{kM - km}.$$

Setting $\Phi_t(A_t) = \langle A_t x_t, x_t\rangle$ for $x_t \in H$ and $t \in T$ in Corollary 9.2 and Remark 9.3 give a generalization of all results from [96, Section 2.4]. For example, we obtain the following two corollaries.

Corollary 9.3 *Let $(A_t)_{t\in T}$ be a continuous field of positive operators on a Hilbert space H defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. We assume the spectra are in $[m,M]$ for some $0 < m < M$. Let furthermore $(x_t)_{t\in T}$ be a continuous field of vectors in H such that $\int_T \|x_t\|^2 \mathrm{d}\mu(t) = k$ for some scalar $k > 0$. Then for any real λ, q, p*

$$\int_T \langle A_t^p x_t, x_t\rangle \mathrm{d}\mu(t) - \lambda \left(\int_T \langle A_t x_t, x_t\rangle \mathrm{d}\mu(t) \right)^q \le C, \tag{9.18}$$

where the constant $C \equiv C(\lambda, m, M, p, q, k)$ is

$$C = \begin{cases} (q-1)\lambda \left(\dfrac{\alpha_p}{\lambda q}\right)^{q/(q-1)} + \beta_p k & \text{if } \lambda q m^{q-1} \le \dfrac{\alpha_p}{k^{q-1}} \le \lambda q M^{q-1}, \\ kM^p - \lambda (kM)^q & \text{if } \dfrac{\alpha_p}{k^{q-1}} \ge \lambda q M^{q-1}, \\ km^p - \lambda (km)^q & \text{if } \dfrac{\alpha_p}{k^{q-1}} \le \lambda q m^{q-1}, \end{cases} \tag{9.19}$$

in the case $\lambda q(q-1) > 0$ and $p \in \mathbb{R} \setminus (0,1)$
or

$$C = \begin{cases} kM^p - \lambda (kM)^q & \text{if } \alpha_p - \lambda k^{q-1} \alpha_q \ge 0, \\ km^p - \lambda (km)^q & \text{if } \alpha_p - \lambda k^{q-1} \alpha_q \le 0, \end{cases} \tag{9.20}$$

in the case $\lambda q(q-1)<0$ *and* $p\in\mathbb{R}\setminus(0,1)$. *In the dual case:* $\lambda q(q-1)<0$ *and* $p\in(0,1)$ *the opposite inequality holds in* (9.18) *with the opposite condition while determining the constant* C *in* (9.19). *But in the dual case:* $\lambda q(q-1)>0$ *and* $p\in(0,1)$ *the opposite inequality holds in* (9.18) *with the opposite condition while determining the constant* C *in* (9.20).

Constants α_p *and* β_p *in the terms above are the constants* α_f *and* β_f *associated with the function* $f(z)=z^p$.

Corollary 9.4 *Let* $(A_t)_{t\in T}$ *and* $(x_t)_{t\in T}$ *be as in Corollary 9.3. Then for any real number* $r\neq 0$ *we have*

$$\int_T\langle\exp(rA_t)x_t,x_t\rangle d\mu(t)-\exp\left(r\int_T\langle A_tx_t,x_t\rangle d\mu(t)\right)\leq C_1, \tag{9.21}$$

$$\int_T\langle\exp(rA_t)x_t,x_t\rangle d\mu(t)\leq C_2\,\exp\left(r\int_T\langle A_tx_t,x_t\rangle d\mu(t)\right), \tag{9.22}$$

where constants $C_1\equiv C_1(r,m,M,k)$ *and* $C_2\equiv C_2(r,m,M,k)$ *are*

$$C_1=\begin{cases}\dfrac{\alpha}{r}\log\left(\dfrac{\alpha}{r\mathrm{e}}\right)+k\beta & \textit{if}\ \ r\mathrm{e}^{rkm}\leq\alpha\leq r\mathrm{e}^{rkM},\\ kM\alpha+k\beta-\mathrm{e}^{rkM} & \textit{if}\ \ r\mathrm{e}^{rkM}\leq\alpha,\\ km\alpha+k\beta-\mathrm{e}^{rkm} & \textit{if}\ \ r\mathrm{e}^{rkm}\geq\alpha,\end{cases}$$

$$C_2=\begin{cases}\dfrac{\alpha}{r\mathrm{e}}\mathrm{e}^{kr\beta/\alpha} & \textit{if}\ \ kr\mathrm{e}^{rm}\leq\alpha\leq kr\mathrm{e}^{rM},\\ k\mathrm{e}^{(1-k)rm} & \textit{if}\ \ kr\mathrm{e}^{rm}\geq\alpha,\\ k\mathrm{e}^{(1-k)rM} & \textit{if}\ \ kr\mathrm{e}^{rM}\leq\alpha.\end{cases}$$

Constants α *and* β *in the terms above are the constants* α_f *and* β_f *associated with the function* $f(z)=\mathrm{e}^{rz}$.

By using subdifferentials we can give an estimation from below in the sense of Theorem 9.11 (compare with Theorem 9.6). It follows from Theorem 9.10 applying LHS of (9.15) for a convex function f (or RHS of (9.15) for a concave function f).

Theorem 9.12 *Let* $(x_t)_{t\in T}$ *be a bounded continuous field of self-adjoint elements in a unital* C^**-algebra* $\mathscr{A}$ *with spectra in* $[m,M]$ *defined on a locally compact Hausdorff space* T *equipped with a bounded Radon measure* μ. *Let* $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$. *Furthermore, let* $f:[m,M]\to\mathbb{R}$, $g:[km,kM]\to\mathbb{R}$ *and* $F:U\times V\to\mathbb{R}$ *be functions such that* $(kf)([m,M])\subset U$, $g([km,kM])\subset V$, F *is bounded and* $f(y)+l(y)(t-y)\in U$ *for every* $y,t\in[m,M]$ *where* l *is the subdifferential of* f. *If* F *is operator monotone in the first variable and* f *is convex on* $[m,M]$, *then*

$$\begin{aligned}&F\left[\int_T\Phi_t\left(f(x_t)\right)d\mu(t),g\left(\int_T\Phi_t(x_t)\,d\mu(t)\right)\right]\\&\geq\inf_{km\leq z\leq kM}F\left[f(y)k+l(y)(z-yk),g(z)\right]\mathbf{1}\end{aligned} \tag{9.23}$$

holds for every $y \in [m,M]$. *In the dual case (when* f *is concave) the opposite inequality holds in* (9.23) *with* sup *instead of* inf.

Proof. We prove only the convex case. Since f is convex we have $f(z) \geq f(y) + l(y)(z-y)$ for every $z,y \in [m,M]$. Thus, by putting $h_1(z) = f(y) + l(y)(z-y)$ in LHS of (9.15) we obtain (9.23). □

Though $f(z) = \log z$ is operator concave, Jensen's inequality $\Phi(f(x)) \leq f(\Phi(x))$ does not hold in the case of non-unital Φ. However, as applications of Corollary 9.2 and Theorem 9.12, we obtain the following corollary.

Corollary 9.5 *Let* $(x_t)_{t\in T}$ *and* $(\Phi_t)_{t\in T}$ *be as in Theorem 9.12 for* $0 < m < M$. *Then*

$$C_1\mathbf{1} \leq \int_T \Phi_t\left(\log(x_t)\right) \mathrm{d}\mu(t) - \log\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right) \leq C_2\mathbf{1}, \tag{9.24}$$

where constants $C_1 \equiv C_1(m,M,k)$ *and* $C_2 \equiv C_2(m,M,k)$ *are*

$$C_1 = \begin{cases} k\beta + \log\left(\mathrm{e}/L(m,M)\right) & \text{if } km \leq L(m,M) \leq kM, \\ \log\left(M^{k-1}/k\right) & \text{if } kM \leq L(m,M), \\ \log\left(m^{k-1}/k\right) & \text{if } km \geq L(m,M), \end{cases}$$

$$C_2 = \begin{cases} \log\left(\dfrac{L(m,M)^k k^{k-1}}{\mathrm{e}^k m}\right) + \dfrac{m}{L(m,M)} & \text{if } m \leq kL(m,M) \leq M \\ \log\left(M^{k-1}/k\right) & \text{if } kL(m,M) \geq M, \\ \log\left(m^{k-1}/k\right) & \text{if } kL(m,M) \leq m, \end{cases}$$

and the logarithmic mean $L(m,M)$, β *is the constant* β_f *associated with the function* $f(z) = \log z$.

Proof. We set $f(z) \equiv g(z) = \log z$ in Corollary 9.2. Then we obtain the lower bound C_1 when we determine $\min\limits_{km \leq z \leq kM} (\alpha z + k\beta - \log z)$.

Next, we shall obtain the upper bound C_2. We set $F(u,v) = u - v$ and $f(z) \equiv g(z) = \log z$ in Theorem 9.12. We obtain

$$\begin{aligned} &\int_T \Phi_t\left(\log(x_t)\right) \mathrm{d}\mu(t) - \log\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right) \\ &\leq \max\left\{\log\left(\frac{y^k}{\mathrm{e}^k km}\right) + \frac{km}{y}, \log\left(\frac{y^k}{\mathrm{e}^k kM}\right) + \frac{kM}{y}\right\}\mathbf{1} \end{aligned}$$

for every $y \in [m,M]$, since $h(z) = k\log y + \frac{1}{y}(z - ky) - \log z$ is a convex function and it implies that

$$\max_{km \leq z \leq kM} h(z) = \max\{h(km), h(kM)\}.$$

Now, if $m \le kL(m,M) \le M$, then we choose $y = kL(m,M)$. In this case we have $h(km) = h(kM)$. But, if $m \ge kL(m,M)$, then it follows $0 < k \le 1$, which implies that $\max\{h(km), h(kM)\} = h(km)$ for every $y \in [m,M]$. In this case we choose $y = m$, since $h(y) = \log\left(\frac{y^k}{\mathrm{e}^k km}\right) + \frac{km}{y}$ is an increasing function in $[m,M]$. If $M \le kL(m,M)$, then the proof is similar to above. □

By using subdifferentials, we also give generalizations of Theorem 9.6.

Theorem 9.13 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$. If the field $t \to \Phi_t(\mathbf{1})$ is integrable with $\int_T \Phi_t(\mathbf{1})\,\mathrm{d}\mu(t) = k\mathbf{1}$ for some positive scalar k and $f : [m,M] \to \mathbb{R}$ is a convex function then*

$$\begin{aligned} f(y)k\mathbf{1} + l(y)\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t) - yk\mathbf{1}\right) &\le \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \\ &\le f(x)k\mathbf{1} - x\int_T \Phi_t(l(x_t))\,\mathrm{d}\mu(t) + \int_T \Phi_t(l(x_t)x_t)\,\mathrm{d}\mu(t) \end{aligned} \tag{9.25}$$

for every $x,y \in [m,M]$, where l is the subdifferential of f. In the dual case (f is concave) the reverse inequality is valid in (9.25).

Remark 9.4 *In the case $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$ we may obtain analogues results as in Theorems 9.5 and 9.8. The interested reader may be read the details in [202].*

9.3 Ratio type inequalities with power functions

In this section we consider the ratio type ordering among the following power functions of operators:

$$F_r(\mathbf{x},\mathbf{\Phi}) := \left(\int_T \Phi_t\left(x_t^r\right)\mathrm{d}\mu(t)\right)^{1/r}, \qquad r \in \mathbb{R}\backslash\{0\} \tag{9.26}$$

under these conditions: $(x_t)_{t\in T}$ is a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0 < m < M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$

As an application, we consider a generalization of the weighted power means of operators:

$$M_r(\mathbf{x},\mathbf{\Phi}) := \left(\int_T \frac{1}{k}\Phi_t\left(x_t^r\right)\mathrm{d}\mu(t)\right)^{1/r}, \qquad r \in \mathbb{R}\backslash\{0\} \tag{9.27}$$

under the same conditions as above.

We need some previous results given in the following three lemmas.

Lemma 9.1 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0<m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$.*

If $0<p\le 1$, then

$$\int_T \Phi_t\left(x_t^p\right) \mathrm{d}\mu(t) \le k^{1-p}\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p. \tag{9.28}$$

If $-1\le p<0$ or $1\le p\le 2$, then the reverse inequality is valid in (9.28).

Proof. We obtain this lemma by applying Theorem 9.9 for the function $f(z)=z^p$ and using the proposition that it is an operator concave function for $0<p\le 1$ and an operator convex one for $-1\le p<0$ and $1\le p\le 2$. □

Lemma 9.2 *Assume that the conditions of Lemma 9.1 hold.*
If $0<p\le 1$, then

$$k^{1-p}K(m,M,p)\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p \le \int_T \Phi_t\left(x_t^p\right) \mathrm{d}\mu(t) \le k^{1-p}\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p, \tag{9.29}$$

if $-1\le p<0$ or $1\le p\le 2$, then

$$k^{1-p}\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p \le \int_T \Phi_t\left(x_t^p\right) \mathrm{d}\mu(t) \le k^{1-p}K(m,M,p)\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p, \tag{9.30}$$

if $p<-1$ or $p>2$, then

$$\begin{aligned} k^{1-p}K(m,M,p)^{-1}\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p &\le \int_T \Phi_t\left(x_t^p\right) \mathrm{d}\mu(t) \\ &\le k^{1-p}K(m,M,p)\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p, \end{aligned} \tag{9.31}$$

where $K(m,M,p)$ is the generalized Kantorovich constant by (2.29).

Proof. We obtain this lemma by applying Corollary 9.2 for the function $f(z)\equiv g(z)=z^p$ and choosing λ such that $C=0$. □

We shall need some properties of the generalized Specht ratio $\Delta(h,r,s)$ (see (8.35) and Figure 9.1).

Lemma 9.3 *Let $M>m>0$, $r\in\mathbb{R}$ and*

$$\Delta(h,r,1)=\frac{r(h-h^r)}{(1-r)(h^r-1)}\left(\frac{h^r-h}{(r-1)(h-1)}\right)^{-1/r}, \qquad h=\frac{M}{m}.$$

(i) *A function $\Delta(r)\equiv\Delta(h,r,1)$ is strictly decreasing for all $r\in\mathbb{R}$,*

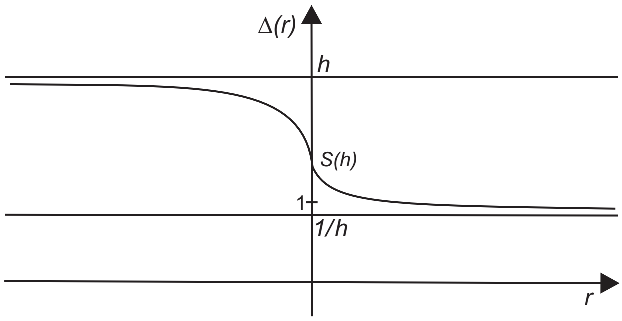


Figure 9.1: Function $\Delta(r) \equiv \Delta(h,r,1)$

(ii) $\lim_{r\to 1} \Delta(h,r,1) = 1$ *and* $\lim_{r\to 0} \Delta(h,r,1) = S(h)$,
where the Specht ratio $S(h)$ *is defined by* (2.35)*.*

(iii) $\lim_{r\to\infty} \Delta(h,r,1) = 1/h$ *and* $\lim_{r\to -\infty} \Delta(h,r,1) = h$*.*

Proof.

(i) We write $\Delta(r) = \Delta_1(r)\cdot\Delta_2(r)$, where

$$\Delta_1(r) = \frac{r(h^r-h)}{(r-1)(h^r-1)}, \quad \Delta_2(r) = \left(\frac{h^r-h}{(r-1)(h-1)}\right)^{-1/r}. \tag{9.32}$$

By using differential calculus we shall prove that a function Δ_1 is strictly decreasing for all $r \neq 0,1$. We have

$$\begin{aligned} \tfrac{d}{dr}\Delta_1(r) &= \tfrac{-1}{(r-1)^2(h^r-1)^2}\left((h^r-1)(h^r-h)-(r-1)rh^r(h-1)\log h\right) \\ &= -\tfrac{h^r(h-1)\log h}{(r-1)^2(h^r-1)^2} f(r), \end{aligned} \tag{9.33}$$

where $f(r) = \frac{(h^r-1)(h^r-h)}{h^r(h-1)\log h} - (r-1)r$. Stationary points of the function f are 0, 0.5, 1 and it is a strictly decreasing function on $(-\infty,0)\cup(0.5,1)$ and strictly increasing on $(0,0.5)\cup(1,\infty)$. Also, $f(0) = f(1) = 0$. So, $f(r) > 0$ for all $r \neq 0,1$. (More exactly, $f'''(r) = \frac{\log^2 h}{h-1}\left(h^r - h^{1-r}\right)$ imply that the function f'' is strictly increasing on $(0.5,\infty)$ and strictly decreasing on $(-\infty,0.5)$. It follows that $f'(r) < 0$ for $r \in (-\infty,0)\cup(0.5,1)$ and $f'(r) > 0$ for $r \in (0,0.5)\cup(1,\infty)$.) Now, using (9.33) we have that $\frac{d}{dr}\Delta_1(r) < 0$ for all $r \neq 0,1$ and it follows that Δ_1 is strictly decreasing function.

Further, in the case of a function Δ_2 in (9.32), we obtain

$$\begin{aligned} \tfrac{d}{dr}\Delta_2(r) &= \tfrac{-1}{(r-1)r^2(h^r-h)}\left(\tfrac{h^r-h}{(r-1)(h-1)}\right)^{-1/r} \\ &\times \left[r(r-1)h^r\log h - r(h^r-h) + (r-1)(h^r-h)\log\left(\tfrac{(r-1)(h-1)}{h^r-h}\right)\right]. \end{aligned}$$

By using differential calculus we can prove that a function

$$r \mapsto r(r-1)h^r \log h - r(h^r - h) + (r-1)(h^r - h)\log\left(\frac{(r-1)(h-1)}{h^r - h}\right)$$

is positive for all $r \neq 0,1$. So $\frac{d}{dr}\Delta_2(r) < 0$ for all $r \neq 0,1$ and it follows that Δ_2 is strictly decreasing function.

(ii) Using the definition of the generalized Specht ratio (8.35), we have $\Delta(h,r,1) = K(h^r, 1/r)$ if $r \neq 0$. Now, we have $K(h,1) = 1$ by using Theorem 2.12 and $\lim_{r\to 0} K(h^r, 1/r) = S(h)$ by using Theorem 2.17.

(iii) We have by L'Hospital's theorem

$$\lim_{r\to\infty} \frac{\log((r-1)(h-1)/(h^r - h))}{r} = \lim_{r\to\infty}\left(\frac{1}{r-1} - \frac{h^r \log h}{h^r - h}\right) = -\log h.$$

So

$$\lim_{r\to\infty} \Delta(h,r,1) = \lim_{r\to\infty} \frac{r}{r-1} \cdot \frac{h^r - h}{h^r - 1} \cdot \left(\frac{(r-1)(h-1)}{h^r - h}\right)^{1/r} = \mathrm{e}^{-\log h} = 1/h.$$

Similarly, we obtain $\lim_{r\to -\infty} \Delta(h,r,1) = h$.

□

Now, we give the ratio type ordering among power functions.

Theorem 9.14 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0 < m < M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$. Let regions* (i) – (v)$_1$ *be as in Figure 9.2.*
If (r,s) in (i)*, then*

$$k^{\frac{s-r}{rs}}\, \Delta(h,r,s)^{-1}\, F_s(\mathbf{x},\mathbf{\Phi}) \le F_r(\mathbf{x},\mathbf{\Phi}) \le k^{\frac{s-r}{rs}}\, F_s(\mathbf{x},\mathbf{\Phi}),$$

if (r,s) in (ii) *or* (iii)*, then*

$$k^{\frac{s-r}{rs}}\, \Delta(h,r,s)^{-1}\, F_s(\mathbf{x},\mathbf{\Phi}) \le F_r(\mathbf{x},\mathbf{\Phi}) \le k^{\frac{s-r}{rs}}\, \Delta(h,r,s)\, F_s(\mathbf{x},\mathbf{\Phi}),$$

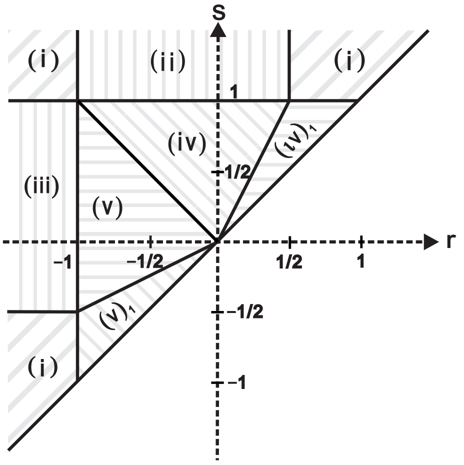
if (r,s) in (iv)*, then*

$$\begin{aligned} &k^{\frac{s-r}{rs}}\, \Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}\, F_s(\mathbf{x},\mathbf{\Phi}) \le F_r(\mathbf{x},\mathbf{\Phi}) \\ &\le k^{\frac{s-r}{rs}}\, \min\{\Delta(h,r,1), \Delta(h,s,1)\Delta(h,r,s)\}\, F_s(\mathbf{x},\mathbf{\Phi}), \end{aligned}$$

if (r,s) in (v) *or* (iv)$_1$ *or* (v)$_1$*, then*

$$k^{\frac{s-r}{rs}}\, \Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}\, F_s(\mathbf{x},\mathbf{\Phi}) \le F_r(\mathbf{x},\mathbf{\Phi}) \le k^{\frac{s-r}{rs}}\, \Delta(h,s,1)\, F_s(\mathbf{x},\mathbf{\Phi}),$$

where $\Delta(h,r,s)$, $rs \neq 0$ is defined by (8.35)*.*



(i)	r ≤ s, s ∉ (-1,1), r ∉ (-1,1) **or** 1/2 ≤ r ≤ 1 ≤ s **or** r ≤ -1 ≤ s ≤ -1/2,
(ii)	s ≥ 1, -1 < r < 1/2, r ≠ 0,
(iii)	r ≤ -1, -1/2 < s < 1, s ≠ 0,
(iv)	-s ≤ r < s/2, r ≠ 0, 0 < s ≤ 1,
(iv)₁	r ≤ s ≤ 2r, 0 < s ≤ 1,
(v)	r/2 < s ≤ -r, s ≠ 0, -1 ≤ r < 0,
(v)₁	2s ≤ r ≤ s, -1 ≤ r < 0.

Figure 9.2: Regions in the (r,s)-plain

Proof. This theorem follows from Lemma 9.2 by putting $p = s/r$ or $p = r/s$ and then using the Löwner-Heinz theorem, Theorem 4.3 and Lemma 9.3. We give the proof for the sake of completeness.

We put $p = s/r$ in Lemma 9.2 and replace x_t by x_t^r. Applying the Löwner-Heinz inequality if $s \geq 1$ or $s \leq -1$ and using that $K(m^r, M^r, s/r)^{1/s} = K(M^r, m^r, s/r)^{1/s} = \Delta(h,r,s)$, we obtain:
(a) If $r \leq s \leq -1$ or $1 \leq s \leq -r$ or $0 < r \leq s \leq 2r,\, s \geq 1$, then

$$k^{\frac{r-s}{sr}} F_r(\mathbf{x},\mathbf{\Phi}) \leq F_s(\mathbf{x},\mathbf{\Phi}) \leq k^{\frac{r-s}{sr}} \Delta(h,r,s)\, F_r(\mathbf{x},\mathbf{\Phi}). \tag{9.34}$$

(b) If $0 < -r < s,\, s \geq 1$ or $0 < 2r < s,\, s \geq 1$, then

$$k^{\frac{r-s}{sr}} \Delta(h,r,s)^{-1}\, F_r(\mathbf{x},\mathbf{\Phi}) \leq F_s(\mathbf{x},\mathbf{\Phi}) \leq k^{\frac{r-s}{sr}} \Delta(h,r,s)\, F_r(\mathbf{x},\mathbf{\Phi}). \tag{9.35}$$

Applying Theorem 4.3 if $-1 \leq s \leq 1$ and using that $K(km^s, kM^s, 1/s) = K(m^s, M^s, 1/s) = \Delta(h,s,1)$, we obtain:
(c) If $r \leq s,\, -1 \leq s < 0$ or $s \leq -r,\, 0 < s \leq 1$ or $0 < r \leq s \leq 2r,\, s \leq 1$, then

$$k^{\frac{r-s}{sr}} \Delta(h,s,1)^{-1}\, F_r(\mathbf{x},\mathbf{\Phi}) \leq F_s(\mathbf{x},\mathbf{\Phi}) \leq k^{\frac{r-s}{sr}} \Delta(h,s,1)\, \Delta(h,r,s)\, F_r(\mathbf{x},\mathbf{\Phi}). \tag{9.36}$$

(d) If $0 < -r < s \leq 1$ or $0 < 2r < s \leq 1$, then

$$\begin{aligned} k^{\frac{r-s}{sr}} \Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}\, F_r(\mathbf{x},\mathbf{\Phi}) &\leq F_s(\mathbf{x},\mathbf{\Phi}) \\ &\leq k^{\frac{r-s}{sr}} \Delta(h,s,1)\Delta(h,r,s)\, F_r(\mathbf{x},\mathbf{\Phi}). \end{aligned} \tag{9.37}$$

Similarly, putting $p = r/s$ in Lemma 9.2 and replace x_t by x_t^s, we obtain:
(a_1) If $1 \leq r \leq s$ or $-s \leq r \leq -1$ or $2s \leq r \leq s < 0,\, r \leq -1$, then

$$k^{\frac{s-r}{sr}} \Delta(h,r,s)^{-1}\, F_s(\mathbf{x},\mathbf{\Phi}) \leq F_r(\mathbf{x},\mathbf{\Phi}) \leq k^{\frac{s-r}{sr}} F_s(\mathbf{x},\mathbf{\Phi}). \tag{9.38}$$

(b_1) If $r < -s < 0, r \leq -1$ or $r < 2s < 0, r \leq -1$, then

$$k^{\frac{s-r}{sr}} \Delta(h,r,s)^{-1} F_s(\mathbf{x},\boldsymbol{\Phi}) \leq F_r(\mathbf{x},\boldsymbol{\Phi}) \leq k^{\frac{s-r}{sr}} \Delta(h,r,s) F_s(\mathbf{x},\boldsymbol{\Phi}). \tag{9.39}$$

(c_1) If $r \leq s, 0 < r \leq 1$ or $-s \leq r, -1 \leq r < 0$ or $2s \leq r \leq s < 0, r \geq -1$, then

$$k^{\frac{s-r}{sr}} \Delta(h,r,1)^{-1}\Delta(h,r,s)^{-1} F_s(\mathbf{x},\boldsymbol{\Phi}) \leq F_r(\mathbf{x},\boldsymbol{\Phi}) \leq k^{\frac{s-r}{sr}} \Delta(h,r,1) F_s(\mathbf{x},\boldsymbol{\Phi}). \tag{9.40}$$

(d_1) If $-1 \leq r < -s < 0$ or $-1 \leq r < 2s < 0$, then

$$\begin{aligned} k^{\frac{s-r}{sr}} \Delta(h,r,1)^{-1}\Delta(h,r,s)^{-1} F_s(\mathbf{x},\boldsymbol{\Phi}) &\leq F_r(\mathbf{x},\boldsymbol{\Phi}) \\ &\leq k^{\frac{s-r}{sr}} \Delta(h,r,1)\Delta(h,r,s) F_s(\mathbf{x},\boldsymbol{\Phi}). \end{aligned} \tag{9.41}$$

Now, we have that in cases (a) and (a_1) the inequality (9.34) holds and in cases (b) and (b_1) the inequality (9.35) holds. If we put $r = 1$ in RHS of (9.38) for $1 \leq r \leq s$ then we obtain

$$\int_T \frac{1}{k} \Phi_t(x_t)\, d\mu(t) \leq \left(\int_T \frac{1}{k} \Phi_t(x_t^s)\, d\mu(t) \right)^{1/s}, \qquad \text{if } s \geq 1.$$

Next, applying LHS of (9.34) for $s = 1$ and $0 < r \leq s \leq 2r$, we have

$$\left(\int_T \frac{1}{k} \Phi_t(x_t^r)\, d\mu(t) \right)^{1/r} \leq \int_T \frac{1}{k} \Phi_t(x_t)\, d\mu(t).$$

The assumption $s \geq 1$ implies

$$\left(\int_T \frac{1}{k} \Phi_t(x_t^r)\, d\mu(t) \right)^{1/r} \leq \int_T \frac{1}{k} \Phi_t(x_t)\, d\mu(t) \leq \left(\int_T \frac{1}{k} \Phi_t(x_t^s)\, d\mu(t) \right)^{1/s} \tag{9.42}$$

for $1/2 \leq r \leq 1 \leq s$. Similarly, putting $s = -1$ in LHS of (9.34) for $r \leq s \leq -1$ and $r = -1$ in RHS of (9.38) for $2s \leq r \leq s < 0$, we can obtain that (9.42) holds for $r \leq -1 \leq s \leq -1/2$. Consequently, we obtain that (9.34) holds in the region (i) and (9.35) holds in the regions (ii) and (iii).

In remainder cases we can choose better bounds. In the region (iv) inequalities (9.37) and (9.40) hold. Now, by Lemma 9.3 we have

$$\Delta(h,r,1) \geq \Delta(h,s,1) \qquad \text{if } r \leq s, \tag{9.43}$$

and we get

$$\Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1} \geq \Delta(h,r,1)^{-1}\Delta(h,r,s)^{-1}.$$

It follows that $k^{\frac{s-r}{sr}} \Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}$ is a better lower bound. The upper bound is equal

$$k^{\frac{s-r}{sr}} \cdot \min\{\Delta(h,r,1), \Delta(h,s,1)\Delta(h,r,s)\}.$$

In the region (v) inequalities (9.36) and (9.41) hold. We have that $k^{\frac{s-r}{sr}} \Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}$ is a better lower bound, since (9.43) holds. The upper bound is equal

$$k^{\frac{s-r}{sr}} \cdot \min\{\Delta(h,s,1), \Delta(h,r,1)\Delta(h,r,s)\} = k^{\frac{s-r}{sr}} \cdot \Delta(h,s,1),$$

since (9.43) holds and $\Delta(h,r,s) \geq 1$ by (2.32).

In the regions $(iv)_1$ and $(v)_1$ inequalities (9.36) and (9.40) hold. Analogously to the case above we obtain that the bounds in the inequality (9.36) are better. □

Finally, we give the ratio type ordering among means (9.27).

Corollary 9.6 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0<m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$. Let regions* (i) – $(\mathrm{v})_1$ *be as in Figure 9.2.*
If (r,s) in (i)*, then*

$$\Delta(h,r,s)^{-1}\,M_s(\mathbf{x},\mathbf{\Phi})\le M_r(\mathbf{x},\mathbf{\Phi})\le M_s(\mathbf{x},\mathbf{\Phi}),$$

if (r,s) in (ii) *or* (iii)*, then*

$$\Delta(h,r,s)^{-1}\,M_s(\mathbf{x},\mathbf{\Phi})\le M_r(\mathbf{x},\mathbf{\Phi})\le \Delta(h,r,s)\,M_s(\mathbf{x},\mathbf{\Phi}),$$

if (r,s) in (iv)*, then*

$$\begin{aligned}&\Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}\,M_s(\mathbf{x},\mathbf{\Phi})\le M_r(\mathbf{x},\mathbf{\Phi})\\&\le\min\{\Delta(h,r,1),\Delta(h,s,1)\Delta(h,r,s)\}\,M_s(\mathbf{x},\mathbf{\Phi}),\end{aligned}$$

if (r,s) in (v) *or* $(\mathrm{iv})_1$ *or* $(\mathrm{v})_1$*, then*

$$\Delta(h,s,1)^{-1}\Delta(h,r,s)^{-1}\,M_s(\mathbf{x},\mathbf{\Phi})\le M_r(\mathbf{x},\mathbf{\Phi})\le \Delta(h,s,1)\,M_s(\mathbf{x},\mathbf{\Phi}),$$

where $\Delta(h,r,s)$, $rs\ne 0$, is defined by (8.35)*.*

Proof. It is sufficient to multiply each inequality in Theorem 9.14 by $k^{-1/r}$. □

9.4 Difference type inequalities with power functions

In this section we consider the difference type ordering among the power functions (9.26). As an application, we consider the weighted power means (9.27).

We need some previous results given in the following two lemmas.

Lemma 9.4 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0<m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$.*
If $0<p\le 1$, then

$$\alpha_p\int_T\Phi_t(x_t)\mathrm{d}\mu(t)+k\beta_p\mathbf{1}\le\int_T\Phi_t(x_t^p)\mathrm{d}\mu(t)\le k^{1-p}\left(\int_T\Phi_t(x_t)\mathrm{d}\mu(t)\right)^p,\tag{9.44}$$

if $-1 \le p < 0$ *or* $1 \le p \le 2$, *then*

$$k^{1-p}\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right)^p \le \int_T \Phi_t(x_t^p)\mathrm{d}\mu(t) \le \alpha_p \int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k\beta_p\mathbf{1}, \tag{9.45}$$

if $p < -1$ *or* $p > 2$, *then*

$$py^{p-1}\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)y^p\mathbf{1} \le \int_T \Phi_t(x_t^p)\mathrm{d}\mu(t) \le \alpha_p \int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k\beta_p\mathbf{1} \tag{9.46}$$

for every $y \in [m,M]$. *Constants* α_p *and* β_p *are the constants* α_f *and* β_f *associated with the function* $f(z) = z^p$.

Proof. RHS of (9.44) and LHS of (9.45) are proven in Lemma 9.1. LHS of (9.44) and RHS of (9.45) and (9.46) follow from Corollary 9.2 for $f(z) = z^p$, $g(z) = z$ and $\lambda = \alpha_p$. LHS of (9.46) follows from LHS of (9.25) in Theorem 9.13 putting $f(y) = y^p$ and $l(y) = py^{p-1}$. □

Remark 9.5 *Setting* $y = (\alpha_p/p)^{1/(p-1)} \in [m,M]$ *the inequality* (9.46) *gives*

$$\begin{aligned}\alpha_p \textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)\,(\alpha_p/p)^{p/(p-1)}\mathbf{1} &\le \textstyle\int_T \Phi_t(x_t^p)\mathrm{d}\mu(t)\\ &\le \alpha_p \textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k\beta_p\mathbf{1}\end{aligned}$$

for $p < -1$ *or* $p > 2$.

Furthermore, setting $y = m$ *or* $y = M$ *gives*

$$\begin{aligned}pm^{p-1}\textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)m^p\mathbf{1} &\le \textstyle\int_T \Phi_t(x_t^p)\mathrm{d}\mu(t)\\ &\le \alpha_p \textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k\beta_p\mathbf{1}\end{aligned} \tag{9.47}$$

or

$$\begin{aligned}pM^{p-1}\textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)M^p\mathbf{1} &\le \textstyle\int_T \Phi_t(x_t^p)\mathrm{d}\mu(t)\\ &\le \alpha_p \textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k\beta_p\mathbf{1}.\end{aligned} \tag{9.48}$$

We remark that the operator in LHS of (9.47) *is positive for* $p > 2$, *since*

$$\begin{aligned}0 < km^p\mathbf{1} &\le pm^{p-1}\textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)m^p\mathbf{1}\\ &\le k(pm^{p-1}M + (1-p)m^p)\mathbf{1} < kM^p\mathbf{1}\end{aligned} \tag{9.49}$$

and the operator in LHS of (9.48) *is positive for* $p < -1$, *since*

$$\begin{aligned}0 < kM^p\mathbf{1} &\le pM^{p-1}\textstyle\int_T \Phi_t(x_t)\mathrm{d}\mu(t) + k(1-p)M^p\mathbf{1}\\ &\le k(pM^{p-1}m + (1-p)M^p)\mathbf{1} < km^p\mathbf{1}.\end{aligned} \tag{9.50}$$

(We have the inequality $pm^{p-1}M + (1-p)m^p < M^p$ *in RHS of* (9.49) *and* $pM^{p-1}m + (1-p)M^p < m^p$ *in RHS of* (9.50) *by using Bernoulli's inequality.)*

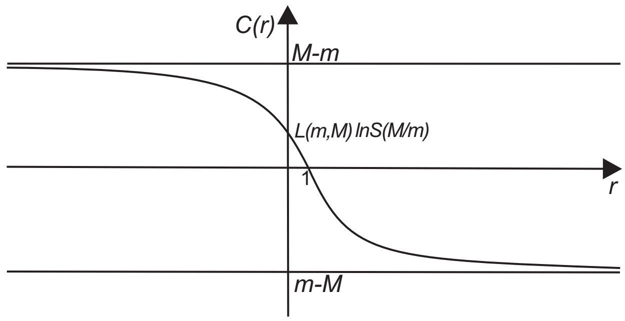


Figure 9.3: Function $C(r) \equiv C(m^r, M^r, 1/r)$

We shall need some properties of the Kantorovich constant for the difference $C(m,M,p)$ (see (2.38) and Figure 9.3).

Lemma 9.5 *Let $M > m > 0$, $r \in \mathbb{R}$ and*

$$C(m^r, M^r, 1/r) := \frac{1-r}{r}\left(r\frac{M-m}{M^r-m^r}\right)^{1/(1-r)} + \frac{M^r m - m^r M}{M^r - m^r}.$$

(i) *A function $C(r) \equiv C(m^r, M^r, 1/r)$ is strictly decreasing for all $r \in \mathbb{R}$,*

(ii) $\lim_{r\to 1} C(m^r, M^r, 1/r) = 0$ *and* $\lim_{r\to 0} C(m^r, M^r, 1/r) = L(m,M)\log S(M/m)$, *where $L(m,M)$ is the logarithmic mean and the Specht ratio $S(h)$ is defined by* (2.35).

(iii) $\lim_{r\to\infty} C(m^r, M^r, 1/r) = m - M$ *and* $\lim_{r\to -\infty} C(m^r, M^r, 1/r) = M - m$.

Proof.

(i) We have by a differential calculation

$$\begin{aligned}
&\frac{d}{dr}C(r) \\
&= \left(r\frac{M-m}{M^r-m^r}\right)^{1/(1-r)}\left(\frac{m^r\log m - M^r\log M}{r(M^r-m^r)} + \frac{1}{r(1-r)}\log\frac{r(M-m)}{M^r-m^r}\right) \\
&+ \frac{M^r m^r (M-m)\log(m/M)}{(M^r-m^r)^2}.
\end{aligned}$$

Both of functions

$$r \mapsto \frac{m^r\log m - M^r\log M}{r(M^r-m^r)} + \frac{1}{r(r-1)}\log\frac{M^r-m^r}{r(M-m)}$$

and

$$r \mapsto \frac{M^r m^r (M-m)\log(m/M)}{(M^r-m^r)^2}$$

are negative for all $r \neq 0, 1$. So $\frac{d}{dr}C(r) < 0$ and the function C is strictly decreasing.

(ii) We have by L'Hospital's theorem

$$\lim_{r\to 1}\frac{\log\left(r(M-m)/(M^r-m^r)\right)}{1-r} = -1+\frac{M\log M - m\log m}{M-m},$$

so

$$\lim_{r\to 1}\frac{1-r}{r}\left(r\frac{M-m}{M^r-m^r}\right)^{1/(1-r)} = 0\cdot \mathrm{e}^{-1+(M\log M-m\log m)/(M-m)} = 0.$$

Also,

$$\lim_{r\to 1}\frac{M^r m - m^r M}{M^r-m^r} = \lim_{r\to 1} m\frac{h^r-h}{h^r-1} = 0, \qquad h=\frac{M}{m}>1.$$

Then, $\lim_{r\to 1} C(m^r,M^r,1/r)=0$. Using Theorem 2.24, we have

$$\lim_{r\to 0} C(m^r,M^r,p/r) = L(m^p,M^p)\log S(h^p) \quad \text{for all } p\in\mathbb{R} \text{ and } h=M/m,$$

so we obtain $\lim_{r\to 0} C(m^r,M^r,1/r) = L(m,M)\log S(M/m)$.

(iii) We have by L'Hospital's theorem

$$\lim_{r\to\infty}\frac{\log\left(r(M-m)/(M^r-m^r)\right)}{1-r} = \lim_{r\to\infty}\frac{M^r m - m^r M}{M^r-m^r} = \log M,$$

so

$$\lim_{r\to\infty}\frac{1-r}{r}\left(r\frac{M-m}{M^r-m^r}\right)^{1/(1-r)} = -1\cdot \mathrm{e}^{\log M} = -M.$$

Also,

$$\lim_{r\to\infty}\frac{M^r m - m^r M}{M^r-m^r} = \lim_{r\to\infty} m\frac{h^r-h}{h^r-1} = m, \qquad h=\frac{M}{m}>1.$$

Then, $\lim_{r\to\infty} C(m^r,M^r,1/r) = m-M$.

Similarly, we obtain $\lim_{r\to-\infty} C(m^r,M^r,1/r) = M-m$.

□

Also, we need the following function order of positive operators.

Theorem 9.15 *Let A,B be positive operators in $\mathscr{B}(H)$.*
If $A\geq B>0$ and the spectrum $\mathsf{Sp}(B)\subseteq[m,M]$ for some scalars $0<m<M$, then

$$A^p + C(m,M,p)\mathbf{1} \geq B^p \qquad \textit{for all} \quad p\geq 1.$$

But, if $A\geq B>0$ and the spectrum $\mathsf{Sp}(A)\subseteq[m,M]$, $0<m<M$, then

$$B^p + C(m,M,p)\mathbf{1} \geq A^p \qquad \textit{for all} \quad p\leq -1,$$

where the Kantorovich constant for the difference $C(m,M,p)$ is defined by (2.38).

Proof. Refer to [191, Corollary 1] for the proof. □

Now, we give the the difference type ordering among power functions.

Theorem 9.16 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0<m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$. Let regions $(\mathrm{i})_1$ – $(\mathrm{v})_1$ be as in Figure 9.2.*

Then

$$C_2\mathbf{1}\le F_s(\mathbf{x},\mathbf{\Phi})-F_r(\mathbf{x},\mathbf{\Phi})\le C_1\mathbf{1}, \tag{9.51}$$

where constants $C_1\equiv C_1(m,M,s,r,k)$ and $C_2\equiv C_2(m,M,s,r,k)$ are

$$C_1=\begin{cases}\widetilde{\Delta}_k & \text{if } (r,s) \text{ in } (\mathrm{i})_1 \text{ or } (\mathrm{ii})_1 \text{ or } (\mathrm{iii})_1;\\ \widetilde{\Delta}_k+\min\{C_k(s),C_k(r)\} & \text{if } (r,s) \text{ in } (\mathrm{iv}) \text{ or } (\mathrm{v}) \text{ or } (\mathrm{iv})_1 \text{ or } (\mathrm{v})_1;\end{cases}$$

$$C_2=\begin{cases}(k^{1/s}-k^{1/r})m, & \text{if } (r,s) \text{ in } (\mathrm{i})_1;\\ \widetilde{D}_k, & \text{if } (r,s) \text{ in } (\mathrm{ii})_1;\\ \overline{D}_k, & \text{if } (r,s) \text{ in } (\mathrm{iii})_1;\\ \max\left\{\widetilde{D}_k-C_k(s),\left(k^{1/s}-k^{1/r}\right)m-C_k(r)\right\}, & \text{if } (r,s) \text{ in } (\mathrm{iv});\\ \max\left\{\overline{D}_k-C_k(r),\left(k^{1/s}-k^{1/r}\right)m-C_k(s)\right\}, & \text{if } (r,s) \text{ in } (\mathrm{v});\\ (k^{1/s}-k^{1/r})m-\min\{C_k(r),C_k(s)\}, & \text{if } (r,s) \text{ in } (\mathrm{iv})_1 \text{ or } (\mathrm{v})_1.\end{cases}$$

A constant $\widetilde{\Delta}_k\equiv\widetilde{\Delta}_k(m,M,r,s)$ is

$$\widetilde{\Delta}_k=\max_{\theta\in[0,1]}\left\{k^{1/s}[\theta M^s+(1-\theta)m^s]^{1/s}-k^{1/r}[\theta M^r+(1-\theta)m^r]^{1/r}\right\},$$

a constant $\widetilde{D}_k\equiv\widetilde{D}_k(m,M,r,s)$ is

$$\widetilde{D}_k=\min\left\{\left(k^{\frac{1}{s}}-k^{\frac{1}{r}}\right)m,\,k^{\frac{1}{s}}m\left(s\frac{M^r-m^r}{rm^r}+1\right)^{\frac{1}{s}}-k^{\frac{1}{r}}M\right\},$$

$\overline{D}_k\equiv\overline{D}_k(m,M,r,s)=-\widetilde{D}_k(M,m,s,r)$ and the Kantorovich constant for the difference $C_k(p)\equiv C_k(m,M,p)$ is defined by (2.38)*.*

Proof. This theorem follows from Lemma 9.4 by putting $p=s/r$ or $p=r/s$ and then using the Löwner-Heinz theorem, Theorem 9.15 and Lemma 9.5. We give the proof for the sake of completeness.

By Lemma 9.4 by putting $p=s/r$ or $p=r/s$ and then using the Löwner-Heinz inequality and Theorem 9.15 we have the following inequalities.
(a) If $r\le s\le -1$ or $1\le s\le -r$ or $0<r\le s\le 2r,\ s\ge 1$, then

$$\begin{aligned}\left(k^{1/s}-k^{1/r}\right)m\mathbf{1}&\le\left(k^{\frac{r-s}{rs}}-1\right)F_r(\mathbf{x},\mathbf{\Phi})\le F_s(\mathbf{x},\mathbf{\Phi})-F_r(\mathbf{x},\mathbf{\Phi})\\&\le\left(\widetilde{\alpha}\int_T\Phi_t(x_t^r)\mathrm{d}\mu(t)+k\widetilde{\beta}\mathbf{1}\right)^{1/s}-F_r(\mathbf{x},\mathbf{\Phi})\le\widetilde{\Delta}_k\mathbf{1}.\end{aligned} \tag{9.52}$$

(b) If $0 < -r < s,\, s \geq 1$ or $0 < 2r < s,\, s \geq 1$, then

$$\begin{aligned} m\left(\frac{s}{r}m^{-r}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\frac{r-s}{r}\mathbf{1}\right)^{1/s} - F_r(\mathbf{x},\boldsymbol{\Phi}) &\leq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \leq \left(\widetilde{\alpha}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\widetilde{\beta}\mathbf{1}\right)^{1/s} - F_r(\mathbf{x},\boldsymbol{\Phi}) &\leq \widetilde{\Delta}_k\mathbf{1}. \end{aligned} \tag{9.53}$$

(c) If $r \leq s,\, -1 \leq s < 0$ or $s \leq -r,\, 0 < s \leq 1$ or $0 < r \leq s \leq 2r,\, s \leq 1$, then

$$\begin{aligned} \left(\left(k^{1/s} - k^{1/r}\right)m - C_k(s)\right)\mathbf{1} &\leq \left(k^{\frac{r-s}{rs}} - 1\right)F_r(\mathbf{x},\boldsymbol{\Phi}) - C_k(s)\mathbf{1} \\ &\leq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \leq \left(\widetilde{\alpha}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\widetilde{\beta}\mathbf{1}\right)^{1/s} &- F_r(\mathbf{x},\boldsymbol{\Phi}) + C_k(s)\mathbf{1} \leq \left(\widetilde{\Delta}_k + C_k(s)\right)\mathbf{1}. \end{aligned} \tag{9.54}$$

(d) If $0 < -r < s \leq 1$ or $0 < 2r < s \leq 1$, then

$$\begin{aligned} m\left(\frac{s}{r}m^{-r}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\frac{r-s}{r}\mathbf{1}\right)^{1/s} &- F_r(\mathbf{x},\boldsymbol{\Phi}) - C_k(s)\mathbf{1} \\ &\leq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \leq \left(\widetilde{\alpha}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\widetilde{\beta}\mathbf{1}\right)^{1/s} &- F_r(\mathbf{x},\boldsymbol{\Phi}) + C_k(s)\mathbf{1} \leq \left(\widetilde{\Delta}_k + C_k(s)\right)\mathbf{1}. \end{aligned} \tag{9.55}$$

Moreover, we can obtain the following inequalities:

(a_1) If $1 \leq r \leq s$ or $-s \leq r \leq -1$ or $2s \leq r \leq s < 0,\, r \leq -1$, then

$$\begin{aligned} \widetilde{\Delta}_k\mathbf{1} \geq F_s(\mathbf{x},\boldsymbol{\Phi}) - \left(\bar{\alpha}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\bar{\beta}\mathbf{1}\right)^{1/r} &\geq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \geq \left(1 - k^{\frac{s-r}{rs}}\right)F_s(\mathbf{x},\boldsymbol{\Phi}) &\geq \left(k^{1/s} - k^{1/r}\right)m\mathbf{1}. \end{aligned} \tag{9.56}$$

(b_1) If $r < -s < 0,\, r \leq -1$ or $r < 2s < 0,\, r \leq -1$, then

$$\begin{aligned} \widetilde{\Delta}_k\mathbf{1} \geq F_s(\mathbf{x},\boldsymbol{\Phi}) - \left(\bar{\alpha}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\bar{\beta}\mathbf{1}\right)^{1/r} &\geq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \geq F_s(\mathbf{x},\boldsymbol{\Phi}) - M\left(\frac{r}{s}M^{-s}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\frac{s-r}{s}\mathbf{1}\right)^{1/r}&. \end{aligned} \tag{9.57}$$

(c_1) If $r \leq s,\, 0 < r \leq 1$ or $-s \leq r,\, -1 \leq r < 0$ or $2s \leq r \leq s < 0,\, r \geq -1$, then

$$\begin{aligned} (\widetilde{\Delta}_k + C_k(r))\mathbf{1} \geq F_s(\mathbf{x},\boldsymbol{\Phi}) - \left(\bar{\alpha}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\bar{\beta}\mathbf{1}\right)^{1/r} &+ C_k(r)\mathbf{1} \\ \geq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi})& \\ \geq \left(1 - k^{\frac{s-r}{rs}}\right)F_s(\mathbf{x},\boldsymbol{\Phi}) - C_k(r)\mathbf{1} \geq \left(\left(k^{1/s} - k^{1/r}\right)m - C_k(r)\right)\mathbf{1}&. \end{aligned} \tag{9.58}$$

(d_1) If $-1 \le r < -s < 0$ or $-1 \le r < 2s < 0$, then

$$\begin{aligned}(\widetilde{\Delta}_k + C_k(r))\mathbf{1} &\geq F_s(\mathbf{x},\boldsymbol{\Phi}) - \left(\overline{\alpha}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\overline{\beta}\mathbf{1}\right)^{1/r} + C_k(r)\mathbf{1} \\ &\geq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ &\geq F_s(\mathbf{x},\boldsymbol{\Phi}) - M\left(\frac{r}{s}M^{-s}\int_T \Phi_t(x_t^s)\mathrm{d}\mu(t) + k\frac{s-r}{s}\mathbf{1}\right)^{1/r} - C_k(r)\mathbf{1},\end{aligned}$$

where we denote

$$\widetilde{\alpha} = \tfrac{M^s - m^s}{M^r - m^r}, \quad \widetilde{\beta} = \tfrac{M^r m^s - M^s m^r}{M^r - m^r}, \quad \overline{\alpha} = \tfrac{M^r - m^r}{M^s - m^s}, \quad \overline{\beta} = \tfrac{M^s m^r - M^r m^s}{M^s - m^s},$$
$$C(km^s, kM^s, 1/s) = k^{1/s}\, C(m^s, M^s, 1/s) = C_k(s),$$

$$\widetilde{\Delta}_k = \max_{z\in\overline{T_1}}\left\{k^{1/s}\left(\widetilde{\alpha}z + \widetilde{\beta}\right)^{1/s} - k^{1/r}z^{1/r}\right\} = \max_{z\in\overline{T_2}}\left\{k^{1/s}z^{1/s} - k^{1/r}\left(\overline{\alpha}z + \overline{\beta}\right)^{1/r}\right\},$$

and $\overline{T_1}$ and $\overline{T_2}$ denote the closed intervals joining m^r to M^r and m^s to M^s, respectively.

We will determine lower bounds in LHS of (b) and (d), in RHS of (b_1) and (d_1).

For LHS of (9.53) we can obtain

$$\begin{aligned} m\left(\frac{s}{r}m^{-r}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\frac{r-s}{r}\mathbf{1}\right)^{1/s} - F_r(\mathbf{x},\boldsymbol{\Phi}) \\ \geq \min\nolimits_{z\in\overline{T_1}}\left\{k^{1/s}m\left(\tfrac{s}{r}m^{-r}z + 1 - \tfrac{s}{r}\right)^{1/s} - k^{1/r}z^{1/r}\right\}\mathbf{1} = \widetilde{D}_k\mathbf{1}. \end{aligned} \tag{9.59}$$

Really, using substitution $z = rm^r\left(x - \frac{1}{s}\right)$, finding the minimum of the function $h(z) = k^{1/s}m\left(\frac{s}{r}m^{-r}z + \frac{r-s}{r}\right)^{1/s} - k^{1/r}z^{1/r}$ on $\overline{T_1}$ is equivalent to finding the minimum of $h_1(x) = k^{1/s}m\left(s(x-\frac{1}{r})\right)^{1/s} - k^{1/r}m\left(r(x-\frac{1}{s})\right)^{1/r}$ on $\overline{T} = [\frac{1}{s}+\frac{1}{r}, \frac{1}{s}+\frac{1}{r}\frac{M^r}{m^r}]$. The domain of h_1 is $\mathrm{S} = [\frac{1}{r},\infty)$ for $r>0$ or $\mathrm{S} = [\frac{1}{r},\frac{1}{s})$ for $r<0$. We have $h_1''(x) = k^{1/s}m(1-s)\left(s(x-\frac{1}{r})\right)^{1/s-2} - k^{1/r}m(1-r)\left(r(x-\frac{1}{s})\right)^{1/r-2}$. If $r<1$ and $s\geq 1$ then $h_1''(x)<0$, since $k^{1/s}m(1-s)\left(s(x-\frac{1}{r})\right)^{1/s-2} \leq 0 < k^{1/r}m(1-r)\left(r(x-\frac{1}{s})\right)^{1/r-2}$. It follows that h_1 is concave on S for $r<1$ and $s\geq 1$. In this case we obtain

$$\min_{z\in\overline{T_1}} h(z) = \min_{x\in\overline{T}} h_1(x) = \min\left\{h_1\left(\frac{1}{s}+\frac{1}{r}\right), h_1\left(\frac{1}{s}+\frac{1}{r}\frac{M^r}{m^r}\right)\right\} = \widetilde{D}_k. \tag{9.60}$$

If $1<r<s$, then we have $\lim_{x\to\frac{1}{r}} h_1(x) = -k^{1/r}m(\frac{s-r}{s})^{\frac{1}{r}} < 0$, $\lim_{x\to\infty} h_1(x) = -\infty$. If $x_0 > \frac{1}{r}$ is the stationary point of the function h_1, then $h_1(x_0)$ is the maximum value, since $h_1''(x_0) = k^{\frac{1}{s}}m\left(s(x_0-\frac{1}{r})\right)^{1/s-2}\left(r(x_0-\frac{1}{s})\right)^{-1}(r-s)(x_0+1-\frac{r+s}{rs}) < 0$. It follows that (9.60) is also true in this case.

So in the case (b) we obtain:

$$\widetilde{D}_k\mathbf{1} \leq F_s(\mathbf{x},\boldsymbol{\Phi}) - F_r(\mathbf{x},\boldsymbol{\Phi}) \leq \widetilde{\Delta}_k\mathbf{1} \tag{9.61}$$

and in the case (d) we obtain:

$$\left(\widetilde{D}_k - C_k(s)\right)\mathbf{1} \le F_s(\mathbf{x},\mathbf{\Phi}) - F_r(\mathbf{x},\mathbf{\Phi}) \le \left(\widetilde{\Delta}_k + C_k(s)\right)\mathbf{1}. \tag{9.62}$$

Similarly, for the RHS of (9.57) we obtain

$$\begin{aligned}
&F_s(\mathbf{x},\mathbf{\Phi}) - M\left(\frac{r}{s}M^{-s}\int_T \Phi_t(x_t^r)\mathrm{d}\mu(t) + k\frac{s-r}{s}\mathbf{1}\right)^{1/r} \\
&\ge \min_{z\in T_2}\left\{k^{1/s}z^{1/s} - k^{1/r}M\left(\frac{r}{s}M^{-s}z + 1 - \frac{r}{s}\right)^{1/r}\right\}\mathbf{1} \\
&= \min\left\{k^{1/s}m - k^{1/r}M\left(\frac{r}{s}\frac{m^s}{M^s} + 1 - \frac{r}{s}\right)^{1/r}, \left(k^{1/s} - k^{1/r}\right)M\right\}\mathbf{1} \\
&= \overline{D}_k\mathbf{1}.
\end{aligned}$$

So in the case (b_1) we obtain:

$$\overline{D}_k\mathbf{1} \le F_s(\mathbf{x},\mathbf{\Phi}) - F_r(\mathbf{x},\mathbf{\Phi}) \le \widetilde{\Delta}_k\mathbf{1} \tag{9.63}$$

and in the case (d_1) we obtain:

$$\left(\overline{D}_k - C_k(r)\right)\mathbf{1} \le F_s(\mathbf{x},\mathbf{\Phi}) - F_r(\mathbf{x},\mathbf{\Phi}) \le \left(\widetilde{\Delta}_k + C_k(r)\right)\mathbf{1}. \tag{9.64}$$

Finally, we can obtain desired bounds C_1 and C_2 in (9.51), taking into account that (9.52) holds in the region $(\mathrm{i})_1$, (9.61) holds in $(\mathrm{ii})_1$, (9.63) holds in $(\mathrm{iii})_1$, (9.62) and (9.58) hold in (iv), (9.54) and (9.64) hold in (v), (9.54) and (9.58) hold in $(\mathrm{iv})_1$ and $(\mathrm{v})_1$. □

Finally, we give the difference type ordering among means (9.27).

Corollary 9.7 *Let $(x_t)_{t\in T}$ be a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $0 < m < M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T} \in P_k[\mathscr{A},\mathscr{B}]$. Let regions* (i) – $(\mathrm{v})_1$ *be as in Figure 9.2.*
If (r,s) in (i), *then*

$$0 \le M_s(\mathbf{x},\mathbf{\Phi}) - M_r(\mathbf{x},\mathbf{\Phi}) \le \tilde{\Delta}\mathbf{1},$$

if (r,s) in (ii), *then*

$$\left(m\left(\frac{s}{r}\frac{M^r}{m^r} + 1 - \frac{s}{r}\right)^{1/s} - M\right)\mathbf{1} \le M_s(\mathbf{x},\mathbf{\Phi}) - M_r(\mathbf{x},\mathbf{\Phi}) \le \tilde{\Delta}\mathbf{1},$$

if (r,s) in (iii), *then*

$$\left(m - M\left(\frac{r}{s}\frac{m^s}{M^s} + 1 - \frac{r}{s}\right)^{1/r}\right)\mathbf{1} \le M_s(\mathbf{x},\mathbf{\Phi}) - M_r(\mathbf{x},\mathbf{\Phi}) \le \tilde{\Delta}\mathbf{1},$$

if (r,s) *in* (iv)*, then*

$$\max\left\{m\left(\tfrac{s}{r}\tfrac{M^r}{m^r}+\tfrac{r-s}{r}\right)^{1/s}-M-\;C(m^s,M^s,1/s),-C(m^r,M^r,1/r)\right\}\mathbf{1}$$
$$\leq M_s(\mathbf{x},\boldsymbol{\Phi})-M_r(\mathbf{x},\boldsymbol{\Phi})\leq\left(\tilde{\Delta}+C(m^s,M^s,1/s)\right)\mathbf{1},$$

if (r,s) *in* (v) *or* $(\mathrm{iv})_1$ *or* $(\mathrm{v})_1$*, then*

$$-C(m^s,M^s,1/s)\mathbf{1}\leq M_s(\mathbf{x},\boldsymbol{\Phi})-M_r(\mathbf{x},\boldsymbol{\Phi})\leq\left(\tilde{\Delta}+C(m^s,M^s,1/s)\right)\mathbf{1},$$

where a constant $\tilde{\Delta}\equiv\tilde{\Delta}(m,M,r,s)$ *is*

$$\tilde{\Delta}=\max_{\theta\in[0,1]}\left\{[\theta M^s+(1-\theta)m^s]^{1/s}-[\theta M^r+(1-\theta)m^r]^{1/r}\right\}$$

and the Kantorovich constant for the difference $C(n,N,p)$ *is defined by* (2.38)*.*

Proof. This corollary follows from Theorem 9.16 putting $k=1$, and then replacing Φ_t by $\frac{1}{k}\Phi_t$, $t\in T$. Finally we choose a better bounds using that $C(m^r,M^r,1/r)\geq C(m^s,M^s,1/s)$ holds for $r\leq s$ by (9.43) and $\widetilde{D}_k=\widetilde{D}_1=m\left(s\frac{M^r-m^r}{rm^r}+1\right)^{\frac{1}{s}}-M$, since $1-\frac{M}{m}<1-\left(\frac{s}{r}\frac{M^r}{m^r}+\left(1-\frac{s}{r}\right)\right)^{\frac{1}{s}}$ holds by (9.49) and (9.50). □

9.5 Quasi-arithmetic means

In this section we give the order among the following generalized quasi-arithmetic operator means

$$M_\varphi(\mathbf{x},\boldsymbol{\Phi})=\varphi^{-1}\left(\int_T\frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)\right),\tag{9.65}$$

under these conditions $(x_t)_{t\in T}$ is a bounded continuous field of positive operators in a unital C^*-algebra $\mathscr{A}$ with spectra in $[m,M]$ for some scalars $m<M$, $(\Phi_t)_{t\in T}\in P_k[\mathscr{A},\mathscr{B}]$ and $\varphi\in\mathscr{C}[m,M]$ is a strictly monotone function.

We denote $M_\varphi(\mathbf{x},\boldsymbol{\Phi})$ shortly with M_φ. It is easy to see that the mean M_φ is well defined.

As a special case of (9.65), we may consider the power operator mean (9.27), which is studied in Sections 9.3 and 9.4.

First, we study the monotonicity of quasi-arithmetic means.

Theorem 9.17 *Let* $(x_t)_{t\in T}$*,* $(\Phi_t)_{t\in T}$ *be as in the definition of the quasi-arithmetic mean* (9.65)*. Let* $\psi,\varphi\in\mathscr{C}[m,M]$ *be strictly monotone functions.*
If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is operator concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$M_\varphi \leq M_\psi. \tag{9.66}$$

If one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is operator convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the reverse inequality is valid in (9.66)*.*

Proof. We prove only the case (i). If we put $f = \psi \circ \varphi^{-1}$ in Theorem 9.9 and replace x_t with $\varphi(x_t)$, then we obtain

$$\psi \circ \varphi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)\right) \leq \int_T \frac{1}{k}\Phi_t\left(\psi(x_t)\right)\mathrm{d}\mu(t). \tag{9.67}$$

Since ψ^{-1} is operator monotone, it follows that

$$\varphi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)\right) \leq \psi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(\psi(x_t)\right)\mathrm{d}\mu(t)\right),$$

which is the desired inequality (9.66). □

We can give the following generalization of the previous theorem.

Corollary 9.8 *Let* $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ *be as in the definition of the quasi-arithmetic mean* (9.65)*. Let* $\psi, \varphi \in \mathscr{C}[m,M]$ *be strictly monotone functions and* $F : [m,M]\times[m,M] \to \mathbb{R}$ *be a bounded and operator monotone function in its first variable, such that* $F(z,z) = C$ *for all* $z \in [m,M]$*.*
If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is operator concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$F\left[M_\psi, M_\varphi\right] \geq C\mathbf{1}. \tag{9.68}$$

If one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is operator convex and* $-\psi^{-1}$ *is operator monotone,*

is satisfied, then the reverse inequality is valid in (9.68)*.*

Proof. Suppose (i) or (i'). Then by Theorem 9.17 we have $M_\varphi \le M_\psi$. Using assumptions about function F, it follows

$$F\left[M_\psi, M_\varphi\right] \ge F\left[M_\varphi, M_\varphi\right] \ge \inf_{m \le z \le M} F(z,z)\mathbf{1} = C\mathbf{1}.$$

In the remaining cases the proof is essentially the same as in previous cases. □

Theorem 9.18 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and $\psi, \varphi \in \mathscr{C}[m,M]$ be strictly monotone functions.*

(i) *If φ^{-1} is operator convex and ψ^{-1} is operator concave, then*

$$M_\varphi \le M_1 \le M_\psi. \tag{9.69}$$

(ii) *If φ^{-1} is operator concave and ψ^{-1} is operator convex then the reverse inequality is valid in* (9.69)*.*

Proof. We prove only the case (i): Using Theorem 9.9 for a operator convex function φ^{-1} on $[\varphi_m, \varphi_M]$, we have

$$M_\varphi = \varphi^{-1}\left(\frac{1}{k}\int_T \Phi_t\left(\varphi(x_t)\right) \mathrm{d}\mu(t)\right) \le \frac{1}{k}\int_T \Phi_t\left(x_t\right) \mathrm{d}\mu(t) = M_1,$$

which gives LHS of (9.69). Similarly, since ψ^{-1} is operator concave on $J = [\psi_m, \psi_M]$, we have

$$M_1 = \frac{1}{k}\int_T \Phi_t\left(x_t\right) \mathrm{d}\mu(t) \le \psi^{-1}\left(\frac{1}{k}\int_T \Phi_t\left(\psi(x_t)\right) \mathrm{d}\mu(t)\right) = M_\psi,$$

which gives RHS of (9.69). □

Theorem 9.19 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and $\psi, \varphi \in \mathscr{C}[m,M]$ be strictly monotone functions. Then*

$$M_\varphi = M_\psi \qquad \text{for all } (x_t)_{t\in T},\ (\Phi_t)_{t\in T}$$

if and only if

$$\varphi = A\psi + B \qquad \text{for some real numbers } A \ne 0 \text{ and } B.$$

Proof. The case $\varphi = A\psi + B \Rightarrow M_\varphi = M_\psi$ is obvious.
$M_\varphi = M_\psi \Rightarrow \varphi = A\psi + B$: Let

$$\varphi^{-1}\left(\frac{1}{k}\int_T \Phi_t\left(\varphi(x_t)\right) \mathrm{d}\mu(t)\right) = \psi^{-1}\left(\frac{1}{k}\int_T \Phi_t\left(\psi(x_t)\right) \mathrm{d}\mu(t)\right)$$

for all $(x_t)_{t\in T}$ and $(\Phi_t)_{t\in T}$. Setting $y_t = \varphi(x_t) \in B(H)$, $\varphi_m\mathbf{1} \le y_t \le \varphi_M\mathbf{1}$, we obtain

$$\psi\circ\varphi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(y_t\right) \mathrm{d}\mu(t)\right) = \int_T \frac{1}{k}\Phi_t\left(\psi\circ\varphi^{-1}(y_t)\right) \mathrm{d}\mu(t)$$

for all $(y_t)_{t\in T}$ and $(\Phi_t)_{t\in T}$. M. D. Choi showed in [35, Theorem 2.5] that if $\Phi : \mathscr{A} \to \mathscr{B}$ is a positive linear mapping, f is a non-affine operator convex function on $(-a,a)$, and $f(\Phi(x)) = \Phi(f(x))$ for all Hermitian x in C^*-algebra $\mathscr{A}$ with spectra in $(-a,a)$, then Φ is a C^*-homomorphism. Similarly as above, in our case we can obtain that $\psi \circ \varphi^{-1}$ is affine, i.e. $\psi \circ \varphi^{-1}(u) = Au + B$ for some real numbers $A \neq 0$ and B, which gives the desired connection: $\psi(v) = A\varphi(v) + B$. □

Using properties of operator monotone or operator convex functions we can obtain some corollaries of Theorems 9.17 and 9.18. E.g. we have the following corollary.

Corollary 9.9 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and $0 \le m < M$. Let φ and ψ be continuous strictly monotone functions from $[0,\infty)$ into itself.*
If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *and* ψ^{-1} *are operator monotone,*

(ii) $\varphi \circ \psi^{-1}$ *is operator convex,* $\varphi \circ \psi^{-1}(0) = 0$ *and* ψ^{-1} *is operator monotone*

is satisfied, then

$$M_\psi \le M_1 \le M_\varphi.$$

Specially, if one of the following conditions

(ii) ψ^{-1} *is operator monotone,*

(ii') ψ^{-1} *is operator convex and* $\varphi(0) = 0$*,*

is satisfied, then

$$M_1 \le M_\psi.$$

Proof. This theorem follows directly from Theorem 9.17.

We prove only the case (i). We use the statement: *a bounded below function $f \in C([\alpha,\infty))$ is operator monotone iff f is operator concave* and we apply Theorem 9.17-(ii). □

Example 9.1 *If we put $\varphi(t) = t^r$, $\psi(t) = t^s$ or $\varphi(t) = t^s$, $\psi(t) = t^r$ in Theorem 9.17 and Theorem 9.18, then we obtain (cf. Corollary 9.6)*

$$M_r(\mathbf{x},\boldsymbol{\Phi}) \le M_s(\mathbf{x},\boldsymbol{\Phi})$$

for either $r \le s$, $r \notin (-1,1)$, $s \notin (-1,1)$ or $1/2 \le r \le 1 \le s$ or $r \le -1 \le s \le -1/2$.

Next, we study the difference and ratio type inequalities among quasi-arithmetic means. With that in mind, we shall prove the following general result.

Theorem 9.20 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65). *Let $\psi,\varphi \in \mathscr{C}[m,M]$ be strictly monotone functions and let $F : [m,M] \times [m,M] \to \mathbb{R}$ be a bounded and operator monotone function in its first variable.*
If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$F\left[M_\psi, M_\varphi\right] \leq \sup_{0\leq\theta\leq 1} F\left[\psi^{-1}\left(\theta\psi(m)+(1-\theta)\psi(M), \varphi^{-1}\left(\theta\varphi(m)+(1-\theta)\varphi(M)\right)\right)\right] \mathbf{1}. \tag{9.70}$$

If one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the opposite inequality is valid in (9.70) *with* inf *instead of* sup.

Proof. We prove only the case (i). Since $f \in \mathscr{C}[m,M]$ is convex then

$$f(z) \leq \frac{M-z}{M-m} f(m) + \frac{z-m}{M-m} f(M)$$

holds for any $z \in [m,M]$. Replacing f by $\psi \circ \varphi^{-1}$, and z by $\varphi(z)$ and introducing the notation $\varphi_m = \min\{\varphi(m), \varphi(M)\}$, $\varphi_M = \max\{\varphi(m), \varphi(M)\}$, we have

$$\psi(z) \leq \frac{\varphi_M - \varphi(z)}{\varphi_M - \varphi_m} \psi \circ \varphi^{-1}(\varphi_m) + \frac{\varphi(z) - \varphi_m}{\varphi_M - \varphi_m} \psi \circ \varphi^{-1}(\varphi_M), \quad \text{for any } z \in [m,M].$$

Thus, replacing z by x_t for $t \in T$, applying the positive linear mappings $\frac{1}{k}\Phi_t$ and integrating, we obtain that

$$\begin{aligned} \int_T \frac{1}{k} \Phi_t\left(\psi(x_t)\right) \mathrm{d}\mu(t) \leq\ & \frac{\varphi_M \mathbf{1} - \int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)}{\varphi_M - \varphi_m} \psi \circ \varphi^{-1}(\varphi_m) \\ & + \frac{\int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t) - \varphi_m \mathbf{1}}{\varphi_M - \varphi_m} \psi \circ \varphi^{-1}(\varphi_M) \end{aligned} \tag{9.71}$$

holds, since $\int_T \frac{1}{k}\Phi_t(\mathbf{1})\,\mathrm{d}\mu(t) = \mathbf{1}$. We denote briefly

$$B = \frac{\varphi(M)\mathbf{1} - \int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)}{\varphi(M) - \varphi(m)}. \tag{9.72}$$

Since $0 \leq \varphi(M)\mathbf{1} - \int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t) \leq (\varphi(M)-\varphi(m))\mathbf{1}$ holds for a increasing function φ or $(\varphi(M)-\varphi(m))\mathbf{1} \leq \varphi(M)\mathbf{1} - \int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t) \leq 0$ holds for a decreasing function φ, then $0 \leq B \leq \mathbf{1}$ holds for any monotone function φ. It is easy to check that the inequality (9.71) becomes

$$\int_T \frac{1}{k}\Phi_t\left(\psi(x_t)\right)\mathrm{d}\mu(t) \leq B\psi(m) + (\mathbf{1}-B)\psi(M).$$

Next, applying an operator monotone function ψ^{-1} to the above inequality, we obtain

$$M_\psi = \psi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(\psi(x_t)\right)\mathrm{d}\mu(t)\right) \leq \psi^{-1}\left(B\psi(m) + (\mathbf{1}-B)\psi(M)\right).$$

Also, using (9.72), we can write

$$M_\varphi = \varphi^{-1}\left(\int_T \frac{1}{k}\Phi_t\left(\varphi(x_t)\right)\mathrm{d}\mu(t)\right) = \varphi^{-1}\left(B\varphi(m) + (\mathbf{1}-B)\varphi(M)\right).$$

Finally using operator monotonicity of $F(\cdot,v)$, we have

$$\begin{aligned} &F\left[M_\psi, M_\varphi\right] \\ &\leq F\left[\psi^{-1}\left(B\psi(m) + (\mathbf{1}-B)\psi(M)\right), \varphi^{-1}\left(B\varphi(m) + (\mathbf{1}-B)\varphi(M)\right)\right] \\ &\leq \sup_{0\leq\theta\leq 1} F\left[\psi^{-1}\left(\theta\psi(m) + (1-\theta)\psi(M), \varphi^{-1}\left(\theta\varphi(m) + (1-\theta)\varphi(M)\right)\right)\right]\mathbf{1}, \end{aligned}$$

which is the desired inequality (9.70). □

Remark 9.6 We can obtain similar inequalities as in Theorem 9.20 when $F : [m,M] \times [m,M] \to \mathbb{R}$ is a bounded and operator monotone function in its second variable.

If the function F in Theorem 9.20 has the form $F(u,v) = u - v$ and $F(u,v) = v^{-1/2}uv^{-1/2}$ $(v > 0)$, we obtain the difference and ratio type inequalities.

Corollary 9.10 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and let $\psi, \varphi \in \mathscr{C}[m,M]$ be strictly monotone functions.*
If one of the following conditions

(i) $\psi\circ\varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(ii) $\psi\circ\varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$M_\psi \leq M_\varphi + \max_{0\leq\theta\leq 1}\left\{\psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right) - \varphi^{-1}\left(\theta\varphi(M) + (1-\theta)\varphi(m)\right)\right\}.$$

If in addition $\varphi > 0$ on $[m,M]$, then

$$M_\psi \leq \max_{0\leq\theta\leq 1}\left\{\frac{\psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right)}{\varphi^{-1}\left(\theta\varphi(M) + (1-\theta)\varphi(m)\right)}\right\} M_\varphi.$$

If one of the following conditions

(ii) $\psi\circ\varphi^{-1}$ *is concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi\circ\varphi^{-1}$ *is convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the opposite inequalities are valid with min *instead of* max.

We will give a complementary result to (i) or (i') of Theorem 9.17 under the assumption that $\psi\circ\varphi^{-1}$ is operator convex and ψ^{-1} is not operator monotone. In the following theorem we give a general result.

Theorem 9.21 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65). *Let $\psi,\varphi\in\mathscr{C}[m,M]$ be strictly monotone functions and $F:[m,M]\times[m,M]\to\mathbb{R}$ be a bounded and operator monotone function in its first variable.*
If one of the following conditions

(i) $\psi\circ\varphi^{-1}$ *is operator convex and* ψ^{-1} *is increasing convex,*

(i') $\psi\circ\varphi^{-1}$ *is operator concave and* ψ^{-1} *is decreasing convex,*

is satisfied, then

$$F\left[M_\varphi,M_\psi\right]\leq\sup_{0\leq\theta\leq1}F\left[\theta M+(1-\theta)m,\psi^{-1}(\theta\psi(M)+(1-\theta)\psi(m))\right]\mathbf{1}. \tag{9.73}$$

If one of the following conditions

(ii) $\psi\circ\varphi^{-1}$ *is operator convex and* ψ^{-1} *is decreasing concave,*

(ii') $\psi\circ\varphi^{-1}$ *is operator concave and* ψ^{-1} *is increasing concave,*

is satisfied, then the opposite inequality is valid in (9.73) *with* inf *instead of* sup.

Proof. We prove only the case (i): If we put $f=\psi\circ\varphi^{-1}$ in Theorem 9.9 and replace x_t with $\varphi(x_t)$, then we obtain (see (9.67))

$$\psi(M_\varphi)\leq\psi(M_\psi) \tag{9.74}$$

Since ψ^{-1} is increasing, then $\psi(m)\mathbf{1}\leq\psi(M_\varphi)\leq\psi(M)\mathbf{1}$, and also since ψ^{-1} is convex we have

$$\begin{aligned}
M_\varphi &= \psi^{-1}(\psi(M_\varphi))\\
&\leq \frac{M-m}{\psi(M)-\psi(m)}\big(\psi(M_\varphi)-\psi(m)\mathbf{1}\big)+m\mathbf{1} \quad \text{by convexity of } \psi^{-1}\\
&\leq \frac{M-m}{\psi(M)-\psi(m)}\big(\psi(M_\psi)-\psi(m)\mathbf{1}\big)+m\mathbf{1} \quad \text{by increase of } \psi \text{ and (9.74).}
\end{aligned}$$

Now, operator monotonicity of $F(\cdot,v)$ give

$$\begin{aligned}
F\left[M_\varphi,M_\psi\right] &\leq F\left[\frac{M-m}{\psi(M)-\psi(m)}\ \big(\psi(M_\psi)-\psi(m)\mathbf{1}\big)+m\mathbf{1},\psi^{-1}\big(\psi(M_\psi)\big)\right]\\
&\leq \sup_{\psi(m)\leq z\leq\psi(M)}F\left[\frac{M-m}{\psi(M)-\psi(m)}\ (z-\psi(m))+m,\psi^{-1}(z)\right]\mathbf{1}\\
&= \sup_{0\leq\theta\leq1}F\left[\theta M+(1-\theta)m,\psi^{-1}(\theta\psi(M)+(1-\theta)\psi(m))\right]\mathbf{1},
\end{aligned}$$

which is the desired inequality (9.73). □

Remark 9.7 *Similar to Corollary 9.10, by using Theorem 9.21 we have the following results.*
Let one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is increasing convex,*

(i') $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is decreasing convex*

be satisfied. Then

$$M_\varphi \le M_\psi + \max_{0\le\theta\le 1} \left\{ \theta M + (1-\theta)m - \psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right)\right\} \mathbf{1},$$

and if, additionally, $\psi > 0$ *on* $[m,M]$*, then*

$$M_\varphi \le \max_{0\le\theta\le 1} \left\{ \frac{\theta M + (1-\theta)m}{\psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right)} \right\} M_\psi.$$

Let one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is decreasing concave,*

(ii') $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is increasing concave*

be satisfied. Then the opposite inequalities are valid with min *instead of* max.

In the following theorem we give the complementary result to the one given in the above remark.

Theorem 9.22 *Let* $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ *be as in the definition of the quasi-arithmetic mean* (9.65) *and* $\psi, \varphi \in \mathscr{C}[m,M]$ *be strictly monotone functions.*

(i) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is decreasing convex,*

(i') $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is increasing convex*

be satisfied. Then

$$M_\psi \le M_\varphi + \max_{0\le\theta\le 1} \left\{ \theta M + (1-\theta)m - \psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right)\right\} \mathbf{1}, \tag{9.75}$$

and if, additionally, $\psi > 0$ *on* $[m,M]$*, then*

$$M_\psi \le \max_{0\le\theta\le 1} \left\{ \frac{\theta M + (1-\theta)m}{\psi^{-1}\left(\theta\psi(M) + (1-\theta)\psi(m)\right)} \right\} M_\varphi. \tag{9.76}$$

Let one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is increasing concave,*

(ii') $\psi \circ \varphi^{-1}$ *is operator concave and* ψ^{-1} *is decreasing concave* *be satisfied. Then the opposite inequality is valid in* (9.75) *with* min *instead of* max.

If, additionally, $\psi > 0$ *on* $[m,M]$*, then the opposite inequality is valid in* (9.76) *with* min *instead of* max.

Proof. We prove only the case (i): Since $\psi \circ \varphi^{-1}$ is operator convex, then $\psi(M_\varphi) \le \psi(M_\psi)$ holds. Next, for every unit vector $x \in H$ we have

$$
\begin{aligned}
&\langle M_\varphi x, x\rangle \\
&= \langle \psi^{-1} \circ \psi(M_\varphi)x, x\rangle \\
&\ge \psi^{-1}\langle \psi(M_\varphi)x, x\rangle \qquad \text{by convexity of } \psi^{-1} \\
&\ge \psi^{-1}\langle \psi(M_\psi)x, x\rangle \qquad \text{by decrease of } \psi^{-1} \text{ and operator convexity } \psi \circ \varphi^{-1} \\
&\ge \langle M_\psi x, x\rangle - \max_{\psi(M) \le z \le \psi(m)} \left\{ \frac{m-M}{\psi^{-1}(m) - \psi^{-1}(M)}(z-m) + \psi^{-1}(m) - \psi^{-1}(z) \right\} \\
&\qquad \text{by convexity of } \psi^{-1} \text{ and using the Mond-Pečarić method} \\
&= \langle M_\psi x, x\rangle - \max_{0 \le \theta \le 1} \left\{ \theta M + (1-\theta)m - \psi^{-1}(\theta\psi(M) + (1-\theta)\psi(m)) \right\}
\end{aligned}
$$

and hence we have the desired inequality (9.75).

Similarly, we can check that (9.76) holds. □

We will give a complementary result to Theorem 9.18. In the following theorem we give a general result.

Theorem 9.23 *Let* $(x_t)_{t\in T}$*,* $(\Phi_t)_{t\in T}$ *be as in the definition of the quasi-arithmetic mean* (9.65) *and* $\psi, \varphi \in \mathscr{C}[m,M]$ *be strictly monotone functions and* $F : [m,M] \times [m,M] \to \mathbb{R}$ *be a bounded and operator monotone function in its first variable.*

(i) *If* φ^{-1} *is operator convex and* ψ^{-1} *is concave, then*

$$F[M_\varphi, M_\psi] \le \sup_{0 \le \theta \le 1} F\left[\theta M + (1-\theta)m, \psi^{-1}(\theta\psi(M) + (1-\theta)\psi(m))\right] \mathbf{1}. \tag{9.77}$$

(ii) *If* φ^{-1} *is convex and* ψ^{-1} *is operator concave, then*

$$F[M_\psi, M_\varphi] \ge \inf_{0 \le \theta \le 1} F\left[\theta M + (1-\theta)m, \varphi^{-1}(\theta\varphi(M) + (1-\theta)\varphi(m))\right] \mathbf{1}. \tag{9.78}$$

Proof. We prove only the case (i): Using LHS of (9.69) for an operator convex function φ^{-1} and then operator monotonicity of $F(\cdot, v)$ we have

$$F[M_\varphi, M_\psi] \le F[M_1, M_\psi].$$

If we put $\psi = \iota$ the identity function and replace φ by ψ in (9.70), we obtain

$$F\left[M_1, M_\psi\right] \le \sup_{0 \le \theta \le 1} F\left[\theta M + (1-\theta)m, \psi^{-1}(\theta\psi(M) + (1-\theta)\psi(m))\right] \mathbf{1}.$$

Combining two above inequalities we have the desired inequality. □

Corollary 9.11 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and $\psi,\varphi \in \mathscr{C}[m,M]$ be strictly monotone functions.*
If φ^{-1} is convex and ψ^{-1} is concave, then

$$\begin{aligned} M_\varphi \le\ & M_\psi \\ &+ \max_{0\le\theta\le1}\left\{\theta M+(1-\theta)m-\psi^{-1}(\theta\psi(M)+(1-\theta)\psi(m))\right\}\mathbf{1} \\ &+ \max_{0\le\theta\le1}\left\{\varphi^{-1}(\theta\varphi(M)+(1-\theta)\varphi(m))-\theta M-(1-\theta)m\right\}\mathbf{1}, \end{aligned} \tag{9.79}$$

and if, additionally, $\varphi >$ and $\psi > 0$ on $[m,M]$, then

$$\begin{aligned} M_\varphi \le\ & \max_{0\le\theta\le1}\left\{\frac{\theta M+(1-\theta)m}{\psi^{-1}(\theta\psi(M)+(1-\theta)\psi(m))}\right\} \\ &\times \max_{0\le\theta\le1}\left\{\frac{\varphi^{-1}(\theta\varphi(M)+(1-\theta)\varphi(m))}{\theta M+(1-\theta)m}\right\} M_\psi. \end{aligned} \tag{9.80}$$

Proof. If we put $F(u,v)=u-v$ and $\varphi=\iota$ in (9.77), then for any concave function ψ^{-1} we have

$$M_1 - M_\psi \le \max_{0\le\theta\le1}\left\{\theta M+(1-\theta)m-\psi^{-1}(\theta\psi(M)+(1-\theta)\psi(m))\right\}\mathbf{1}.$$

Similarly, if we put $\psi=\iota$ in (9.78), then for any convex function φ^{-1} we have

$$M_1 - M_\varphi \ge \min_{0\le\theta\le1}\left\{\theta M+(1-\theta)m-\varphi^{-1}(\theta\varphi(M)+(1-\theta)\varphi(m))\right\}\mathbf{1}.$$

Combining two above inequalities we have the inequality (9.79).
We have (9.80) by a similar method. □

If we use conversions of Jensen's inequality (9.1), we obtain the following two corollaries.

Corollary 9.12 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ be as in the definition of the quasi-arithmetic mean* (9.65) *and $\psi,\varphi \in \mathscr{C}[m,M]$ be strictly monotone functions. Let $\psi\circ\varphi^{-1}$ be convex (resp. concave).*

(i) *If ψ^{-1} is operator monotone and operator subadditive (resp. operator superadditive) on $\mathbb{R}$, then*

$$M_\psi \le M_\varphi + \psi^{-1}(\beta)\mathbf{1} \qquad \left(\text{resp. } M_\psi \ge M_\varphi + \psi^{-1}(\beta)\mathbf{1}\right), \tag{9.81}$$

(i') *if $-\psi^{-1}$ is operator monotone and operator subadditive (resp. operator superadditive) on $\mathbb{R}$, then the reverse inequality is valid in* (9.73)*,*

(ii) *if ψ^{-1} is operator monotone and operator superadditive (resp. operator subadditive) on $\mathbb{R}$, then*

$$M_\psi \le M_\varphi - \varphi^{-1}(-\beta)\mathbf{1} \qquad \left(\text{resp. } M_\psi \ge M_\varphi - \varphi^{-1}(-\beta)\mathbf{1}\right), \tag{9.82}$$

(ii') *if* $-\psi^{-1}$ *is operator monotone and operator superadditive (resp. operator subadditive) on* $\mathbb{R}$*, then the reverse inequality is valid in* (9.81)*,*

where

$$\beta = \max_{0\le\theta\le1}\left\{\theta\psi(M)+(1-\theta)\psi(m)-\psi\circ\varphi^{-1}\left(\theta\varphi(M)+(1-\theta)\varphi(m)\right)\right\} \tag{9.83}$$

$$\left(resp.\ \beta = \min_{0\le\theta\le1}\left\{\theta\psi(M)+(1-\theta)\psi(m)-\psi\circ\varphi^{-1}\left(\theta\varphi(M)+(1-\theta)\varphi(m)\right)\right\}.\right)$$

Proof. We prove the case (i) only and when $\psi\circ\varphi^{-1}$ is convex: Putting $F(u,v)=u-v$ and $f=g=\psi\circ\varphi^{-1}$ in Theorem 9.11, we have:

$$\psi(M_\psi)=\int_T \frac{1}{k}\Phi_t\left(\psi\circ\varphi^{-1}\left(\varphi(x_t)\right)\right)\mathrm{d}\mu(t)\le\psi\circ\varphi^{-1}\left(\varphi(M_\varphi)\right)+\beta\mathbf{1}, \tag{9.84}$$

where

$$\beta=\max_{\varphi_m\le z\le\varphi_M}\left\{\frac{\psi(M)-\psi(m)}{\varphi(M)-\varphi(m)}(z-\varphi_m)+\psi\circ\varphi^{-1}(\varphi_m)-\psi\circ\varphi^{-1}(z)\right\}$$

which gives (9.83). Since ψ^{-1} is operator monotone and subadditive on $\mathbb{R}$, then by using (9.84) we obtain

$$M_\psi\le\psi^{-1}\left(\psi(M_\varphi)+\beta\mathbf{1}\right)\le M_\varphi+\psi^{-1}(\beta)\mathbf{1}.$$

□

Corollary 9.13 *Let* $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ *be as in the definition of the quasi-arithmetic mean* (9.65) *and* $\psi,\varphi\in\mathscr{C}[m,M]$ *be strictly monotone functions. Let* $\psi\circ\varphi^{-1}$ *be convex and* $\psi>0$ *(resp.* $\psi<0$*) on* $[m,M]$.

(i) *If* ψ^{-1} *is operator monotone and operator submultiplicative on* $\mathbb{R}$*, then*

$$M_\psi\le\psi^{-1}(\alpha)M_\varphi, \tag{9.85}$$

(i') *if* $-\psi^{-1}$ *is operator monotone and operator submultiplicative on* $\mathbb{R}$*, then the reverse inequality is valid in* (9.85)*,*

(ii) *if* ψ^{-1} *is operator monotone and operator supermultiplicative on* $\mathbb{R}$*, then*

$$M_\psi\le\left[\psi^{-1}(\alpha^{-1})\right]^{-1}M_\varphi, \tag{9.86}$$

(ii') *if* $-\psi^{-1}$ *is operator monotone and operator supermultiplicative on* $\mathbb{R}$*, then the reverse inequality is valid in* (9.86)*,*

where

$$\alpha=\max_{0\le\theta\le1}\left\{\frac{\theta\psi(M)+(1-\theta)\psi(m)}{\psi\circ\varphi^{-1}\left(\theta\varphi(M)+(1-\theta)\varphi(m)\right)}\right\} \tag{9.87}$$

$$\left(resp.\quad \alpha=\min_{0\le\theta\le1}\left\{\frac{\theta\psi(M)+(1-\theta)\psi(m)}{\psi\circ\varphi^{-1}\left(\theta\varphi(M)+(1-\theta)\varphi(m)\right)}\right\}.\quad\right)$$

Proof. The proof is essentially the same as that of Corollary 9.12 and we omit details. □

Remark 9.8 *We note that if $\psi \circ \varphi^{-1}$ is a concave function, we can obtain similar inequalities as in Corollary 9.13. We use the same way as we did in Corollary 9.12.*
E.g. if $\psi > 0$ (resp. $\psi < 0$) on $[m,M]$ is operator monotone and operator supermultiplicative on $\mathbb{R}$, then

$$M_\psi \geq \psi^{-1}(\alpha) M_\varphi,$$

with min *instead of* max *in* (9.87)*.*

Example 9.2 *If we put $\varphi(t) = t^s$ and $\psi(t) = t^r$ in inequalities involving the complementary order among quasi-arithmetic means, we can obtain the complementary order among power means.*
E.g. using Corollary 9.10, we obtain that (compare with Theorem 9.14)

$$M_s(\mathbf{x},\mathbf{\Phi}) \leq k^{\frac{r-s}{rs}} \max_{0\leq\theta\leq 1} \left\{ \frac{\sqrt[r]{(\theta M^r + (1-\theta)m^r)}}{\sqrt[s]{(\theta M^s + (1-\theta)m^s)}} \right\} M_r(\mathbf{x},\mathbf{\Phi})$$

holds for $r \leq s$, $s \geq 1$ or $r \leq s \leq -1$, where

$$\max_{0\leq\theta\leq 1} \left\{ \frac{\sqrt[r]{(\theta M^r + (1-\theta)m^r)}}{\sqrt[s]{(\theta M^s + (1-\theta)m^s)}} \right\} = \Delta(h,r,s)$$

is the generalized Specht ratio defined by (9.3)*, i.e.*

$$\Delta(h,r,s) = \left\{ \frac{r(h^s - h^r)}{(s-r)(h^r - 1)} \right\}^{\frac{1}{s}} \left\{ \frac{s(h^r - h^s)}{(r-s)(h^s - 1)} \right\}^{-\frac{1}{r}}, \quad h = \frac{M}{m}.$$

9.6 Some better bounds

In this section we study converses of a generalized Jensen's inequality for a continuous field of self-adjoint operators, a unital field of positive linear mappings and real values continuous convex functions. We obtain some better bounds than the ones calculated in Section 9.1 and a series of papers in which these inequalities are studied. As an application, we provide a refined calculation of bounds in the case of power functions.

In the following theorem we give a general form of converses of Jensen's inequality which give a better bound than the one in Theorem 9.2.

Theorem 9.24 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with the spectra in $[m,M]$, $m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let m_x and M_x, $m_x \le M_x$, be the bounds of the self-adjoint element $x = \int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)$ and $f:[a,b]\to\mathbb{R}$, $g:[m_x,M_x]\to\mathbb{R}$, $F: U\times V\to\mathbb{R}$, where $f([a,b])\subseteq U$, $g([m_x,M_x])\subseteq V$ and F be bounded.*

If f is convex and F is operator monotone in the first variable, then

$$F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\,,\, g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right] \le C_1\,\mathbf{1}_K \le C\,\mathbf{1}_K, \tag{9.88}$$

where constants $C_1 \equiv C_1(F,f,g,m,M,m_x,M_x)$ and $C\equiv C(F,f,g,m,M)$ are

$$\begin{aligned} C_1 &:= \sup_{m_x\le z\le M_x}\left\{F\left[\frac{M-z}{M-m}f(m)+\frac{z-m}{M-m}f(M),g(z)\right]\right\} \\ &= \sup_{\frac{M-M_x}{M-m}\le p\le \frac{M-m_x}{M-m}} \{F[pf(m)+(1-p)f(M)\,,\,g(pm+(1-p)M)]\}, \\ C &:= \sup_{m\le z\le M}\left\{F\left[\frac{M-z}{M-m}f(m)+\frac{z-m}{M-m}f(M),g(z)\right]\right\} \\ &= \sup_{0\le p\le 1}\{F[pf(m)+(1-p)f(M)\,,\,g(pm+(1-p)M)]\}. \end{aligned}$$

If f is concave, then the opposite inequality holds in (9.88) *with* inf *instead of* sup *in bounds C_1 and C.*

Proof. We prove only the convex case. Since $m\Phi_t(\mathbf{1}_H)\le\Phi_t(x_t)\le M\Phi_t(\mathbf{1}_H)$ and $\int_T\Phi_t(\mathbf{1}_H)\mathrm{d}\mu(t)=\mathbf{1}_K$, then $m\mathbf{1}_K\le\int_T\Phi_t(x_t)\mathrm{d}\mu(t)\le M\mathbf{1}_K$. Next, since m_x and M_x, are the bounds of the operator $\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)$ it follows that $[m_x,M_x]\subseteq[m,M]$.

By using convexity of f and functional calculus, we obtain

$$\begin{aligned} \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) &\le \int_T \Phi_t\left(\frac{M\mathbf{1}_H-x_t}{M-m}f(m)+\frac{x_t-m\mathbf{1}_H}{M-m}f(M)\right)\mathrm{d}\mu(t) \\ &= \frac{M\mathbf{1}_K-\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)}{M-m}f(m)+\frac{\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)-m\mathbf{1}_K}{M-m}f(M). \end{aligned}$$

Using operator monotonicity of $u\mapsto F(u,v)$ and boundedness of F, it follows

$$\begin{aligned} &F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\,,\, g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right] \\ &\le F\left[\frac{M\mathbf{1}_K-\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)}{M-m}f(m)+\frac{\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)-m\mathbf{1}_K}{M-m}f(M)\,,\, g\left(\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right] \\ &\le \sup_{m_x\le z\le M_x}\left\{F\left[\frac{M-z}{M-m}f(m)+\frac{z-m}{M-m}f(M),g(z)\right]\right\}\mathbf{1}_K \\ &\le \sup_{m\le z\le M}\left\{F\left[\frac{M-z}{M-m}f(m)+\frac{z-m}{M-m}f(M),g(z)\right]\right\}\mathbf{1}_K. \end{aligned}$$

□

Remark 9.9 *We can obtain an inequality similar to the one in Theorem 9.24 in the case when $(\Phi_t)_{t\in T}$ is a non-unit field of positive linear mappings, i.e. when $\int_T \Phi_t(\mathbf{1})\,\mathrm{d}\mu(t) = k\mathbf{1}$ for some positive scalar k. Then,*

$$\begin{aligned} &F\left[\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t)\,,\, g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)\right] \\ &\le \sup_{km_x\le z\le kM_x}\left\{F\left[\frac{kM-z}{M-m}f(m)+\frac{z-km}{M-m}f(M)\,,\,g(z)\right]\right\}\mathbf{1}_K \\ &\le \sup_{km\le z\le kM}\left\{F\left[\frac{kM-z}{M-m}f(m)+\frac{z-km}{M-m}f(M)\,,\,g(z)\right]\right\}\mathbf{1}_K. \end{aligned}$$

This means that we obtain a better upper bound than the one given in Theorem 9.11.

We recall that the following generalization of Jensen's inequality (9.1) holds. If f is an operator convex function on $[m,M]$ and $\lambda g \le f$ on $[m,M]$ for some function g and real number λ, then

$$0 \le \int_T \Phi_t\left(f(x_t)\right)\mathrm{d}\mu(t) - \lambda g\left(\int_T \Phi_t(x_t)\mathrm{d}\mu(t)\right).$$

In the following we consider the difference type converses of the above inequality.

We introduce some abbreviations. Let $f:[m,M]\to\mathbb{R}$, $m<M$, be a convex or a concave function. We denote a linear function through $(m,f(m))$ and $(M,f(M))$ by $f^{cho}_{[m,M]}$, i.e.

$$f^{cho}_{[m,M]}(z) = \frac{M-z}{M-m}f(m)+\frac{z-m}{M-m}f(M), \qquad z\in\mathbb{R}$$

and the slope and the intercept by α_f and β_f as in (9.2).

The following Theorem 9.25 and Corollary 9.14 are refinements of [124, Theorem 2.4].

Theorem 9.25 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with the spectra in $[m,M]$, $m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t:\mathscr{A}\to\mathscr{B}$ from $\mathscr{A}$ to another unital C^*−algebra $\mathscr{B}$. Let m_x and M_x, $m_x\le M_x$, be the bounds of $x=\int_T\Phi_t(x_t)\,\mathrm{d}\mu(t)$ and $f:[m,M]\to\mathbb{R}$, $g:[m_x,M_x]\to\mathbb{R}$ be continuous functions.*

If f is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) - \lambda g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le \max_{m_x\le z\le M_x}\left\{\alpha_f z+\beta_f-\lambda g(z)\right\}\mathbf{1}_K \tag{9.89}$$

holds and the bound in RHS of (9.89) *exists for any m,M,m_x and M_x.*

If f is concave, then the reverse inequality with min *instead of* max *is valid in* (9.89). *The bound in RHS of this inequality exists for any m,M,m_x and M_x.*

Proof. We put $F(u,v) = u - \lambda v$, $\lambda \in \mathbb{R}$ in Theorem 9.24. A function $z \mapsto \alpha_f z + \beta_f - \lambda g(z)$ is continuous on $[m_x, M_x]$, so the global extremes exist. □

In the following corollary we give a way of determining the bounds placed in Theorem 9.25.

Corollary 9.14 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$, A, f and g be as in Theorem 9.25.*

(i) *Let $\lambda \le 0$.*

If f is convex and g is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) - \lambda g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le C_\lambda \mathbf{1}_K \tag{9.90}$$

holds with

$$C_\lambda = \max\left\{ f^{cho}_{[m,M]}(m_x) - \lambda g(m_x),\, f^{cho}_{[m,M]}(M_x) - \lambda g(M_x) \right\}. \tag{9.91}$$

But, if f is convex and g is concave, then the inequality (9.90) *holds with*

$$C_\lambda = \begin{cases} f^{cho}_{[m,M]}(m_x) - \lambda g(m_x) & \text{if } \lambda g'_-(z) \ge \alpha_f \text{ for every } z \in (m_x, M_x), \\ f^{cho}_{[m,M]}(z_0) - \lambda g(z_0) & \text{if } \lambda g'_-(z_0) \le \alpha_f \le \lambda g'_+(z_0) \text{ for some } z_0 \in (m_x, M_x), \\ f^{cho}_{[m,M]}(M_x) - \lambda g(M_x) & \text{if } \lambda g'_+(z) \le \alpha_f \text{ for every } z \in (m_x, M_x). \end{cases} \tag{9.92}$$

If f is concave and g is convex, then

$$c_\lambda \mathbf{1}_K \le \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) - \lambda g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \tag{9.93}$$

holds with c_λ which equals the right side in (9.92) *with reverse inequality signs.*

But, if f is concave and g is concave, then the inequality (9.93) *holds with c_λ which equals the right side in* (9.91) *with* min *instead of* max.

(ii) *Let $\lambda \ge 0$.*

If f is convex and g is convex, then the inequality (9.90) *holds with C_λ defined by* (9.92). *But if f is convex and g is concave, then* (9.90) *holds with C_λ defined by* (9.91).

If f is concave and g is convex, then the inequality (9.93) *holds with c_λ which equals the right side in* (9.91) *with* min *instead of* max. *But, if f is concave and g is concave, then* (9.93) *holds with c_λ which equals the right side in* (9.92) *with reverse inequality signs.*

Proof. (i): We prove only the cases when f is convex. If g is convex (resp. concave) we apply Proposition 9.2 (resp. Proposition 9.1) on the convex (resp. concave) function $h_\lambda = f^{cho}_{[m,M]}(z) - \lambda g(z)$, and get (9.91) (resp. (9.92)).

In the remaining cases the proof is essentially the same as in the above cases. □

Corollary 9.14 applied on the functions $f(z) = z^p$ and $g(z) = z^q$ gives the following corollary, which is a refinement of [124, Corollary 2.6].

Corollary 9.15 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.25, and additionally let operators x_t be strictly positive with the spectra in $[m,M]$, where $0<m<M$.*

(i) *Let $\lambda \le 0$.*

If $p,q \in (-\infty,0]\cup[1,\infty)$, then

$$\int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) - \lambda\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^q \le C_\lambda^\star \mathbf{1}_K \tag{9.94}$$

holds with

$$C_\lambda^\star = \max\{\alpha_{t^p} m_x + \beta_{t^p} - \lambda m_x^q,\, \alpha_{t^p} M_x + \beta_{t^p} - \lambda M_x^q\}. \tag{9.95}$$

If $p \in (-\infty,0)$ and $q\in(0,1)$, then the inequality (9.94) *holds with*

$$C_\lambda^\star = \begin{cases} \alpha_{t^p}\, m_x + \beta_{t^p} - \lambda\, m_x^q & \text{if } (\lambda\, q/\alpha_{t^p})^{1/(1-q)} \le m_x, \\ \beta_{t^p} + \lambda(q-1)\,(\lambda\, q/\alpha_{t^p})^{q/(1-q)} & \text{if } m_x \le (\lambda\, q/\alpha_{t^p})^{1/(1-q)} \le M_x, \\ \alpha_{t^p}\, M_x + \beta_{t^p} - \lambda\, M_x^q & \text{if } (\lambda\, q/\alpha_{t^p})^{1/(1-q)} \ge M_x. \end{cases} \tag{9.96}$$

If $p\in(0,1)$ and $q\in(-\infty,0)$, then

$$c_\lambda^\star \mathbf{1}_K \le \int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) - \lambda\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^q \tag{9.97}$$

holds with $c_\lambda^\star$ which equals the right side in (9.96).

If $p,q\in[0,1]$, then the inequality (9.97) *holds with $c_\lambda^\star$ which equals the right side in* (9.95) *with* min *instead of* max.

(ii) *Let $\lambda \ge 0$.*

If $p,q\in(-\infty,0)\cup(1,\infty)$, then (9.94) *holds with $C_\lambda^\star$ defined by* (9.96)*. But, if $p\in(-\infty,0]\cup[1,+\infty)$ and $q\in[0,1]$, then* (9.94) *holds with $C_\lambda^\star$ defined by* (9.95).

If $p\in[0,1]$ and $q\in(-\infty,0]\cup[1,\infty)$, then (9.97) *holds with $c_\lambda^\star$ which equals the right side in* (9.95) *with* min *instead of* max. *But, if $p\in(0,1)$ and $q\in(0,1)$, then* (9.97) *holds with $c_\lambda^\star$ which equals the right side in* (9.96).

Using Theorem 9.25 and Corollary 9.14 with $g = f$ and $\lambda = 1$ we have the following theorem.

Theorem 9.26 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with the spectra in $[m,M]$, $m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A}\to\mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let m_x and M_x, $m_x \le M_x$, be the bounds of $x = \int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)$ and $f:[m,M]\to\mathbb{R}$ be a continuous function.*

If f is convex, then

$$0 \le \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) - f\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le \max_{m_x \le z \le M_x} \left\{ f^{cho}_{[m,M]}(z) - f(z) \right\} \mathbf{1}_K \quad (9.98)$$

holds and the bound in RHS of (9.98) *exists for any m, M, m_x and M_x.*

The value of the constant

$$\bar{C} \equiv \bar{C}(f, m, M, m_x, M_x) := \max_{m_x \le z \le M_x} \left\{ f^{cho}_{[m,M]}(z) - f(z) \right\}$$

can be determined as follows

$$\bar{C} = \begin{cases} f^{cho}_{[m,M]}(m_x) - f(m_x) & \text{if } f'_-(z) \ge \alpha_f \text{ for every } z \in (m_x, M_x), \\ f^{cho}_{[m,M]}(z_0) - f(z_0) & \text{if } g'_-(z_0) \le \alpha_f \le g'_+(z_0) \text{ for some } z_0 \in (m_x, M_x), \\ f^{cho}_{[m,M]}(M_x) - f(M_x) & \text{if } g'_+(z) \le \alpha_f \text{ for every } z \in (m_x, M_x). \end{cases} \quad (9.99)$$

If f is concave, then the reverse inequality with min *instead of* max *is valid in* (9.98)*. The bound in this inequality exists for any m, M, m_x and M_x. The value of the constant*

$$\bar{c} \equiv \bar{c}(f, m, M, m_x, M_x) := \min_{m_x \le z \le M_x} \left\{ f^{cho}_{[m,M]}(z) - f(z) \right\}$$

can be determined as in the right side in (9.99) *with reverse inequality signs.*

If f is a strictly convex differentiable function on $[m_x, M_x]$, then we obtain the following corollary of Theorem 9.26. This is a refinement of [124, Corollary 2.16].

Corollary 9.16 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.26. Let $f : [m,M] \to \mathbb{R}$ be a continuous function. If f is strictly convex differentiable on $[m_x, M_x]$, then*

$$0 \le \int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) - f\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \le \left(\alpha_f z_0 + \beta_f - f(z_0)\right) \mathbf{1}_K, \quad (9.100)$$

where

$$z_0 = \begin{cases} m_x & \text{if } f'(m_x) \ge \alpha_f, \\ f'^{-1}(\alpha_f) & \text{if } f'(m_x) \le \alpha_f \le f'(M_x), \\ M_x & \text{if } f'(M_x) \le \alpha_f. \end{cases} \quad (9.101)$$

The global upper bound is $C(m,M,f) = \alpha_f \bar{z}_0 + \beta_f - f(\bar{z}_0)$, where $\bar{z}_0 = (f')^{-1}(\alpha_f) \in (m,M)$. The upper bound in RHS of (9.100) *is better than the global upper bound provided that either $f'(m_x) \ge \alpha_f$ or $f'(M_x) \le \alpha_f$.*

In the dual case, when f is strictly concave differentiable on $[m_x, M_x]$, then the reverse inequality is valid in (9.100)*, with z_0 which equals the right side in* (9.101) *with reverse inequality signs. The global lower bound is defined as the global upper bound in the convex case. The lower bound in the reverse inequality in* (9.100) *is better than the global lower bound provided that either $f'(m_x) \le \alpha_f$ or $f'(M_x) \ge \alpha_f$.*

Proof. We prove only the cases when f is strictly convex differentiable on $[m_x, M_x]$. The inequality (9.100) follows from Theorem 9.26 by using the differential calculus. Since $h(z) = \alpha_f z + \beta_f - f(z)$ is a continuous strictly concave function on $[m, M]$, then there is exactly one point $z_0 \in [m, M]$ which achieves the global maximum. If neither of these points is in the interval $[m_x, M_x]$, then the global maximum in $[m_x, M_x]$ is less than the global maximum in $[m, M]$. □

Using Corollary 9.15 with $q = p$, $\lambda = 1$ or applying Corollary 9.16 we have the following corollary, which is a refinement of [124, Corollary 2.18].

Corollary 9.17 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.26, and additionally let operators x_t be strictly positive with the spectra in $[m, M]$, where $0 < m < M$. Then*

$$0 \le \int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) - \left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p \le \overline{C}(m_x, M_x, m, M, p)\mathbf{1}_K \le C(m, M, p)\mathbf{1}_K$$

for $p \notin (0,1)$, and

$$0 \ge \int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) - \left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p \ge \overline{c}(m_x, M_x, m, M, p)\mathbf{1}_K \ge C(m, M, p)\mathbf{1}_K$$

for $p \in (0,1)$, where

$$\overline{C}(m_x, M_x, m, M, p) = \begin{cases} \alpha_{t^p}\, m_x + \beta_{t^p} - m_x^p & \text{if } \; p m_x^{p-1} \ge \alpha_{t^p}, \\ C(m, M, p) & \text{if } \; p m_x^{p-1} \le \alpha_{t^p} \le p M_x^{p-1}, \\ \alpha_{t^p}\, M_x + \beta_{t^p} - M_x^p & \text{if } \; p M_x^{p-1} \le \alpha_{t^p}, \end{cases} \tag{9.102}$$

and $\overline{c}(m_x, M_x, m, M, p)$ equals the right side in (9.102) *with reverse inequality signs. The constant $C(m, M, p)$ is defined by* (2.38)*.*

In the same way in the following we consider the ratio type converses of Jensen's inequality. The following Theorem 9.27 and Corollary 9.18 are refinements of [124, Theorem 2.9].

Theorem 9.27 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with the spectra in $[m, M]$, $m < M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let m_x and M_x, $m_x \le M_x$, be the bounds of $x = \int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)$ and $f : [m, M] \to \mathbb{R}$ be a continuous function and $g : [m_x, M_x] \to \mathbb{R}$ be a strictly positive continuous function.*

If f is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \le \max_{m_x \le z \le M_x} \left\{ \frac{\alpha_f z + \beta_f}{g(z)} \right\} g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \tag{9.103}$$

holds and the bound in RHS of (9.103) *exists for any m, M, m_x and M_x.*

If f is concave, then the reverse inequality with min *instead of* max *is valid in* (9.103)*. The bound in RHS of this inequality exists for any m, M, m_x and M_x.*

Proof. We put $F(u,v) = v^{-\frac{1}{2}} u v^{-\frac{1}{2}}$ in Theorem 9.24.
A function $z \mapsto \frac{\alpha_f z+\beta_f}{g(z)}$ is continuous on $[m_x, M_x]$, so the global extremes exist. □

Remark 9.10 *If f is convex and g is strictly negative on $[m_x, M_x]$, then the inequality with* min *instead of* max *is valid in* (9.103)*. If f is concave and g is strictly negative on $[m_x, M_x]$, then the reverse inequality is valid in* (9.103)*.*

In the following corollary, we give a way of determining the bounds placed in Theorem 9.27.

Corollary 9.18 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$, A, f and g be as in Theorem 9.27. Additionally, let $f^{cho}_{[m,M]}$ and g be strictly positive on $[m_x, M_x]$.*

If f is convex and g is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \le C\,g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \tag{9.104}$$

holds with

$$C = \begin{cases} \dfrac{f^{cho}_{[m,M]}(m_x)}{g(m_x)} & \text{if } g'_-(z) \ge \dfrac{\alpha_f g(z)}{\alpha_f z+\beta_f} \text{ for every } z \in (m_x, M_x), \\ \dfrac{f^{cho}_{[m,M]}(z_0)}{g(z_0)} & \text{if } g'_-(z_0) \le \dfrac{\alpha_f g(z_0)}{\alpha_f z_0+\beta_f} \le g'_+(z_0) \text{ for some } z_0 \in (m_x, M_x), \\ \dfrac{f^{cho}_{[m,M]}(M_x)}{g(M_x)} & \text{if } g'_+(z) \le \dfrac{\alpha_f g(z)}{\alpha_f z+\beta_f} \text{ for every } z \in (m_x, M_x). \end{cases} \tag{9.105}$$

If f is convex and g is concave, then the inequality (9.104) *holds with*

$$C = \max\left\{ \frac{f^{cho}_{[m,M]}(m_x)}{g(m_x)}, \frac{f^{cho}_{[m,M]}(M_x)}{g(M_x)} \right\}. \tag{9.106}$$

If f is concave and g is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \ge c\,g\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \tag{9.107}$$

holds with c which equals the right side in (9.106) *with* min *instead of* max*.*

If f is concave and g is concave, then the inequality (9.107) *holds with c which equals the right side in* (9.105) *with reverse inequality signs.*

Proof. We prove only the cases when f is convex. If g is convex (resp. concave) we apply Proposition 9.3 (resp. Proposition 9.5) on the ratio function $h(z) = \frac{f^{cho}_{[m,M]}(z)}{g(z)}$ with the convex (resp. concave) denominator g, and so we get (9.105) (resp. (9.106)). □

Corollary 9.18 applied on the functions $f(z) = z^p$ and $g(z) = z^q$ gives the following corollary, which is a refinement of [124, Corollary 2.11].

Corollary 9.19 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.27, and additionally let operators x_t be strictly positive with the spectra in $[m,M]$, where $0<m<M$.*

If $p,q\in(-\infty,0)\cup(1,\infty)$, then

$$\int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) \le C^\star \left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^q \tag{9.108}$$

holds with

$$C^\star = \begin{cases} \dfrac{\alpha_{t^p}\, m_x + \beta_{t^p}}{m_x^q} & \text{if } \dfrac{q}{1-q}\dfrac{\beta_{t^p}}{\alpha_{t^p}} \le m_x, \\ \dfrac{\beta_{t^p}}{1-q}\left(\dfrac{1-q}{q}\dfrac{\alpha_{t^p}}{\beta_{t^p}}\right)^q & \text{if } m_x \le \dfrac{q}{1-q}\dfrac{\beta_{t^p}}{\alpha_{t^p}} \le M_x, \\ \dfrac{\alpha_{t^p}\, M_x + \beta_{t^p}}{M_x^q} & \text{if } \dfrac{q}{1-q}\dfrac{\beta_{t^p}}{\alpha_{t^p}} \ge M_x. \end{cases} \tag{9.109}$$

If $p\in(-\infty,0]\cup[1,\infty)$ and $q\in[0,1]$, then the inequality (9.108) *holds with*

$$C^\star = \max\left\{ \frac{\alpha_{t^p} m_x + \beta_{t^p}}{m_x^q}, \frac{\alpha_{t^p} M_x + \beta_{t^p}}{M_x^q} \right\}. \tag{9.110}$$

If $p\in[0,1]$ and $q\in(-\infty,0]\cup[1,\infty)$, then

$$\int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) \ge c^\star \left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^q \tag{9.111}$$

holds with c_λ which equals the right side in (9.110) *with* min *instead of* max.

If $p,q\in(0,1)$, then the inequality (9.111) *holds with $c^\star$ which equals the right side in* (9.109).

Using Theorem 9.27, Proposition 9.4 and 9.6 with $g=f$ we have the following theorem.

Theorem 9.28 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ with the spectra in $[m,M]$, $m<M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ, and let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t:\mathscr{A}\to\mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let m_x and M_x, $m_x\le M_x$, be the bounds of $x=\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)$. let $f:[m,M]\to\mathbb{R}$ be a continuous function and strictly positive on $[m_x,M_x]$.*

If f is convex, then

$$\int_T \Phi_t(f(x_t))\,\mathrm{d}\mu(t) \le \max_{m_x\le z\le M_x}\left\{\frac{f^{cho}_{[m,M]}(z)}{f(z)}\right\} f\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right) \tag{9.112}$$

holds and the bound in RHS of (9.112) *exists for any m,M,m_x and M_x.*

The value of the constant

$$\bar{C} \equiv \bar{C}(f,m,M,m_x,M_x) := \max_{m_x\le z\le M_x}\left\{\frac{f^{cho}_{[m,M]}(z)}{f(z)}\right\}$$

can be determined as follows:

$$\overline{C} = \begin{cases} \dfrac{f^{cho}_{[m,M]}(m_x)}{f(m_x)} & \text{if } f'_-(z) \geq \dfrac{\alpha_f f(z)}{\alpha_f z + \beta_f} \text{ for every } z \in (m_x, M_x), \\ \dfrac{f^{cho}_{[m,M]}(z_0)}{f(z_0)} & \text{if } f'_-(z_0) \leq \dfrac{\alpha_f f(z_0)}{\alpha_f z_0 + \beta_f} \leq f'_+(z_0) \text{ for some } z_0 \in (m_x, M_x), \\ \dfrac{f^{cho}_{[m,M]}(M_x)}{f(M_x)} & \text{if } f'_+(z) \leq \dfrac{\alpha_f f(z)}{\alpha_f z + \beta_f} \text{ for every } z \in (m_x, M_x). \end{cases} \tag{9.113}$$

If f is concave, then the reverse inequality with min *instead of* max *is valid in* (9.112)*. The bound in this inequality exists for any m, M, m_x and M_x. The value of the constant*

$$\overline{c} \equiv \overline{c}(f, m, M, m_x, M_x) := \min_{m_x \leq z \leq M_x} \left\{ \frac{f^{cho}_{[m,M]}(z)}{f(z)} \right\}$$

can be determined as in the right side in (9.112) *with reverse inequality signs.*

Remark 9.11 *If f is convex and strictly negative on $[m_x, M_x]$, then the inequality with* min *instead of* max *is valid in* (9.112)*. If f is concave and strictly negative on $[m_x, M_x]$, then the reverse inequality is valid in* (9.112)*.*

If f is a strictly convex differentiable function on $[m_x, M_x]$, then we obtain the following corollary of Theorem 9.28. This is a refinement of [124, Corollary 2.10].

Corollary 9.20 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.28. Let $f : [m, M] \to \mathbb{R}$ be a continuous function and $f(m), f(M) > 0$. If f is strictly positive and strictly convex twice differentiable on $[m_x, M_x]$, then*

$$\int_T \Phi_t(f(x_t))\, d\mu(t) \leq \left(\frac{\alpha_f z_0 + \beta_f}{f(z_0)} \right) f\left(\int_T \Phi_t(x_t)\, d\mu(t) \right), \tag{9.114}$$

where $z_0 \in (m_x, M_x)$ is defined as the unique solution of $\alpha_f f(z) = (\alpha_f z + \beta_f) f'(z)$ provided $(\alpha_f m_x + \beta_f) f'(m_x)/f(m_x) \leq \alpha_f \leq (\alpha_f M_x + \beta_f) f'(M_x)/f(M_x)$, otherwise z_0 is defined as m_x or M_x provided $\alpha_f \leq (\alpha_f m_x + \beta_f) f'(m_x)/f(m_x)$ or $\alpha_f \geq (\alpha_f M_x + \beta_f) f'(M_x)/f(M_x)$, respectively.

The global upper bound is $C(m, M, f) = (\alpha_f \overline{z}_0 + \beta_f)/f(\overline{z}_0)$, where $\overline{z}_0 \in (m, M)$ is defined as the unique solution of $\alpha_f f(z) = (\alpha_f z + \beta_f) f'(z)$. The upper bound in RHS of (9.114) *is better than the global upper bound provided that either $\alpha_f \leq (\alpha_f m_x + \beta_f) f'(m_x)/f(m_x)$ or $\alpha_f \geq (\alpha_f M_x + \beta_f) f'(M_x)/f(M_x)$.*

In the dual case, when f is positive and strictly concave differentiable on $[m_x, M_x]$, then the reverse inequality is valid in (9.114)*, with z_0 is defined as in* (9.114) *with reverse inequality signs. The global lower bound is defined as the global upper bound in the convex case. The lower bound in the reverse inequality in* (9.114) *is better than the global lower bound provided that either $\alpha_f \geq (\alpha_f m_x + \beta_f) f'(m_x)/f(m_x)$ or $\alpha_f \leq (\alpha_f M_x + \beta_f) f'(M_x)/f(M_x)$.*

Proof. We prove only the cases when f is strictly convex differentiable on $[m_x,M_x]$. The inequality (9.114) follows from Theorem 9.28 by using the differential calculus.

Next, we put $h(z) = (\alpha_f z+\beta_f)/f(z)$. Then $h'(z) = H(z)/f(z)^2$, where $H(z) = \alpha_f f(z) - (\alpha_f z+\beta_f)f'(z)$. Due to the strict convexity of f on $[m_x,M_x]$ and since $f(m), f(M) > 0$, it follows that $H'(z) = -(\alpha_f z+\beta_f)f''(z) < 0$. Hence $H(z)$ is decreasing on $[m_x,M_x]$. If $H(m_x)H(M_x) \le 0$, then the minimum value of the function h on $[m_x,M_x]$ is attained in z_0 which is the unique solution of the equation $H(z) = 0$. Otherwise, if $H(m_x)H(M_x) \ge 0$, then this minimum value is attained in m_x or M_x according to $H(m_x) \le 0$ or $H(M_x) \ge 0$.

Since $h(z) = (\alpha_f z+\beta_f)/f(z)$ is a continuous function on $[m,M]$, then the global maximum in $[m_x,M_x]$ is less than the global maximum in $[m,M]$. □

Using Corollary 9.19 with $q = p$ or applying Corollary 9.20 we have the following corollary, which is a refinement of [124, Corollary 2.12].

Corollary 9.21 *Let $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$ and x be as in Theorem 9.28, and additionally let operators x_t be strictly positive with the spectra in $[m,M]$, where $0 < m < M$. Then*

$$\int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) \le \overline{K}(m_x,M_x,m,M,p)\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p \le K(m,M,p)\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p$$

for $p \notin (0,1)$, and

$$\int_T \Phi_t(x_t^p)\,\mathrm{d}\mu(t) \ge \overline{k}(m_x,M_x,m,M,p)\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p \ge K(m,M,p)\left(\int_T \Phi_t(x_t)\,\mathrm{d}\mu(t)\right)^p$$

for $p \in (0,1)$, where

$$\overline{K}(m_x,M_x,m,M,p) = \begin{cases} \dfrac{\alpha_{t^p}\, m_x + \beta_{t^p}}{m_x^p} & \text{if } p\,\beta_{t^p}/m_x \ge (1-p)\,\alpha_{t^p}, \\ K(m,M,p) & \text{if } p\,\beta_{t^p}/m_x < (1-p)\,\alpha_{t^p} < p\,\beta_{t^p}/M_x, \\ \dfrac{\alpha_{t^p}\, M_x + \beta_{t^p}}{M_x^p} & \text{if } p\,\beta_{t^p}/M_x \le (1-p)\,\alpha_{t^p}, \end{cases} \tag{9.115}$$

and $\overline{k}(m_x,M_x,m,M,p)$ equals the right side in (9.115) with reverse inequality signs. $K(m,M,p)$ is the Kantorovich constant defined by (2.29).

Remark 9.12 *We can obtain similar inequalities to above in the case when $(\Phi_t)_{t\in T}$ is a field of positive linear mappings such that $\int_T \Phi_t(\mathbf{1})\,\mathrm{d}\mu(t) = k\mathbf{1}$ for some positive scalar k. The details are left to the interested reader.*

9.7 Appendix

In appendix of this section we give the calculation of extreme values of a difference or ratio function $y = h(z)$, of a linear function $y = kx + l$ and a continuous convex or concave function $y = g(x)$ on a closed interval. The basic facts about the convex and concave functions can be found e.g. in books [239, 249].

We first examine two cases for the difference.

Proposition 9.1 *Let $g : [a,b] \to \mathbb{R}$ be a continuous function and let $h(z) = kz + l - g(z)$ be a difference function. If g is convex, then*

$$\min_{a \le z \le b} h(z) = \min\{h(a)\,, h(b)\} \tag{9.116}$$

and

$$\max_{a \le z \le b} h(z) = \begin{cases} h(a) & \text{if } g'_-(z) \ge k \text{ for every } z \in (a,b), \\ h(z_0) & \text{if } g'_-(z_0) \le k \le g'_+(z_0) \text{ for some } z_0 \in (a,b), \\ h(b) & \text{if } g'_+(z) \le k \text{ for every } z \in (a,b). \end{cases} \tag{9.117}$$

Additionally, if g is strictly convex and h is not monotone, then a unique number $z_0 \in (a,b)$ exists so that

$$h(z_0) = \max_{a \le z \le b} h(z). \tag{9.118}$$

Proof. A function $y = h(z)$ is continuously concave because it is the sum of two continuous concave functions $y = kz + l$ and $y = -g(z)$. Since a function h is lower bounded by the chord line through endpoints $P_a(a, h(a))$ and $P_b(b, h(b))$, then (9.116) holds. Next, (9.117) follows from the global maximum property for concave functions. With additional assumptions the equality (9.118) follows from the strict concavity of h. □

Proposition 9.2 *Let $g : [a,b] \to \mathbb{R}$ be a continuous function and let $h(z) = kz + l - g(z)$ be a difference function. If g is concave, then*

$$\max_{a \le z \le b} h(z) = \max\{h(a)\,, h(b)\}$$

and

$$\min_{a \le z \le b} h(z) = \begin{cases} h(a) & \text{if } g'_-(z) \le k \text{ for every } z \in (a,b), \\ h(z_0) & \text{if } g'_+(z_0) \le k \le g'_-(z_0) \text{ for some } z_0 \in (a,b), \\ h(b) & \text{if } g'_+(z) \ge k \text{ for every } z \in (a,b). \end{cases}$$

Additionally, if g is strictly concave and h is not monotone, then a unique number $z_0 \in (a,b)$ exists so that

$$h(z_0) = \min_{a \le z \le b} h(z).$$

Proof. The proof is essentially the same as the one in Proposition 9.1. □

We now examine four cases for the ratio.

Proposition 9.3 *Let $g : [a,b] \to \mathbb{R}$ be either a strictly positive or strictly negative continuous function and let $h(z) = (kz+l)/g(z)$ be a ratio function with strictly positive numerator. If g is convex, then*

$$\min_{a \le z \le b} h(z) = \min\{h(a), h(b)\} \tag{9.119}$$

and

$$\max_{a \le z \le b} h(z) = \begin{cases} h(a) & \textit{if } g'_-(z) \ge \dfrac{kg(z)}{kz+l} \textit{ for every } z \in (a,b), \\ h(z_0) & \textit{if } g'_-(z_0) \le \dfrac{kg(z_0)}{kz_0+l} \le g'_+(z_0) \textit{ for some } z_0 \in (a,b), \\ h(b) & \textit{if } g'_+(z) \le \dfrac{kg(z)}{kz+l} \textit{ for every } z \in (a,b). \end{cases} \tag{9.120}$$

Additionally, if g is strictly convex and h is not monotone, then a unique number $z_0 \in (a,b)$ exists so that

$$h(z_0) = \max_{a \le z \le b} h(z). \tag{9.121}$$

Proof. Maximum value: A function $y = h(z)$ is continuous on $[a,b]$ because it is the ratio of two continuous functions. Then there exists $z_0 \in [a,b]$ such that $h(z_0) = \max_{a \le z \le b} h(z)$. Also, since g is convex, then $g'_-(z)$ and $g'_+(z)$ exist and $g'_-(z) \le g'_+(z)$ on (a,b). Then h'_- and h'_+ exist and

$$h'_{\mp}(z) = \frac{k\,g(z) - (k\,z+l)\,g'_{\mp}(z)}{(g(z))^2}.$$

First we observe the case when h is not monotone on $[a,b]$. Then there exists $z_0 \in (a,b)$ such that $h(z_0) = \max_{a \le z \le b} h(z)$. So for every $z \in (a,b)$ we have

$$(k\,z+l)/g(z) \le (k\,z_0+l)/g(z_0) \qquad \text{(because } h(z_0) \text{ is maximum)},$$
$$(k\,z+l)\,g(z_0) \le (k\,z+l)\,g(z) + k\,g(z)\,(z_0 - z) \qquad \text{(because } g>0 \text{ or } g<0),$$
$$(k\,z+l)\,\mu_g(z)\,(z_0 - z) \le (k\,z+l)\,(g(z_0) - g(z)) \le k\,g(z)\,(z_0 - z) \quad \text{(because } g \text{ is convex)},$$

$g'_-(z) \le \mu_g(z) \le k\,g(z)/(k\,z+l)$ for $a < z < z_0$ and $g'_+(z) \ge \mu_g(z) \ge k\,g(z)/(k\,z+l)$ for $b > z > z_0$, where $\mu_g(z)$ is a subdifferential of the function g in z, i.e. $\mu_g(z) \in [g'_-(z), g'_+(z)]$. So

$$h'_-(z) \ge 0 \quad \text{for } a < z < z_0 \qquad \text{and} \qquad h'_+(z) \le 0 \quad \text{for } b > z > z_0. \tag{9.122}$$

It follows that for each number z_0 at which the function h has the global maximum on $[a,b]$ the conditione $g'_-(z_0) \le kg(z_0)/(kz_0+l) \le g'_+(z_0)$ is valid.

In the case when h is monotonically decreasing on $[a,b]$, we have $\max_{a \le z \le b} h(z) = h(a)$ and $h'_-(z) \le 0$ for all $z \in (a,b)$, which imply that $g'_-(z) \ge kg(z)/(kz+l)$ for every $z \in (a,b)$. In the same way we can observe the case when h is monotonically increasing.

With additional assumptions it follows by using (9.122) that the function h is strictly increasing on $(a, z_0]$ and strictly decreasing on $[z_0, b)$. Hence the equality (9.121) is valid.

Minimum value: There does not exist $z_0 \in (a,b)$ at which the function h has the global minimum. Indeed, if h is not a monotone function on $[a,b]$, it follows by using (9.122) that h is increasing on $(a, \bar{z}_0]$ and decreasing on $[\bar{z}_0, b)$, where $\bar{z}_0 \in (a,b)$ is the point at which the function h has the global maximum. It follows that the function h does not have a global minimum on (a,b), and consequently (9.119) is valid. □

Similarly to Proposition 9.3 we obtain the following result.

Proposition 9.4 *Let $g : [a,b] \to \mathbb{R}$ be either a strictly positive or strictly negative continuous function and let $h(z) = (kz+l)/g(z)$ be a ratio function with a strictly negative numerator. If g is convex, then the equality* (9.119) *is valid with* max *instead of* min*, and the equality* (9.120) *is valid with* min *instead of* max .

Additionally, if g is strictly convex and h is not monotone, then the equality (9.121) *is valid with* min *instead of* max.

Proposition 9.5 *Let $g : [a,b] \to \mathbb{R}$ be either a strictly positive or strictly negative continuous function and let $h(z) = (kz+l)/g(z)$ be a ratio function with a strictly positive numerator. If g is concave, then*

$$\max_{a \le z \le b} h(z) = \max\{h(a), h(b)\}. \tag{9.123}$$

and

$$\min_{a \le z \le b} h(z) = \begin{cases} h(a) & \text{if } g'_-(z) \le \dfrac{kg(z)}{kz+l} \text{ for every } z \in (a,b), \\ h(z_0) & \text{if } g'_+(z_0) \le \dfrac{kg(z_0)}{kz_0+l} \le g'_-(z_0) \text{ for some } z_0 \in (a,b), \\ h(b) & \text{if } g'_+(z) \ge \dfrac{kg(z)}{kz+l} \text{ for every } z \in (a,b). \end{cases} \tag{9.124}$$

Additionally, if g is strictly concave and h is not monotone, then a unique number $z_0 \in (a,b)$ exists so that

$$h(z_0) = \min_{a \le z \le b} h(z). \tag{9.125}$$

Proof. The proof is the same as the one in Proposition 9.3. □

Similarly to the above proposition we obtain the following result.

Proposition 9.6 *Let $g : [a,b] \to \mathbb{R}$ be either a strictly positive or strictly negative continuous function and let $h(z) = (kz+l)/g(z)$ be a ratio function with a strictly negative numerator. If g is concave, then the equality* (9.123) *is valid with* min *instead of* max*, and the equality* (9.124) *is valid with* max *instead of* min.

Additionally, if g is strictly concave and h is not monotone, then the equality (9.125) *is valid with* max *instead of* min .

9.8 Notes

A version of Jensen's operator inequality and its converses for a bounded continuous field self-adjoint elements in C^*-algebra and a unital field of positive linear mappings is firstly discussed by Hansen, Pečarić and I. Perić [135] based on [137] by Hansen and Pederson. A generalization of the previous results on non-unital fields of positive linear mappings is presented by Mićić, Pečarić and Seo [202]. The results in Sections 9.3 and 9.4 for power operator means are given by Mićić, and Pečarić [193]. A version of these results is presented in Sections 9.5 for quasi-arithmetic means, which is based on the results of Mićić, Pečarić and Seo [203, 204]. Results with some better bounds in Section 9.6 are given by Mićić, Pavić and Pečarić [189].

Chapter 10

Jensen's Operator Inequality Without Operator Convexity

In this chapter, we study Jensen's operator inequality without operator convexity. We observe this inequality for an $n-$tuples of self-adjoint operators, a unital $n-$tuples of positive linear mappings and a general convex function with conditions on the operators bounds. In the present context, we also study an extension and a refinement of Jensen's operator inequality. As an application we give the order among quasi-arithmetic operator means.

10.1 Jensen's operator inequality with a general convex function

In this section we give our main result about Jensen's operator inequality without operator convexity. We give this inequality with conditions on the bounds of the operators (defined by (1.2)), but for a general convex function. We also study monotonicity of quasi-arithmetic operator means under the same conditions.

Suppose that J is an arbitrary interval in $\mathbb{R}$.

We recall that operator convexity plays an essential role in the Davis-Choi-Jensen inequality: $f(\Phi(A)) \leq \Phi(f(A))$. In fact, this inequality will be false if we replace an operator convex function by a general convex function (see the example given by M.D. Choi in [35]). Furthermore, if $f : J \to \mathbb{R}$ be an operator convex function, then the generalized discrete Jensen's operator inequality (8.12):

$$f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right) \le \sum_{i=1}^{n}\Phi_i(f(A_i)) \tag{$\clubsuit$}$$

holds for every $n-$tuple $(A_1,\ldots,A_n)$ of self-adjoint operators in $B(H)$ with spectra in J and every unital $n-$tuple $(\Phi_1,\ldots,\Phi_n)$ of positive linear mappings $\Phi_i : B(H) \to B(K)$, $i = 1,\ldots,n$, (i.e. $(\Phi_1,\ldots,\Phi_n)$ is such that $\sum_{i=1}^{n}\Phi_i(I_H) = I_K$).

Next we observe an example when the above Jensen's inequality is valid for some non-operator convex function.

Example 10.1 *It appears that the above inequality will be false if we replace the operator convex function by a general convex function. For example, we define mappings* $\Phi_1,\Phi_2 : M_3(\mathbb{C}) \to M_2(\mathbb{C})$ *by* $\Phi_1((a_{ij})_{1\le i,j\le 3}) = \frac{1}{2}(a_{ij})_{1\le i,j\le 2}$, $\Phi_2 = \Phi_1$. *Then* $\Phi_1(I_3)+\Phi_2(I_3) = I_2$.

I) *If*

$$A_1 = 2\begin{pmatrix}1&0&1\\0&0&1\\1&1&1\end{pmatrix} \quad \text{and} \quad A_2 = 2\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix},$$

then

$$(\Phi_1(A_1)+\Phi_2(A_2))^4 = \begin{pmatrix}16&0\\0&0\end{pmatrix} \not\le \begin{pmatrix}80&40\\40&24\end{pmatrix} = \Phi_1\left(A_1^4\right)+\Phi_2\left(A_2^4\right).$$

Given the above, there is no relation between $(\Phi_1(A_1)+\Phi_2(A_2))^4$ *and* $\Phi_1\left(A_1^4\right) + \Phi_2\left(A_2^4\right)$ *under the operator order. We observe that in the above case the following stands* $A = \Phi_1(A_1)+\Phi_2(A_2) = \begin{pmatrix}2&0\\0&0\end{pmatrix}$ *and* $[m_A,M_A] = [0,2]$, $[m_1,M_1] \subset [-1.60388, 4.49396]$, $[m_2,M_2] = [0,2]$, *i.e.*

$$(m_A,M_A) \subset [m_1,M_1] \cup [m_2,M_2]$$

similarly as in Figure 10.1.a).

II) *If*

$$A_1 = \begin{pmatrix}-14&0&1\\0&-2&-1\\1&-1&-1\end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix}15&0&0\\0&2&0\\0&0&15\end{pmatrix},$$

then

$$(\Phi_1(A_1)+\Phi_2(A_2))^4 = \begin{pmatrix}\frac{1}{16}&0\\0&0\end{pmatrix} \le \begin{pmatrix}89660&-247\\-247&51\end{pmatrix} = \Phi_1\left(A_1^4\right)+\Phi_2\left(A_2^4\right).$$

So we have that an inequality of type ($\clubsuit$) *now is valid. In the above case the following stands* $A = \Phi_1(A_1)+\Phi_2(A_2) = \begin{pmatrix}\frac{1}{2}&0\\0&0\end{pmatrix}$ *and* $[m_A,M_A] = [0,0.5]$, $[m_1,M_1] \subset [-14.077,-0.328566]$, $[m_2,M_2] = [2,15]$, *i.e.*

$$(m_A,M_A)\cap[m_1,M_1] = \emptyset \quad \text{and} \quad (m_A,M_A)\cap[m_2,M_2] = \emptyset.$$

similarly as in Figure 10.1.b).

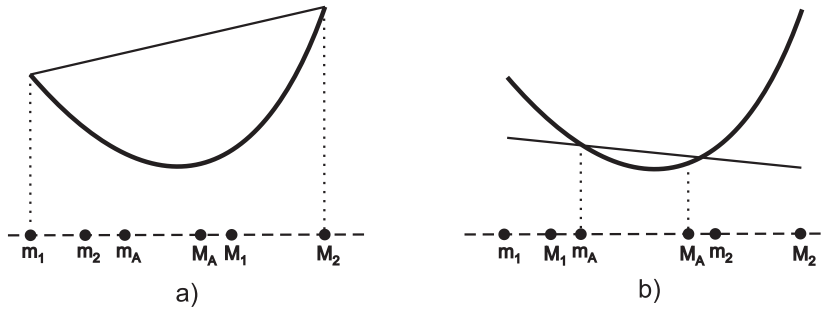


Figure 10.1: Spectral conditions for a convex function f

It is no coincidence that the inequality (♣) is valid in Example 10.1-II). In the following theorem we prove a general result when Jensen's operator inequality holds for convex functions.

Theorem 10.1 *Let $(A_1,\dots,A_n)$ be an $n-$tuple of self-adjoint operators $A_i \in B(H)$ with the bounds m_i and M_i, $m_i \le M_i$, $i=1,\dots,n$. Let $(\Phi_1,\dots,\Phi_n)$ be an $n-$tuple of positive linear mappings $\Phi_i : B(H) \to B(K)$, $i=1,\dots,n$, such that $\sum_{i=1}^n \Phi_i(I_H) = I_K$. If*

$$(m_A, M_A) \cap [m_i, M_i] = \emptyset \quad \text{for } i=1,\dots,n, \tag{10.1}$$

where m_A and M_A, $m_A \le M_A$, are the bounds of the self-adjoint operator $A = \sum_{i=1}^n \Phi_i(A_i)$, then

$$f\left(\sum_{i=1}^n \Phi_i(A_i)\right) \le \sum_{i=1}^n \Phi_i\left(f(A_i)\right) \tag{10.2}$$

holds for every continuous convex function $f : J \to \mathbb{R}$ provided that the interval J contains all m_i, M_i.

If $f : J \to \mathbb{R}$ is concave, then the reverse inequality is valid in (10.2)*.*

Proof. We prove only the case when f is a convex function.

If we denote $m = \min\{m_1,\dots,m_n\}$ and $M = \max\{M_1,\dots,M_n\}$, then $[m,M] \subseteq I$ and $mI_H \le A_i \le MI_H$, $i=1,\dots,n$. It follows $mI_K \le \sum_{i=1}^n \Phi_i(A_i) \le MI_K$. Therefore $[m_A, M_A] \subseteq [m,M] \subseteq I$.

a) Let $m_A < M_A$. Since f is convex on $[m_A, M_A]$, then

$$f(t) \le \frac{M_A - t}{M_A - m_A} f(m_A) + \frac{t - m_A}{M_A - m_A} f(M_A), \quad t \in [m_A, M_A], \tag{10.3}$$

but since f is convex on $[m_i, M_i]$ and since $(m_A, M_A) \cap [m_i, M_i] = \emptyset$, then

$$f(t) \ge \frac{M_A - t}{M_A - m_A} f(m_A) + \frac{t - m_A}{M_A - m_A} f(M_A), \quad t \in [m_i, M_i] \text{ for } i=1,\dots,n. \tag{10.4}$$

Since $m_A I_K \le \sum_{i=1}^n \Phi_i(A_i) \le M_A I_K$, then by using functional calculus, it follows from (10.3)

$$f\left(\sum_{i=1}^n \Phi_i(A_i)\right) \le \frac{M_A I_K - \sum_{i=1}^n \Phi_i(A_i)}{M_A - m_A} f(m_A) + \frac{\sum_{i=1}^n \Phi_i(A_i) - m_A I_K}{M_A - m_A} f(M_A). \qquad (10.5)$$

On the other hand, since $m_i I_H \le A_i \le M_i I_H$, $i = 1,\ldots,n$, then by using functional calculus, it follows from (10.4)

$$f(A_i) \ge \frac{M_A I_H - A_i}{M_A - m_A} f(m_A) + \frac{A_i - m_A I_H}{M_A - m_A} f(M_A), \qquad i = 1,\ldots,n.$$

Applying a positive linear mapping Φ_i and summing, we obtain

$$\sum_{i=1}^n \Phi_i(f(A_i)) \ge \frac{M_A I_K - \sum_{i=1}^n \Phi_i(A_i)}{M_A - m_A} f(m_A) + \frac{\sum_{i=1}^n \Phi_i(A_i) - m_A I_K}{M_A - m_A} f(M_A), \qquad (10.6)$$

since $\sum_{i=1}^n \Phi_i(I_H) = I_K$. Combining the two inequalities (10.5) and (10.6), we have the desired inequality (10.2).

b) Let $m_A = M_A$. Since f is convex on $[m,M]$, we have

$$f(t) \ge f(m_A) + l(m_A)(t - m_A) \quad \text{for every } t \in [m,M], \qquad (10.7)$$

where l is the subdifferential of f. Since $mI_H \le A_i \le MI_H$, $i = 1,\ldots,n$, then by using functional calculus, applying a positive linear mapping Φ_i and summing, we obtain from (10.7)

$$\sum_{i=1}^n \Phi_i(f(A_i)) \ge f(m_A)I_K + l(m_A)\left(\sum_{i=1}^n \Phi_i(A_i) - m_A I_K\right).$$

Since $m_A I_K = \sum_{i=1}^n \Phi_i(A_i)$, it follows

$$\sum_{i=1}^n \Phi_i(f(A_i)) \ge f(m_A)I_K = f\left(\sum_{i=1}^n \Phi_i(A_i)\right),$$

which is the desired inequality (10.2). □

We have the following obvious corollary of Theorem 10.1 with the convex combination of operators A_i, $i = 1,\ldots,n$.

Corollary 10.1 *Let $(A_1,\ldots,A_n)$ be an $n-$tuple of self-adjoint operators $A_i \in B(H)$ with the bounds m_i and M_i, $m_i \le M_i$, $i = 1,\ldots,n$. Let $(\alpha_1,\ldots,\alpha_n)$ be an $n-$tuple of nonnegative real numbers such that $\sum_{i=1}^n \alpha_i = 1$. If*

$$(m_A, M_A) \cap [m_i, M_i] = \emptyset \quad \textit{for } i = 1,\ldots,n,$$

where m_A and M_A, $m_A \le M_A$, are the bounds of $A = \sum_{i=1}^n \alpha_i A_i$, then

$$f\left(\sum_{i=1}^n \alpha_i A_i\right) \le \sum_{i=1}^n \alpha_i f(A_i) \qquad (10.8)$$

holds for every continuous convex function $f : J \to \mathbb{R}$ provided that the interval J contains all m_i, M_i.

Proof. We apply Theorem 10.1 for positive linear mappings $\Phi_i : B(H) \to B(H)$ determined by $\Phi_i : B \mapsto \alpha_i B$, $i = 1, \ldots, n$. □

In the present context we can study the monotonicity of the discrete version of quasi-arithmetic mean (9.65) defined as follows

$$M_\varphi(\mathbf{A}, \mathbf{\Phi}, n) = \varphi^{-1}\left(\sum_{i=1}^{n} \Phi_i\left(\varphi(A_i)\right)\right), \tag{10.9}$$

where $(A_1, \ldots, A_n)$ is an $n-$tuple of self-adjoint operators in $B(H)$ with spectra in J, $(\Phi_1, \ldots, \Phi_n)$ is a unital $n-$tuple of positive linear mappings $\Phi_i : B(H) \to B(K)$ and $\varphi : J \to \mathbb{R}$ is a continuous strictly monotone function.

Example 10.2 *Theorem 9.17 will not true if we replace the operator convex function by a general convex function in* (9.66)*. Indeed, we put for $T = \{1, 2\}$, $\varphi(t) = \sqrt[3]{t}$ and $\psi = \iota$ (the identity function) in* (10.9) *($\psi \circ \varphi^{-1}(t) = t^3$ is not operator convex) and we define mappings $\Phi_1, \Phi_2 : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ by $\Phi_1(B) = \Phi_2(B) = \frac{1}{2}B$ for $B \in M_2(\mathbb{C})$ (then $\Phi_1(I_2) + \Phi_2(I_2) = I_2$). If*

$$A_1 = \begin{pmatrix} 34 & 14 \\ 14 & 6 \end{pmatrix} \quad \textit{and} \quad A_2 = \begin{pmatrix} 36 & 28 \\ 28 & 36 \end{pmatrix},$$

then

$$\begin{aligned} M_{\sqrt[3]{t}}(A, \Phi, 2) &= \left(\Phi_1\left(\sqrt[3]{A_1}\right) + \Phi_2\left(\sqrt[3]{A_2}\right)\right)^3 \\ &= \left(\frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\right)^3 = \begin{pmatrix} 35 & 20 \\ 20 & 15 \end{pmatrix}, \\ M_1(A, \Phi, 2) &= \Phi_1(A_1) + \Phi_2(A_2) = \begin{pmatrix} 35 & 21 \\ 21 & 35 \end{pmatrix}, \end{aligned}$$

$$M_1(A, \Phi, 2) - M_{\sqrt[3]{}}(A, \Phi, 2) = \begin{pmatrix} 0 & 1 \\ 1 & 20 \end{pmatrix} \not\geq 0.$$

Given the above, there is no relation between $M_1(A, \Phi, 2)$ and $M_{\sqrt[3]{}}(A, \Phi, 2)$ under the operator order. For the bounds of A_1, A_2 and the mean $M_{\sqrt[3]{}}(A, \Phi, 2)$ the following stands $[m_1, M_1] \subset [0.2, 39.8]$, $[m_2, M_2] = [8, 64]$ and $[m, M] \subset [2.63, 47.37]$, respectively. We observe that in the above case the following stands

$$(m, M) \cap [m_1, M_1] \neq \emptyset, \quad (\textit{and} \quad (m, M) \cap [m_2, M_2] \neq \emptyset.)$$

In the case when $(m, M) \cap [m_1, M_1] = \emptyset$ and $(m, M) \cap [m_2, M_2] = \emptyset$ for some A_1 and A_2, then the relation $M_{\sqrt[3]{}}(A, \Phi, 2) \leq M_1(A, \Phi, 2)$ holds according to Theorem 10.2.

In the next theorem we will examine the order among quasi-arithmetic means without operator convexity in Theorem 9.17 when $T = \{1,\dots,n\}$ and $k = 1$.

Theorem 10.2 *Let* $(A_1,\dots,A_n)$ *and* $(\Phi_1,\dots,\Phi_n)$ *be as in the definition of the quasi-arithmetic mean* (10.9)*. Let* m_i *and* M_i, $m_i \le M_i$ *be the bounds of* A_i, $i = 1,\dots,n$. *Let* $\varphi,\psi : J \to \mathbb{R}$ *be continuous strictly monotone functions on an interval* J *which contains all* m_i, M_i. *Let* m_φ *and* M_φ, $m_\varphi \le M_\varphi$, *be the bounds of the mean* $M_\varphi(\mathbf{A},\mathbf{\Phi},n)$*, such that*

$$\big(m_\varphi, M_\varphi\big) \cap [m_i, M_i] = \emptyset \qquad \text{for } i = 1,\dots,n. \tag{10.10}$$

If one of the following conditions

(i) $\psi\circ\varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(i') $\psi\circ\varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$M_\varphi(\mathbf{A},\mathbf{\Phi},n) \le M_\psi(\mathbf{A},\mathbf{\Phi},n). \tag{10.11}$$

If one of the following conditions

(ii) $\psi\circ\varphi^{-1}$ *is concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi\circ\varphi^{-1}$ *is convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the reverse inequality is valid in (10.11)*.*

Proof. We prove the case (i) only. Suppose that φ is a strictly increasing function. Since $m_i I_H \le A_i \le M_i I_H$, $i = 1,\dots,n$, and $m_\varphi I_K \le M_\varphi(\mathbf{A},\mathbf{\Phi},n) \le M_\varphi I_K$, then

$$\varphi(m_i) I_H \le \varphi(A_i) \le \varphi(M_i) I_H, \quad i = 1,\dots,n,$$
$$\text{and} \qquad \varphi(m_\varphi) I_K \le \sum_{i=1}^{n} \Phi_i(\varphi(A_i)) \le \varphi(M_\varphi) I_K.$$

Then,

$$\big(m_\varphi, M_\varphi\big) \cap [m_i, M_i] = \emptyset \quad \text{ for } i = 1,\dots,n$$

implies

$$\big(\varphi(m_\varphi), \varphi(M_\varphi)\big) \cap [\varphi(m_i), \varphi(M_i)] = \emptyset \quad \text{ for } i = 1,\dots,n. \tag{10.12}$$

Replacing A_i by $\varphi(A_i)$ in (10.2) and taking into account (10.12), we obtain

$$f\left(\sum_{i=1}^{n} \Phi_i(\varphi(A_i))\right) \le \sum_{i=1}^{n} \Phi_i\left(f(\varphi(A_i))\right) \tag{10.13}$$

for every convex function $f : J \to \mathbb{R}$ on an interval J which contains all $[\varphi(m_i), \varphi(M_i)] = \varphi([m_i, M_i])$. Also, if φ is strictly decreasing, then we check that (10.13) holds for convex $f : J \to \mathbb{R}$ on J which contains all $[\varphi(M_i), \varphi(m_i)] = \varphi([m_i, M_i])$.

Putting $f = \psi \circ \varphi^{-1}$ in (10.13), we obtain

$$\psi \circ \varphi^{-1}\left(\sum_{i=1}^{n} \Phi_i\left(\varphi(A_i)\right)\right) \leq \sum_{i=1}^{n} \Phi_i\left(\psi(A_i)\right).$$

Applying an operator monotone function ψ^{-1} on the above inequality, we get the desired inequality (10.11). □

We can give the following version of Corollary 9.8 without operator convexity and operator concavity.

Corollary 10.2 *Let the assumptions of Theorem 10.2 hold. Let $F : J \times J \to \mathbb{R}$ be a bounded and operator monotone function in its first variable, such that $F(t,t) = C$ for all $t \in [m_\varphi, M_\varphi]$.*
If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$F\left[M_\psi(\mathbf{A}, \mathbf{\Phi}, n), M_\varphi(\mathbf{A}, \mathbf{\Phi}, n)\right] \geq C I_K. \tag{10.14}$$

(ii) $\psi \circ \varphi^{-1}$ *is concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the reverse inequality is valid in (10.14).

Proof. The proof is the same as the one of Corollary 9.8 and we omit it. □

Now, we will examine the order among quasi-arithmetic means (10.9) without operator convexity and operator concavity in Theorem 9.18.

Corollary 10.3 *Let $(A_1, \ldots, A_n)$ and $(\Phi_1, \ldots, \Phi_n)$ be as in the definition of the quasi-arithmetic mean* (10.9)*. Let m_i and M_i, $m_i \leq M_i$ be the bounds of A_i, $i = 1, \ldots, n$. Let $\varphi, \psi : J \to \mathbb{R}$ be continuous strictly monotone functions on an interval J which contains all m_i, M_i and M_1 be generated by the identity function on J.*

(i) *If m_φ and M_φ, $m_\varphi \leq M_\varphi$ are the bounds of $M_\varphi(\mathbf{A}, \mathbf{\Phi}, n)$, such that*

$$\left(m_\varphi, M_\varphi\right) \cap [m_i, M_i] = \emptyset \qquad \textit{for } i = 1, \ldots, n \tag{10.15}$$

and φ^{-1} is convex, then

$$M_\varphi(\mathbf{A}, \mathbf{\Phi}, n) \leq M_1(\mathbf{A}, \mathbf{\Phi}, n). \tag{10.16}$$

(ii) *If* (10.15) *is satisfied and φ^{-1} is concave, then the reverse inequality is valid in* (10.16).

(iii) *If* (10.15) *is satisfied and* φ^{-1} *is convex, and if* m_ψ *and* M_ψ, $m_\psi \le M_\psi$ *are the bounds of* $M_\psi(\mathbf{A},\mathbf{\Phi},n)$, *such that*

$$\left(m_\psi, M_\psi\right) \cap [m_i, M_i] = \emptyset \qquad \textit{for } i = 1,\ldots,n$$

and ψ^{-1} *is concave, then*

$$M_\varphi(\mathbf{A},\mathbf{\Phi},n) \le M_1(\mathbf{A},\mathbf{\Phi},n) \le M_\psi(\mathbf{A},\mathbf{\Phi},n). \tag{10.17}$$

Proof. (i) – (ii): Putting $\psi = M_1$ in Theorem 10.2(i) and (ii), we obtain (10.16) and its reverse inequality, respectively.

(iii): Replacing ψ by φ in (ii) and combining this with (i), we obtain the desired inequality (10.17). □

Remark 10.1 Results given in the previous section we can generalize for continuous fields of operators. E.g. the continuous version of Theorem 10.1 is given below (see also Theorem 10.7 in Section 10.5).

Let $(x_t)_{t\in T}$ *be a bounded continuous field of self-adjoint elements in an unital* C^*-*algebra* $\mathscr{A}$ *defined on a locally compact Hausdorff space* T *equipped with a bounded Radon measure* μ. *Let* m_t *and* M_t, $m_t \le M_t$, *be the bounds of* x_t, $t \in T$. *Let* $(\phi_t)_{t\in T}$ *be an unital field of positive linear mappings* $\phi_t : \mathscr{A} \to \mathscr{B}$ *from* $\mathscr{A}$ *to another unital* C^*-*algebra* $\mathscr{B}$. *Let*

$$(m_x, M_x) \cap [m_t, M_t] = \emptyset, \qquad t \in T,$$

where m_x *and* M_x, $m_x \le M_x$, *are the bounds of the operator* $x = \int_T \phi_t(x_t)\, d\mu(t)$. *If* $f : J \to \mathbb{R}$ *is a continuous convex function provided that the interval* J *contains all* m_t, M_t, *then*

$$f\left(\int_T \phi_t(x_t)\, d\mu(t)\right) \le \int_T \phi_t(f(x_t))\, d\mu(t). \tag{10.18}$$

If f *is concave, the reverse inequality is valid in* (10.18).

10.2 Order among power means

The operator power mean $M_r(A,\mathbf{\Phi})$ defined by (8.33) is a special case of the quasi-arithmetic mean (10.9). As a continuation of our previous considerations about the order among quasi-arithmetic operator means, in this section we observe the order among operator power means.

We recall the known result as follows (see Example 9.1).

Corollary 10.4 *Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an $n-$tuple of positive invertible operators in $B(H)$ with $Sp(A_i) \subseteq [m, M]$ for some scalars $0 < m < M$, and let $\mathbf{\Phi} = (\Phi_1, \ldots, \Phi_n)$ be a unital $n-$tuple positive linear mappings $\Phi_i : B(H) \to B(K)$, $i = 1, \ldots, n$.*

If either $r \leq s$, $r \notin (-1,1)$, $s \notin (-1,1)$ or $1/2 \leq r \leq 1 \leq s$ or $r \leq -1 \leq s \leq -1/2$ (see Figure 10.1.a), then

$$M_r(\mathbf{A}, \mathbf{\Phi}) \leq M_s(\mathbf{A}, \mathbf{\Phi}). \tag{10.19}$$

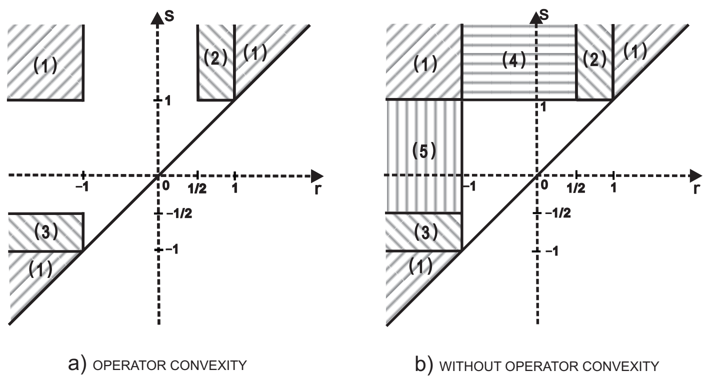


Figure 10.2: Regions for the order among power means

Remark 10.2 *Corollary 10.4 is not valid if r, s are not in the regions (1)-(2) in Figure 10.2.a) (see Example 10.2).*

Applying Theorem 10.2 we obtain that (10.19) holds in a broader region (see Figure 10.2.b).

Corollary 10.5 *Let $(A_1, \ldots, A_n)$ and $(\Phi_1, \ldots, \Phi_n)$ be as in Corollary 10.4. Let m_i and M_i, $0 < m_i \leq M_i$ be the bounds of A_i, $i = 1, \ldots, n$.*
If one of the following conditions

(i) *$r \leq s$, $s \geq 1$ or $r \leq s \leq -1$ (Figure 10.2.b (1),(2),(4)) and*

$$\left(m^{[r]}, M^{[r]}\right) \cap [m_i, M_i] = \emptyset, \qquad i = 1, \ldots, n,$$

where $m^{[r]}$ and $M^{[r]}$, $m^{[r]} \leq M^{[r]}$ are the bounds of $M_r(\mathbf{A}, \mathbf{\Phi})$,

(ii) *$r \leq s$, $r \leq -1$ or $1 \leq r \leq s$ (Figure 10.2.b (1),(3),(5)) and*

$$\left(m^{[s]}, M^{[s]}\right) \cap [m_i, M_i] = \emptyset, \qquad i = 1, \ldots, n,$$

where $m^{[s]}$ and $M^{[s]}$, $m^{[s]} \leq M^{[s]}$ are the bounds of $M_s(\mathbf{A}, \mathbf{\Phi})$, *is satisfied, then*

$$M_r(\mathbf{A},\mathbf{\Phi}) \le M_s(\mathbf{A},\mathbf{\Phi}). \tag{10.20}$$

Proof. We prove the case (i) only. We put $\varphi(t) = t^r$ and $\psi(t) = t^s$ for $t > 0$.

Then $\psi \circ \varphi^{-1}(t) = t^{s/r}$ is concave for $r \le s$, $s \le 0$ and $r \ne 0$. Since $-\psi^{-1}(t) = -t^{1/s}$ is operator monotone for $s \le -1$ and $\left(m^{[r]}, M^{[r]}\right) \cap [m_i, M_i] = \emptyset$ is satisfied, then by applying Theorem 10.2-(i') we obtain (10.20) for $r \le s \le -1$.

But, $\psi \circ \varphi^{-1}(t) = t^{s/r}$ is convex for $r \le s$, $s \ge 0$ and $r \ne 0$. Since $\psi^{-1}(t) = t^{1/s}$ is operator monotone for $s \ge 1$, then by applying Theorem 10.2-(i) we obtain (10.20) for $r \le s$, $s \ge 1$, $r \ne 0$.

If $r = 0$ and $s \ge 1$, we put $\varphi(t) = \log t$ and $\psi(t) = t^s$, $t > 0$. Since $\psi \circ \varphi^{-1}(t) = \exp(st)$ is convex, then similarly as above we obtain the desired inequality.

In the case (ii) we put $\varphi(t) = t^s$ and $\psi(t) = t^r$ for $t > 0$ and we use the same technique as in the case (i). □

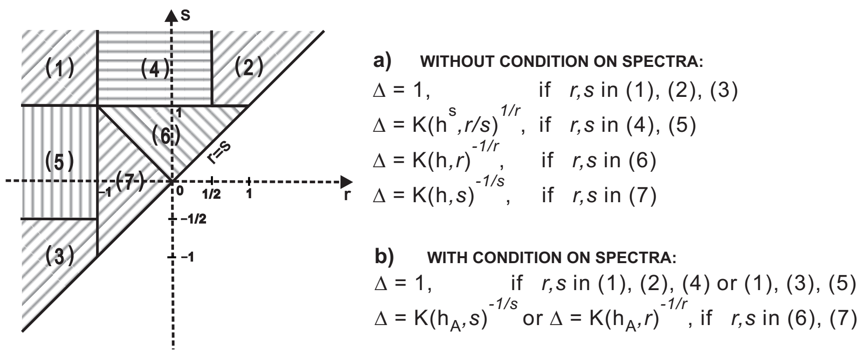


Figure 10.3: Regions for the order $M_r(\mathbf{A},\mathbf{\Phi}) \le \Delta M_s(\mathbf{A},\mathbf{\Phi})$

Figure 10.3 shows regions (1),(2),(3) in which the monotonicity of the power mean (♣) holds true and regions (1)-(5) which this holds true with the condition on spectra. In the next theorem we observe the order among power operator means with the condition on spectra in regions (6) and (7) in Figure 10.3.

Theorem 10.3 *Let $(A_1,\ldots,A_n)$ and $(\Phi_1,\ldots,\Phi_n)$ be as in Corollary 10.4. Let m_i and M_i, $0 < m_i \le M_i$ be bounds of A_i, $i = 1,\ldots,n$. Let $r,s \in (-1,1)$, $r \le s$ (Figure 10.3 (6),(7)).*

(i) *If*

$$\left(m^{[r]}, M^{[r]}\right) \cap [m_i, M_i] = \emptyset, \qquad i = 1,\ldots,n,$$

where $m^{[r]}$ and $M^{[r]}$, $m^{[r]} \le M^{[r]}$ are bounds of $M_r(\mathbf{A},\mathbf{\Phi})$, then

$$M_r(\mathbf{A},\mathbf{\Phi}) \le C(h^{[r]}, s)\, M_s(\mathbf{A},\mathbf{\Phi}), \qquad h^{[r]} = M^{[r]}/m^{[r]}. \tag{10.21}$$

(ii) *If*

$$\left(m^{[s]}, M^{[s]}\right) \cap [m_i, M_i] = \emptyset, \qquad i = 1, \ldots, n,$$

where $m^{[s]}$ *and* $M^{[s]}$, $m^{[s]} \leq M^{[s]}$ *are bounds of* $M_s(\mathbf{A}, \mathbf{\Phi})$, *then*

$$M_r(\mathbf{A}, \mathbf{\Phi}) \leq C(h^{[s]}, r) M_s(\mathbf{A}, \mathbf{\Phi}), \qquad h^{[s]} = M^{[s]}/m^{[s]}. \tag{10.22}$$

The constant $C(h,p)$, $h > 0$, *is a generalization of the Specht ratio* (2.35) *defined as follows*

$$C(h,p) := \begin{cases} \frac{p(h-h^p)}{(1-p)(h^p-1)} \left(\frac{(p-1)(h-1)}{h^p-h}\right)^{\frac{1}{p}}, & \text{if } p \neq 0 \text{ and } h \neq 1, \\ \frac{(h-1)h^{\frac{1}{h-1}}}{e \log h}, & \text{if } p = 0 \text{ and } h \neq 1, \\ 1, & \text{if } h = 1. \end{cases}$$

In order to prove Theorem 10.3, we need the operator order given in the following theorem.

Theorem 10.4 *Self-adjoint operators* $A, B \in B(H)$ *with* $Sp(A) \subseteq [m_A, M_A]$ *where* $0 < m_A < M_A$ *satisfy the following implication:*

$$A \leq B \implies \mathrm{e}^A \leq S(\mathrm{e}^{M_A - m_A})\, \mathrm{e}^B$$

where $S(h)$ *is the Specht ratio defined by* (2.35)*.*

Proof. Refer to [124, Corollary 8.24] for the proof. □

Proof of Theorem 10.3. We prove the case (i) only.
a) Let $m^{[r]} < M^{[r]}$.

· Suppose that $0 < r \leq s \leq 1$. Since $m_i I_H \leq A_i \leq M_i I_H$, $i = 1, \ldots, n$, and $m^{[r]} I_K \leq M_r(\mathbf{A}, \mathbf{\Phi}) \leq M^{[r]} I_K$, then

$$m_i^r I_H \leq A_i^r \leq M_i^r I_H, \quad i = 1, \ldots, n, \tag{10.23}$$

$$(m^{[r]})^r I_K \leq \textstyle\sum_{i=1}^n \Phi_i(A_i^r) \leq (M^{[r]})^r I_K. \tag{10.24}$$

Then,

$$\left(m^{[r]}, M^{[r]}\right) \cap [m_i, M_i] = \emptyset \quad \text{for } i = 1, \ldots, n$$

implies

$$\left((m^{[r]})^r, (M^{[r]})^r\right) \cap [m_i^r, M_i^r] = \emptyset \quad \text{for } i = 1, \ldots, n. \tag{10.25}$$

Putting $f(t) = t^{s/r}$, which is convex, in Theorem 10.1 and replacing A_i by A_i^r, we obtain

$$\left(\sum_{i=1}^n \Phi_i(A_i^r)\right)^{s/r} \leq \sum_{i=1}^n \Phi_i(A_i^s). \tag{10.26}$$

Now, applying Theorem 4.3 for $p = \frac{1}{s} \geq 1$ and using that

$$(m^{[r]})^s I_K \leq \left(\sum_{i=1}^{n} \Phi_i(A_i^r) \right)^{s/r} \leq (M^{[r]})^s I_K, \tag{10.27}$$

we obtain

$$M_r(\mathbf{A},\mathbf{\Phi}) \leq K\big((m^{[r]})^s, (M^{[r]})^s, 1/s\big)\, M_s(\mathbf{A},\mathbf{\Phi}), \tag{10.28}$$

which gives the desired inequality by using $K\big((m^{[r]})^s, (M^{[r]})^s, 1/s\big) = C\big(h^{[r]}, s\big)$.

· Suppose that $-1 \leq r < 0 < s \leq 1$. Then the reverse inequality is valid in (10.23) and (10.24). It follows that

$$\big((M^{[r]})^r, (m^{[r]})^r\big) \cap [M_i^r, m_i^r] = \emptyset \quad \text{for } i = 1, \ldots, n \tag{10.29}$$

holds. Putting $f(t) = t^{s/r}$, which is convex, in Theorem 10.1 and replacing A_i by A_i^r, we again obtain (10.26). Now, applying Theorem 4.3 for $p = \frac{1}{s} \geq 1$ and since (10.27) holds, then we obtain again (10.28).

· Suppose that $-1 \leq r \leq s < 0$. Then the reverse inequality is valid in (10.23) and (10.24). It follows that (10.29) holds. Putting $f(t) = t^{s/r}$, which is concave, in Theorem 10.1 and replacing A_i by A_i^r, we obtain that the reverse inequality holds in (10.26). Now, applying Theorem 4.3 for $p = \frac{1}{s} \leq -1$ and using that reverse inequalities is valid in (10.27), then we obtain

$$M_r(\mathbf{A},\mathbf{\Phi}) \leq K\big((M^{[r]})^s, (m^{[r]})^s, 1/s\big)\, M_s(\mathbf{A},\mathbf{\Phi}).$$

Since $K\big((M^{[r]})^s, (m^{[r]})^s, 1/s\big) = K\big((m^{[r]})^s, (M^{[r]})^s, 1/s\big)$ we get again the desired inequality.

· Suppose that $0 = r < s \leq 1$. Putting the operator concave function $f(t) = \frac{1}{s}\log t$ in reverse of Jensen's operator inequality given in Theorem 9.1 and replace A_i by A_i^s, we obtain

$$\log\big(M_0(\mathbf{A},\mathbf{\Phi})\big) \leq \log\big(M_s(\mathbf{A},\mathbf{\Phi})\big).$$

The spectrum of $\log\big(M_0(\mathbf{A},\mathbf{\Phi})\big)$ is contained in $[\log m^{[0]}, \log M^{[0]}]$, and then after use Theorem 10.4 we get

$$M_0(\mathbf{A},\mathbf{\Phi}) \leq S\left(\frac{M^{[0]}}{m^{[0]}}, 1\right) M_s(\mathbf{A},\mathbf{\Phi}) = C\big(h^{[0]}, 0\big)\, M_s(\mathbf{A},\mathbf{\Phi}),$$

which is the desired inequality.

· Suppose that $-1 \leq r < s = 0$. Putting the operator convex function $f(t) = \frac{1}{r}\log t$ in Jensen's operator inequality given in Theorem 9.1 and replace A_i by A_i^r, we obtain

$$\log\big(M_r(\mathbf{A},\mathbf{\Phi})\big) \leq \log\big(M_0(\mathbf{A},\mathbf{\Phi})\big).$$

The spectrum of $\log\big(M_r(\mathbf{A},\mathbf{\Phi})\big)$ is contained in $[\log m^{[r]}, \log M^{[r]}]$. Then applying Theorem 10.4 we get

$$M_r(\mathbf{A},\mathbf{\Phi}) \le S\left(\frac{M^{[r]}}{m^{[r]}},1\right) M_0(\mathbf{A},\mathbf{\Phi}) = C\big(h^{[r]},0\big)\, M_0(\mathbf{A},\mathbf{\Phi}),$$

which is the desired inequality.

b) Let $m^{[r]} \to M^{[r]}$ in inequalities

$$M_r(\mathbf{A},\mathbf{\Phi}) \le K\big((m^{[r]})^s,(M^{[r]})^s,1/s\big)\, M_s(\mathbf{A},\mathbf{\Phi}),$$
$$\text{or } M_r(\mathbf{A},\mathbf{\Phi}) \le S\left(\frac{M^{[r]}}{m^{[r]}}\right) M_s(\mathbf{A},\mathbf{\Phi}).$$

Since $K\big(m^s,M^s,1/s\big) = K\big(m,M,s\big)^{-1/s}$ and $\lim_{m\to M} K\big(m,M,s\big) = 1$ for all $s\in\mathbb{R}$; $\lim_{h\to 1} S(h) = 1$, we obtain the desired inequalities in the case $m^{[r]} = M^{[r]}$. □

Remark 10.3 *The constant $C(h^{[r]},s)$ in RHS of (10.21) in Theorem 10.3 is not worse than the constants in RHS of the inequalities in Corollary 9.6, i.e. if $r,s\in(-1,1)$, $r\le s$, then*

$$C(h^{[r]},s) \le \min\{\Delta(h,s,1), \Delta(h,s,1)\cdot\Delta(h,r,s), \Delta(h,r,1)\},$$

where $h^{[r]} = M^{[r]}/m^{[r]}$, $h = M/m$ and $m^{[r]}$ and $M^{[r]}$, $m^{[r]} \le M^{[r]}$ are bounds of $M_r(\mathbf{A},\mathbf{\Phi})$, such that $\big(m^{[r]},M^{[r]}\big)\cap[m_i,M_i] = \emptyset$, $i=1,\ldots,n$ and

$$m = \min\{m_1,\ldots,m_n\}, \qquad M = \max\{M_1,\ldots,M_n\}.$$

Indeed, we should just use the following properties of the function $(h,s)\mapsto C(h,s) = \Delta(h,s,1)$.

(r1) *$C(h,s)$ is strictly increasing in the first variable for $h>1$ and $s<1$ by [124, Theorem 2.62 (i)],*

(r2) *$C(h,s)$ is strictly decreasing in the second variable for $h>1$ and $s\in\mathbb{R}$ by Lemma 9.3.*

So, let $r,s\in(-1,1)$, $r\le s$. Since $[m^{[r]},M^{[r]}]\subseteq[m,M]$, it follows by (r1) that

$$C(h^{[r]},s) = \Delta(h^{[r]},s,1) \le \Delta(h,s,1);$$

since $\Delta(h,r,s)\ge 1$, then

$$C(h^{[r]},s) \le \Delta(h,s,1)\cdot\Delta(h,r,s);$$

and it follows by (r2) that

$$C(h^{[r]},s) \le C(h^{[r]},r) \le \Delta(h,r,1).$$

The three inequalities above give the desired relation.

Similarly, we can observe that the constant $C(h^{[s]},r)$ in RHS of (10.22) in Theorem 10.3 is not worse than the constants in RHS of the inequalities in Corollary 9.6.

10.3 Extension of Jensen's operator inequality without operator convexity

In this section, we give an extension of Jensen's operator inequality without operator convexity. As an application of this result, we give an extension of our previous results for a version of the quasi-arithmetic mean (10.9) with an $n-$tuple of positive linear mappings which is non-unital.

In Theorem 10.1 we prove that Jensen's operator inequality holds for every continuous convex function and for every $n-$tuple of self-adjoint operators $(A_1,\ldots,A_n)$, for every $n-$tuple of positive linear mappings $(\Phi_1,\ldots,\Phi_n)$ in the case when the interval with bounds of the operator $A=\sum_{i=1}^n \Phi_i(A_i)$ has no intersection points with the interval with bounds of the operator A_i for each $i=1,\ldots,n$, i.e. when

$$(m_A,M_A)\cap[m_i,M_i]=\emptyset \quad \text{for } i=1,\ldots,n,$$

where m_A and M_A, $m_A\le M_A$, are the bounds of A, and m_i and M_i, $m_i\le M_i$, are the bounds of A_i, $i=1,\ldots,n$.

It is interesting to consider the case when $(m_A,M_A)\cap[m_i,M_i]=\emptyset$ is valid for several $i\in\{1,\ldots,n\}$, but not for all $i=1,\ldots,n$. We study it in the following theorem.

Theorem 10.5 *Let $(A_1,\ldots,A_n)$ be an $n-$tuple of self-adjoint operators $A_i\in B(H)$ with the bounds m_i and M_i, $m_i\le M_i$, $i=1,\ldots,n$. Let $(\Phi_1,\ldots,\Phi_n)$ be an $n-$tuple of positive linear mappings $\Phi_i: B(H)\to B(K)$, such that $\sum_{i=1}^n \Phi_i(I_H)=I_K$. For $1\le n_1<n$, we denote $m=\min\{m_1,\ldots,m_{n_1}\}$, $M=\max\{M_1,\ldots,M_{n_1}\}$ and $\sum_{i=1}^{n_1}\Phi_i(I_H)=\alpha I_K$, $\sum_{i=n_1+1}^{n}\Phi_i(I_H)=\beta I_K$, where $\alpha,\beta>0$, $\alpha+\beta=1$. If*

$$(m,M)\cap[m_i,M_i]=\emptyset \qquad \text{for} \quad i=n_1+1,\ldots,n,$$

and one of two equalities

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i)=\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i)=\sum_{i=1}^{n}\Phi_i(A_i)$$

is valid, then

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i))\le\sum_{i=1}^{n}\Phi_i(f(A_i))\le\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)) \tag{10.30}$$

holds for every continuous convex function $f: J\to\mathbb{R}$ provided that the interval J contains all m_i,M_i, $i=1,\ldots,n$,.

If $f: J\to\mathbb{R}$ is concave, then the reverse inequality is valid in (10.30).

Proof. We prove only the case when f is a convex function.

Let us denote

$$A=\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i), \qquad B=\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i), \qquad C=\sum_{i=1}^{n}\Phi_i(A_i).$$

It is easy to verify that $A=B$ or $B=C$ or $A=C$ implies $A=B=C$.

a) Let $m<M$. Since f is convex on $[m,M]$ and $[m_i,M_i]\subseteq[m,M]$ for $i=1,\dots,n_1$, then

$$f(t)\le \frac{M-t}{M-m}f(m)+\frac{t-m}{M-m}f(M),\quad t\in[m_i,M_i] \text{ for } i=1,\dots,n_1, \tag{10.31}$$

but since f is convex on all $[m_i,M_i]$ and $(m,M)\cap[m_i,M_i]=\emptyset$ for $i=n_1+1,\dots,n$, then

$$f(t)\ge \frac{M-t}{M-m}f(m)+\frac{t-m}{M-m}f(M),\quad t\in[m_i,M_i] \text{ for } i=n_1+1,\dots,n. \tag{10.32}$$

Since $m_iI_H\le A_i\le M_iI_H$, $i=1,\dots,n_1$, it follows from (10.31)

$$f(A_i)\le \frac{MI_H-A_i}{M-m}f(m)+\frac{A_i-mI_H}{M-m}f(M),\qquad i=1,\dots,n_1.$$

Applying a positive linear mapping Φ_i and summing, we obtain

$$\sum_{i=1}^{n_1}\Phi_i(f(A_i))\le \frac{M\alpha I_K-\sum_{i=1}^{n_1}\Phi_i(A_i)}{M-m}f(m)+\frac{\sum_{i=1}^{n_1}\Phi_i(A_i)-m\alpha I_K}{M-m}f(M),$$

since $\sum_{i=1}^{n_1}\Phi_i(I_H)=\alpha I_K$. It follows

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i))\le \frac{MI_K-A}{M-m}f(m)+\frac{A-mI_K}{M-m}f(M). \tag{10.33}$$

Similarly to (10.33) in the case $m_iI_H\le A_i\le M_iI_H$, $i=n_1+1,\dots,n$, it follows from (10.32)

$$\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i))\ge \frac{MI_K-B}{M-m}f(m)+\frac{B-mI_K}{M-m}f(M). \tag{10.34}$$

Combining (10.33) and (10.34) and taking into account that $A=B$, we obtain

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i))\le \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)). \tag{10.35}$$

It follows

$$\begin{aligned}
\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i)) &= \sum_{i=1}^{n_1}\Phi_i(f(A_i))+\frac{\beta}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i)) && (\text{by } \alpha+\beta=1)\\
&\le \sum_{i=1}^{n_1}\Phi_i(f(A_i))+\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)) && (\text{by (10.35)})\\
&= \sum_{i=1}^{n}\Phi_i(f(A_i))\\
&\le \frac{\alpha}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i))+\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)) && (\text{by (10.35)})\\
&= \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)) && (\text{by } \alpha+\beta=1)
\end{aligned}$$

which gives the desired double inequality (10.30).

b) Let $m = M$. Since $[m_i, M_i] \subseteq [m, M]$ for $i = 1, \ldots, n_1$, then $A_i = mI_H$ and $f(A_i) = f(m)I_H$ for $i = 1, \ldots, n_1$. It follows

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i) = mI_K \quad \text{and} \quad \frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i\left(f(A_i)\right) = f(m)I_K. \tag{10.36}$$

On the other hand, since f is convex on J, we have

$$f(t) \geq f(m) + l(m)(t-m) \quad \text{for every } t \in I, \tag{10.37}$$

where l is the subdifferential of f. Replacing t by A_i for $i = n_1+1, \ldots, n$, applying Φ_i and summing, we obtain from (10.37) and (10.36)

$$\begin{aligned}\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i\left(f(A_i)\right) &\geq f(m)I_K + l(m)\left(\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i) - mI_K\right)\\ &= f(m)I_K = \frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i\left(f(A_i)\right).\end{aligned}$$

So (10.35) holds again. The remaining part of the proof is the same as in the case a).

□

Remark 10.4 *We obtain the equivalent inequality to the one in Theorem 10.5 in the case when* $\sum_{i=1}^{n}\Phi_i(I_H) = \gamma I_K$, *for some positive scalar* γ. *If* $\alpha + \beta = \gamma$ *and one of two equalities*

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i) = \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i) = \frac{1}{\gamma}\sum_{i=1}^{n}\Phi_i(A_i)$$

is valid, then

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i)) \leq \frac{1}{\gamma}\sum_{i=1}^{n}\Phi_i(f(A_i)) \leq \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i))$$

holds for every continuous convex function f.

Remark 10.5 *Let the assumptions of Theorem 10.5 be valid.*

(1) *We observe that the following inequality*

$$f\left(\frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i)\right) \leq \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)),$$

holds for every continuous convex function $f: J \to \mathbb{R}$.

Indeed, by the assumptions of Theorem 10.5 we have

$$m\alpha I_H \leq \sum_{i=1}^{n_1} \Phi_i(f(A_i)) \leq M\alpha I_H \quad \text{and} \quad \frac{1}{\alpha}\sum_{i=1}^{n_1} \Phi_i(A_i) = \frac{1}{\beta}\sum_{i=n_1+1}^{n} \Phi_i(A_i)$$

which implies

$$mI_H \leq \sum_{i=n_1+1}^{n} \frac{1}{\beta}\Phi_i(f(A_i)) \leq MI_H.$$

Also $(m,M)\cap[m_i,M_i]=\emptyset$ *for* $i=n_1+1,\ldots,n$ *and* $\sum_{i=n_1+1}^{n}\frac{1}{\beta}\Phi_i(I_H)=I_K$ *hold. So we can apply Theorem 10.1 on operators* $A_{n_1+1},\ldots,A_n$ *and mappings* $\frac{1}{\beta}\Phi_i$ *and obtain the desired inequality.*

(2) *We denote by* m_C *and* M_C *the bounds of* $C=\sum_{i=1}^{n}\Phi_i(A_i)$. *If* $(m_C,M_C)\cap[m_i,M_i]=\emptyset$, $i=1,\ldots,n_1$ *or* f *is an operator convex function on* $[m,M]$, *then the double inequality* (10.30) *can be extended from the left side if we use Jensen's operator inequality in Theorem 9.9*

$$\begin{aligned} f\left(\sum_{i=1}^{n}\Phi_i(A_i)\right) &= f\left(\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i)\right) \\ &\leq \frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(f(A_i)) \leq \sum_{i=1}^{n}\Phi_i(f(A_i)) \leq \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(f(A_i)). \end{aligned}$$

Example 10.3 *If neither assumptions* $(m_C,M_C)\cap[m_i,M_i]=\emptyset$, $i=1,\ldots,n_1$ *nor* f *is operator convex in Remark 10.5 (2) is satisfied and if* $1<n_1<n$, *then* (10.30) *can not be extended by Jensen's operator inequality, since it is not valid. Indeed, for* $n_1=2$ *we define mappings* $\Phi_1,\Phi_2 : M_3(\mathbb{C})\to M_2(\mathbb{C})$ *by* $\Phi_1((a_{ij})_{1\leq i,j\leq 3})=\frac{\alpha}{2}(a_{ij})_{1\leq i,j\leq 2}$, $\Phi_2=\Phi_1$. *Then* $\Phi_1(I_3)+\Phi_2(I_3)=\alpha I_2$. *If*

$$A_1 = 2\begin{pmatrix}1&0&1\\0&0&1\\1&1&1\end{pmatrix} \quad \text{and} \quad A_2 = 2\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix},$$

then

$$\left(\frac{1}{\alpha}\Phi_1(A_1)+\frac{1}{\alpha}\Phi_2(A_2)\right)^4 = \frac{1}{\alpha^4}\begin{pmatrix}16&0\\0&0\end{pmatrix} \not\leq \frac{1}{\alpha}\begin{pmatrix}80&40\\40&24\end{pmatrix} = \frac{1}{\alpha}\Phi_1\left(A_1^4\right)+\frac{1}{\alpha}\Phi_2\left(A_2^4\right)$$

for every $\alpha\in(0,1)$. *We observe that* $f(t)=t^4$ *is not operator convex and* $(m_C,M_C)\cap[m_i,M_i]\neq\emptyset$, *since* $C=A=\frac{1}{\alpha}\Phi_1(A_1)+\frac{1}{\alpha}\Phi_2(A_2)=\frac{1}{\alpha}\begin{pmatrix}2&0\\0&0\end{pmatrix}$, $[m_C,M_C]=[0,2/\alpha]$, $[m_1,M_1]\subset[-1.60388,4.49396]$ *and* $[m_2,M_2]=[0,2]$.

With respect to Remark 10.4, we obtain the following obvious corollary of Theorem 10.5.

Corollary 10.6 *Let $(A_1,\ldots,A_n)$ be an $n-$tuple of self-adjoint operators $A_i \in B(H)$ with the bounds m_i and M_i, $m_i \le M_i$, $i=1,\ldots,n$. For some $1 \le n_1 < n$, we denote $m = \min\{m_1,\ldots,m_{n_1}\}$, $M = \max\{M_1,\ldots,M_{n_1}\}$. Let $(p_1,\ldots,p_n)$ be an $n-$tuple of non-negative numbers, such that $0 < \sum_{i=1}^{n_1} p_i = \mathbf{p_{n_1}} < \mathbf{p_n} = \sum_{i=1}^{n} p_i$. If*

$$(m,M)\cap[m_i,M_i]=\emptyset \qquad \textit{for} \quad i=n_1+1,\ldots,n,$$

and one of two equalities

$$\frac{1}{\mathbf{p_{n_1}}}\sum_{i=1}^{n_1} p_iA_i = \frac{1}{\mathbf{p_n}}\sum_{i=1}^{n} p_iA_i = \frac{1}{\mathbf{p_n}-\mathbf{p_{n_1}}}\sum_{i=n_1+1}^{n} p_iA_i$$

is valid, then

$$\frac{1}{\mathbf{p_{n_1}}}\sum_{i=1}^{n_1} p_if(A_i) \le \frac{1}{\mathbf{p_n}}\sum_{i=1}^{n} p_if(A_i) \le \frac{1}{\mathbf{p_n}-\mathbf{p_{n_1}}}\sum_{i=n_1+1}^{n} p_if(A_i) \tag{10.38}$$

holds for every continuous convex function $f: J \to \mathbb{R}$ provided that the interval J contains all m_i, M_i, $i=1,\ldots,n$.

If $f: J \to \mathbb{R}$ is concave, then the reverse inequality is valid in (10.38)*.*

By applying Corollary 10.6 we obtain the following special case of Theorem 10.1.

Corollary 10.7 *Let $(A_1,\ldots,A_n)$ be an $n-$tuple of self-adjoint operators $A_i \in B(H)$ with the bounds m_i and M_i, $m_i \le M_i$, $i=1,\ldots,n$. Let $(\alpha_1,\ldots,\alpha_n)$ be an $n-$tuple of nonnegative real numbers such that $\sum_{i=1}^{n}\alpha_i = 1$. If*

$$(m_A,M_A)\cap[m_i,M_i]=\emptyset \quad \textit{for } i=1,\ldots,n, \tag{10.39}$$

where m_A and M_A, $m_A \le M_A$, are the bounds of $A = \sum\limits_{i=1}^{n}\alpha_iA_i$, then

$$f\left(\sum_{i=1}^{n}\alpha_iA_i\right) \le \sum_{i=1}^{n}\alpha_i f(A_i) \tag{10.40}$$

holds for every continuous convex function $f: J \to \mathbb{R}$ provided that the interval J contains all m_i, M_i.

Proof. We prove only the convex case. We define $(n+1)-$tuple of operators $(B_1,\ldots,B_{n+1})$, $B_i \in B(H)$, by $B_1 = A = \sum_{i=1}^{n}\alpha_iA_i$ and $B_i = A_{i-1}$, $i=2,\ldots,n+1$. Then $m_{B_1} = m_A$, $M_{B_1} = M_A$ are the bounds of B_1 and $m_{B_i} = m_{i-1}$, $M_{B_i} = M_{i-1}$ are the ones of B_i, $i=2,\ldots,n+1$. Also, we define $(n+1)-$tuple of non-negative numbers $(p_1,\ldots,p_{n+1})$ by $p_1 = 1$ and $p_i = \alpha_{i-1}$, $i=2,\ldots,n+1$. Then $\sum\limits_{i=1}^{n+1} p_i = 2$ and by using (10.39) we have

$$(m_{B_1},M_{B_1})\cap[m_{B_i},M_{B_i}]=\emptyset, \ \text{for } i=2,\ldots,n+1. \tag{10.41}$$

Since

$$\sum_{i=1}^{n+1} p_iB_i = B_1 + \sum_{i=2}^{n+1} p_iB_i = \sum_{i=1}^{n} \alpha_iA_i + \sum_{i=1}^{n} \alpha_iA_i = 2B_1,$$

then

$$p_1B_1 = \frac{1}{2}\sum_{i=1}^{n+1} p_iB_i = \sum_{i=2}^{n+1} p_iB_i. \tag{10.42}$$

Taking into account (10.41) and (10.42), we can apply Corollary 10.6 for $n_1 = 1$ and B_i, p_i as above, and we get

$$p_1 f(B_1) \leq \frac{1}{2}\sum_{i=1}^{n+1} p_i f(B_i) \leq \sum_{i=2}^{n+1} p_i f(B_i),$$

which gives the desired inequality (10.40). □

10.4 Extension of order among quasi-arithmetic means

In this section we study an application of Theorem 10.5 to the quasi-arithmetic mean with weight. For a subset $\{A_{p_1},\dots,A_{p_2}\}$ of $\{A_1,\dots,A_n\}$ we denote the quasi-arithmetic mean by

$$M_\varphi(\gamma,\mathbf{A},\mathbf{\Phi},p_1,p_2) = \varphi^{-1}\left(\frac{1}{\gamma}\sum_{i=p_1}^{p_2} \Phi_i(\varphi(A_i))\right), \tag{10.43}$$

where $(A_{p_1},\dots,A_{p_2})$ are self-adjoint operators in $B(H)$ with the spectra in J, $(\Phi_{p_1},\dots,\Phi_{p_2})$ are positive linear mappings $\Phi_i : B(H) \to B(K)$ such that $\sum_{i=p_1}^{p_2} \Phi_i(I_H) = \gamma I_K$, and $\varphi : J \to \mathbb{R}$ is a continuous strictly monotone function.

The following theorem is an extension of Theorem 10.2.

Theorem 10.6 *Let $(A_1,\dots,A_n)$ be an n-tuple of self-adjoint operators in $B(H)$ with the spectra in J, $(\Phi_1,\dots,\Phi_n)$ be an n-tuple of positive linear mappings $\Phi_i : B(H) \to B(K)$ such that $\sum_{i=1}^{n} \Phi_i(I_H) = I_K$. Let m_i and M_i, $m_i \leq M_i$ be the bounds of A_i, $i = 1,\dots,n$. Let $\varphi,\psi : J \to \mathbb{R}$ be continuous strictly monotone functions on an interval J which contains all m_i, M_i. For $1 \leq n_1 < n$, we denote $m = \min\{m_1,\dots,m_{n_1}\}$, $M = \max\{M_1,\dots,M_{n_1}\}$ and $\sum_{i=1}^{n_1} \Phi_i(I_H) = \alpha I_K$, $\sum_{i=n_1+1}^{n} \Phi_i(I_H) = \beta I_K$, where $\alpha,\beta > 0$, $\alpha+\beta = 1$. Let*

$$(m,M) \cap [m_i,M_i] = \emptyset \qquad \text{for} \quad i = n_1+1,\dots,n,$$

and one of two equalities

$$M_\varphi(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) = M_\varphi(1,\mathbf{A},\mathbf{\Phi},1,n) = M_\varphi(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n) \tag{10.44}$$

be valid.

If one of the following conditions

(i) $\psi \circ \varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then

$$M_\psi(\alpha, \mathbf{A}, \mathbf{\Phi}, 1, n_1) \le M_\psi(1, \mathbf{A}, \mathbf{\Phi}, 1, n) \le M_\psi(\beta, \mathbf{A}, \mathbf{\Phi}, n_1+1, n). \tag{10.45}$$

If one of the following conditions

(ii) $\psi \circ \varphi^{-1}$ *is concave and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is convex and* $-\psi^{-1}$ *is operator monotone*

is satisfied, then the reverse inequality is valid in (10.45)*.*

Proof. We prove the case (i) only. Suppose that φ is a strictly increasing function. Since $mI_H \le A_i \le MI_H$, $i = 1, \ldots, n_1$, implies $\varphi(m)I_K \le \varphi(A_i) \le \varphi(M)I_K$, then

$$(m, M) \cap [m_i, M_i] = \emptyset \qquad \text{for} \quad i = n_1+1, \ldots, n$$

implies

$$\big(\varphi(m), \varphi(M)\big) \cap [\varphi(m_i), \varphi(M_i)] = \emptyset \quad \text{for } i = n_1+1, \ldots, n. \tag{10.46}$$

Also, by using (10.44), we have

$$\frac{1}{\alpha}\sum_{i=1}^{n_1} \Phi_i(\varphi(A_i)) = \sum_{i=1}^{n} \Phi_i(\varphi(A_i)) = \frac{1}{\beta}\sum_{i=n_1+1}^{n} \Phi_i(\varphi(A_i)).$$

Taking into account (10.46) and the above double equality, we obtain by Theorem 10.5

$$\frac{1}{\alpha}\sum_{i=1}^{n_1} \Phi_i(f(\varphi(A_i))) \le \sum_{i=1}^{n} \Phi_i(f(\varphi(A_i))) \le \frac{1}{\beta}\sum_{i=n_1+1}^{n} \Phi_i(f(\varphi(A_i))) \tag{10.47}$$

for every continuous convex function $f : J \to \mathbb{R}$ on an interval J which contains all $[\varphi(m_i), \varphi(M_i)] = \varphi([m_i, M_i])$, $i = 1, \ldots, n$.

Also, if φ is strictly decreasing, then we check that (10.47) holds for convex $f : J \to \mathbb{R}$ on J which contains all $[\varphi(M_i), \varphi(m_i)] = \varphi([m_i, M_i])$.

Putting $f = \psi \circ \varphi^{-1}$ in (10.47), we obtain

$$\frac{1}{\alpha}\sum_{i=1}^{n_1} \Phi_i(\psi(A_i)) \le \sum_{i=1}^{n} \Phi_i(\psi(A_i)) \le \frac{1}{\beta}\sum_{i=n_1+1}^{n} \Phi_i(\psi(A_i)).$$

Applying an operator monotone function ψ^{-1} on the above double inequality, we obtain the desired inequality (10.45). □

Remark 10.6 *Let the assumptions of Theorem 10.6 be valid.*

(1) *We observe that if one of the following conditions*

(i) $\psi \circ \varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(ii) $\psi \circ \varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone,*

is satisfied, then the following obvious inequality (see Remark 10.5 (1))

$$M_{\varphi}(\beta, \mathbf{A}, \mathbf{\Phi}, n_1+1, n) \le M_{\psi}(\beta, \mathbf{A}, \mathbf{\Phi}, n_1+1, n)$$

holds.

(2) *We denote by* m_{φ} *and* M_{φ} *the bounds of* $M_{\varphi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n)$. *If* $(m_{\varphi}, M_{\varphi}) \cap [m_i, M_i] = \emptyset$, $i = 1, \dots, n_1$, *and one of two following conditions*

(i) $\psi \circ \varphi^{-1}$ *is convex and* ψ^{-1} *is operator monotone,*

(ii) $\psi \circ \varphi^{-1}$ *is concave and* $-\psi^{-1}$ *is operator monotone,*

(i') $\psi \circ \varphi^{-1}$ *is operator convex and* ψ^{-1} *is operator monotone,*

(ii') $\psi \circ \varphi^{-1}$ *is operator concave and* $-\psi^{-1}$ *is operator monotone,*

is satisfied (see Theorem 9.17), then the double inequality (10.45) *can be extended from the left side as follows*

$$\begin{aligned} M_{\varphi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n) &= M_{\varphi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n_1) \\ \le M_{\psi}(\alpha, \mathbf{A}, \mathbf{\Phi}, 1, n_1) &\le M_{\psi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n) \le M_{\psi}(\beta, \mathbf{A}, \mathbf{\Phi}, n_1+1, n). \end{aligned}$$

(3) *If neither assumptions* $(m_{\psi}, M_{\psi}) \cap [m_i, M_i] = \emptyset$, $i = 1, \dots, n_1$ *nor* $\psi \circ \varphi^{-1}$ *is operator convex (or operator concave) is satisfied and if* $1 < n_1 < n$, *then* (10.45) *can not be extended from the left side by* $M_{\varphi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n_1)$ *as above. It is easy to check it with a counterexample similarly to Example 10.2.*

We now give some particular results of interest that can be derived from Theorem 10.6.

Corollary 10.8 *Let* $(A_1, \dots, A_n)$ *and* $(\Phi_1, \dots, \Phi_n)$, m_i, M_i, m, M, α *and* β *be as in Theorem 10.6. Let* J *be an interval which contains all* m_i, M_i *and*

$$(m, M) \cap [m_i, M_i] = \emptyset \qquad \text{for} \quad i = n_1+1, \dots, n.$$

If one of two equalities

$$M_{\varphi}(\alpha, \mathbf{A}, \mathbf{\Phi}, 1, n_1) = M_{\varphi}(1, \mathbf{A}, \mathbf{\Phi}, 1, n) = M_{\varphi}(\beta, \mathbf{A}, \mathbf{\Phi}, n_1+1, n)$$

is valid, then

$$\frac{1}{\alpha} \sum_{i=1}^{n_1} \Phi_i(A_i) \le \sum_{i=1}^{n} \Phi_i(A_i) \le \frac{1}{\beta} \sum_{i=n_1+1}^{n} \Phi_i(A_i) \tag{10.48}$$

holds for every continuous strictly monotone function $\varphi : J \to \mathbb{R}$ *such that* φ^{-1} *is convex on* J*. But, if* φ^{-1} *is concave, then the reverse inequality is valid in* (10.48).

On the other hand, if one of two equalities

$$\frac{1}{\alpha}\sum_{i=1}^{n_1}\Phi_i(A_i) = \sum_{i=1}^{n}\Phi_i(A_i) = \frac{1}{\beta}\sum_{i=n_1+1}^{n}\Phi_i(A_i)$$

is valid, then

$$M_\varphi(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) \le M_\varphi(1,\mathbf{A},\mathbf{\Phi},1,n) \le M_\varphi(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n) \tag{10.49}$$

holds for every continuous strictly monotone function $\varphi : J \to \mathbb{R}$ *such that one of the following conditions*

(i) φ *is convex and* φ^{-1} *is operator monotone,*

(i') φ *is concave and* $-\varphi^{-1}$ *is operator monotone*

is satisfied.

But, if one of the following conditions

(ii) φ *is concave and* φ^{-1} *is operator monotone,*

(ii') φ *is convex and* $-\varphi^{-1}$ *is operator monotone,*

is satisfied, then the reverse inequality is valid in (10.49).

Proof. The proof of (10.48) follows from Theorem 10.6 by replacing ψ with the identity function, while the proof of (10.49) follows from the same theorem by replacing φ with the identity function and ψ with φ. □

As a special case of the quasi-arithmetic mean (10.43) we can study the weighted power mean as follows. For a subset $\{A_{p_1},\dots,A_{p_2}\}$ of $\{A_1,\dots,A_n\}$ we denote this mean by

$$M_r(\gamma,\mathbf{A},\mathbf{\Phi},p_1,p_2) = \begin{cases} \left(\dfrac{1}{\gamma}\displaystyle\sum_{i=p_1}^{p_2}\Phi_i\left(A_i^r\right)\right)^{1/r}, & r\in\mathbb{R}\backslash\{0\}, \\ \exp\left(\dfrac{1}{\gamma}\displaystyle\sum_{i=p_1}^{p_2}\Phi_i\left(\log\left(A_i\right)\right)\right), & r=0, \end{cases}$$

where $(A_{p_1},\dots,A_{p_2})$ are strictly positive operators, $(\Phi_{p_1},\dots,\Phi_{p_2})$ are positive linear mappings $\Phi_i : B(H) \to B(K)$ such that $\sum_{i=p_1}^{p_2}\Phi_i(I_H) = \gamma I_K$.

We obtain the following corollary by applying Theorem 10.6 to the above mean.

Corollary 10.9 *Let* $(A_1,\dots,A_n)$ *be an* $n-$*tuple of strictly positive operators in* $B(H)$ *and* $(\Phi_1,\dots,\Phi_n)$ *be an* $n-$*tuple of positive linear mappings* $\Phi_i : B(H) \to B(K)$ *such that* $\sum_{i=1}^{n}\Phi_i(I_H) = I_K$*. Let* m_i *and* M_i*,* $0 < m_i \le M_i$ *be the bounds of* A_i*,* $i = 1,\dots,n$*. For* $1 \le n_1 < n$*, we denote* $m = \min\{m_1,\dots,m_{n_1}\}$*,* $M = \max\{M_1,\dots,M_{n_1}\}$ *and* $\sum_{i=1}^{n_1}\Phi_i(I_H) = \alpha I_K$*,* $\sum_{i=n_1+1}^{n}\Phi_i(I_H) = \beta I_K$*, where* $\alpha,\beta > 0$*,* $\alpha+\beta = 1$*.*

(i) *If either $r \le s$, $s \ge 1$ or $r \le s \le -1$ and also one of two equalities*

$$M_r(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) = M_r(1,\mathbf{A},\mathbf{\Phi},1,n) = M_r(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n) \tag{10.50}$$

is valid, then

$$M_s(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) \le M_s(1,\mathbf{A},\mathbf{\Phi},1,n) \le M_s(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n). \tag{10.51}$$

(ii) *If either $r \le s$, $r \le -1$ or $1 \le r \le s$ and also one of two equalities*

$$M_s(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) = M_s(1,\mathbf{A},\mathbf{\Phi},1,n) = M_s(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n)$$

is valid, then

$$M_r(\alpha,\mathbf{A},\mathbf{\Phi},1,n_1) \ge M_r(1,\mathbf{A},\mathbf{\Phi},1,n) \ge M_r(\beta,\mathbf{A},\mathbf{\Phi},n_1+1,n). \tag{10.52}$$

Proof. We take $\varphi(t) = t^r$ and $\psi(t) = t^s$ or $\varphi(t) = t^s$ and $\psi(t) = t^r$ for $t > 0$ and apply Theorem 10.6. We omit the details. □

10.5 Refinements

In this section we present a refinement of Jensen's inequality (10.18) and a refined the general form of its converses (9.88).

For convenience we introduce the abbreviation

$$\delta_f(m,M) := f(m) + f(M) - 2f\left(\frac{m+M}{2}\right), \tag{10.53}$$

where $f : [m,M] \to \mathbb{R}$, $m < M$, is a continue function. It is obvious that, if f is *convex* (resp. *concave*) then $\delta_f \ge 0$ (resp. $\delta_f \le 0$).

To obtain our results we need the following three lemmas.

Lemma 10.1 *Let f be a convex function on an interval J, $m,M \in J$ and $p_1,p_2 \in [0,1]$ such that $p_1 + p_2 = 1$. Then*

$$\begin{aligned}\min\{p_1,p_2\}\left[f(m) + f(M) - 2f\left(\frac{m+M}{2}\right)\right] \\ \le p_1 f(m) + p_2 f(M) - f(p_1 m + p_2 M).\end{aligned} \tag{10.54}$$

Proof. These results follows from [208, Theorem 1, p. 717] for $n = 2$ and replacing x_1 and x_2 with m and M, respectively. □

Lemma 10.2 *Let x be a bounded self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ of operators on some Hilbert space H. If the spectrum of x is in $[m,M]$, for some scalars $m < M$, then*

$$f(x) \leq \frac{M\mathbf{1}_H - x}{M-m} f(m) + \frac{x - m\mathbf{1}_H}{M-m} f(M) - \delta_f(m,M)\overline{x} \tag{10.55}$$

holds for every continuous convex function $f : [m,M] \to \mathbb{R}$, where $\delta_f(m,M)$ is defined by (10.53) *and*

$$\overline{x} = \frac{1}{2}\mathbf{1}_H - \frac{1}{M-m}\left| x - \frac{m+M}{2}\mathbf{1}_H \right|.$$

If f is concave, then the reverse inequality is valid in (10.55)*.*

Proof. We prove the convex case only. By using (10.54) we get

$$f(p_1 m + p_2 M) \leq p_1 f(m) + p_2 f(M) - \min\{p_1, p_2\}\delta_f(m,M) \tag{10.56}$$

for every $p_1, p_2 \in [0,1]$ such that $p_1 + p_2 = 1$. Let functions $p_1, p_2 : [m,M] \to [0,1]$ be defined by

$$p_1(z) = \frac{M-z}{M-m}, \qquad p_2(z) = \frac{z-m}{M-m}.$$

Than for any $z \in [m,M]$ we can write

$$f(z) = f\left(\frac{M-z}{M-m}m + \frac{z-m}{M-m}M\right) = f\left(p_1(z)m + p_2(z)M\right).$$

By using (10.56) we get

$$f(z) \leq \frac{M-z}{M-m} f(m) + \frac{z-m}{M-m} f(M) - \tilde{z}\delta_f(m,M), \tag{10.57}$$

where

$$\tilde{z} = \frac{1}{2} - \frac{1}{M-m}\left| z - \frac{m+M}{2} \right|,$$

since

$$\min\left\{\frac{M-z}{M-m}, \frac{z-m}{M-m}\right\} = \frac{1}{2} - \frac{1}{M-m}\left| z - \frac{m+M}{2} \right|.$$

Finally by utilizing the functional calculus to (10.57) we obtain the desired inequality (10.55). □

In the following lemma we present an improvement of the Mond-Pečarić method.

Lemma 10.3 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in an unital C^*-algebra $\mathscr{A}$ with the spectra in $[m,M]$, $m < M$, defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ and $(\Phi_t)_{t\in T}$ be an unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Then*

$$\int_T \Phi_t(f(x_t))\,d\mu(t) \leq \alpha_f \int_T \Phi_t(x_t)\,d\mu(t) + \beta_f \mathbf{1}_K - \delta_f \widetilde{x} \leq \alpha_f \int_T \Phi_t(x_t)\,d\mu(t) + \beta_f \mathbf{1}_K \tag{10.58}$$

for every continuous convex function $f : [m,M] \to \mathbb{R}$*, where* $\delta_f \equiv \delta_f(m,M)$ *is defined by* (10.53)*,*

$$\widetilde{x} \equiv \widetilde{x}_{x_t,\Phi_t}(m,M) := \frac{1}{2}\mathbf{1}_K - \frac{1}{M-m}\int_T \Phi_t\left(|x_t - \frac{m+M}{2}\mathbf{1}_H|\right) d\mu(t) \tag{10.59}$$

and $\alpha_f = \dfrac{f(M)-f(m)}{M-m}$, $\beta_f = \dfrac{Mf(m)-mf(M)}{M-m}$ *(the same as in Chapter 9).*

If f is concave, then the reverse inequality is valid in (10.58)*.*

Proof. We prove the convex case only. Since $\mathsf{Sp}(x_t) \subseteq [m,M]$, then by utilizing the functional calculus to (10.57) in Lemma 10.2, we obtain

$$f(x_t) \le \frac{M-x_t}{M-m}f(m) + \frac{x_t-m}{M-m}f(M) - \widetilde{x}_t\delta_f(m,M),$$

where

$$\widetilde{x}_t = \frac{1}{2}\mathbf{1}_H - \frac{1}{M-m}\left|x_t - \frac{m+M}{2}\mathbf{1}_H\right|.$$

Applying a positive linear mapping Φ_t, integrating and using that $\int_T \Phi_t(\mathbf{1}_H)\,d\mu(t) = \mathbf{1}_K$, we get the first inequality in (10.58), since

$$\int_T \Phi_t(\widetilde{x}_t)\,d\mu(t) = \frac{1}{2}\mathbf{1}_K - \frac{1}{M-m}\int_T \Phi_t\left(|x_t - \frac{m+M}{2}\mathbf{1}_H|\right) d\mu(t) = \widetilde{x}.$$

Also, $m\mathbf{1}_H \le x_t \le M\mathbf{1}_H$, $t \in T$, implies $\int_T \Phi_t\left(|x_t - \frac{m+M}{2}\mathbf{1}_H|\right) d\mu(t) \le \frac{M-m}{2}\mathbf{1}_K$. It follows $\widetilde{x} \ge 0$. Then the second inequality in (10.58) holds, since $\delta_f\widetilde{x} \ge 0$. □

Now, we present a refinement of Jensen's inequality.

Theorem 10.7 *Let $(x_t)_{t\in T}$ be a bounded continuous field of self-adjoint elements in a unital C^*-algebra $\mathscr{A}$ defined on a locally compact Hausdorff space T equipped with a bounded Radon measure μ. Let m_t and M_t, $m_t \le M_t$, be the bounds of x_t, $t \in T$. Let $(\Phi_t)_{t\in T}$ be a unital field of positive linear mappings $\Phi_t : \mathscr{A} \to \mathscr{B}$ from $\mathscr{A}$ to another unital C^*-algebra $\mathscr{B}$. Let*

$$(m_x, M_x) \cap [m_t, M_t] = \emptyset, \quad t \in T, \qquad \text{and} \qquad m < M,$$

where m_x and M_x, $m_x \le M_x$, be the bounds of the operator $x = \int_T \Phi_t(x_t)\,d\mu(t)$ and

$$a = \sup\{M_t : M_t \le m_x, t \in T\}, \quad b = \inf\{m_t : m_t \ge M_x, t \in T\}.$$

If $f : J \to \mathbb{R}$ is a continuous convex function provided that the interval J contains all m_t, M_t, then

$$f\left(\int_T \Phi_t(x_t)\,d\mu(t)\right) \le \int_T \Phi_t(f(x_t))\,d\mu(t) - \delta_f(\overline{m},\overline{M})\overline{x} \le \int_T \Phi_t(f(x_t))\,d\mu(t) \tag{10.60}$$

holds, where $\delta_f(\overline{m},\overline{M})$ *is defined by* (10.53),

$$\bar{x} \equiv \bar{x}_x(\overline{m},\overline{M}) := \frac{1}{2}\mathbf{1}_K - \frac{1}{\overline{M}-\overline{m}}\left| x - \frac{\overline{m}+\overline{M}}{2}\mathbf{1}_K \right| \tag{10.61}$$

and $\overline{m} \in [a, m_x]$, $\overline{M} \in [M_x, b]$, $\overline{m} < \overline{M}$, *are arbitrary numbers. If* f *is concave, then the reverse inequality is valid in* (10.60).

Proof. We prove only the convex case. Since $x = \int_T \Phi_t(x_t)\,d\mu(t) \in \mathscr{B}$ is the self-adjoint elements such that $\overline{m}\mathbf{1}_K \le m_x\mathbf{1}_K \le \int_T \Phi_t(x_t)\,d\mu(t) \le M_x\mathbf{1}_K \le \overline{M}\mathbf{1}_K$ and f is convex on $[\overline{m},\overline{M}] \subseteq J$, then by Lemma 10.2 we obtain

$$f\left(\int_T \Phi_t(x_t)\,d\mu(t)\right) \le \frac{\overline{M}\mathbf{1}_K - \int_T \Phi_t(x_t)\,d\mu(t)}{\overline{M}-\overline{m}} f(\overline{m}) + \frac{\int_T \Phi_t(x_t)\,d\mu(t) - \overline{m}\mathbf{1}_K}{\overline{M}-\overline{m}} f(\overline{M}) - \delta_f \bar{x}, \tag{10.62}$$

where $\delta_f \equiv \delta_f(\overline{m},\overline{M})$ and $\bar{x}$ are defined by (10.53) and (10.61), respectively. On the other hand, since $(m_x, M_x) \cap [m_t, M_t] = \emptyset$ implies $(\overline{m},\overline{M}) \cap [m_t, M_t] = \emptyset$ and f is convex on $[m_t, M_t]$, then

$$f(x_t) \ge \frac{\overline{M}\mathbf{1}_H - x_t}{\overline{M}-\overline{m}} f(\overline{m}) + \frac{x_t - \overline{m}\mathbf{1}_H}{\overline{M}-\overline{m}} f(\overline{M}), \quad t \in T.$$

Applying a positive linear mapping Φ_t, integrating and adding $-\delta_f\bar{x}$, we obtain

$$\int_T \Phi_t\left(f(x_t)\right) d\mu(t) - \delta_f\bar{x} \ge \frac{\overline{M}\mathbf{1}_K - \int_T \Phi_t(x_t)\,d\mu(t)}{\overline{M}-\overline{m}} f(\overline{m}) + \frac{\int_T \Phi_t(x_t)\,d\mu(t) - \overline{m}\mathbf{1}_K}{\overline{M}-\overline{m}} f(\overline{M}) - \delta_f\bar{x}. \tag{10.63}$$

Combining two inequalities (10.62) and (10.63), we have LHS of (10.60). Also, since $\delta_f \ge 0$ and $\bar{x} \ge 0$, we have RHS of (10.60). □

Finally, we present a refinement of (9.88).

Theorem 10.8 *Let* $(x_t)_{t\in T}$, $(\Phi_t)_{t\in T}$, m, M, $\delta_f(m,M)$, $\widetilde{x}$, α_f *and* β_f *be as in Lemma 10.3. Let* m_x *and* M_x, $m_x \le M_x$, *be the bounds of the operator* $x = \int_T \Phi_t(x_t)\,d\mu(t)$, *and* $m_{\widetilde{x}}$ *be the lower bound of the operator* $\widetilde{x}$.

Let $f : [m,M] \to \mathbb{R}$, $g : [m_x, M_x] \to \mathbb{R}$, $F : U \times V \to \mathbb{R}$, *where* $f([m,M]) \subseteq U$, $g([m_x,M_x]) \subseteq V$ *and* F *be bounded. If* f *is convex and* F *is operator monotone in the first variable, then*

$$\begin{aligned} &F\left[\int_T \Phi_t(f(x_t))\,d\mu(t),\, g\Big(\int_T \Phi_t(x_t)\,d\mu(t)\Big)\right] \\ &\le F\left[\alpha_f x + \beta_f - \delta_f(m,M)\widetilde{x},\, g\Big(\int_T \Phi_t(x_t)\,d\mu(t)\Big)\right] \\ &\le \sup_{m_x \le z \le M_x} F\left[\alpha_f z + \beta_f - \delta_f(m,M)m_{\widetilde{x}}, g(z)\right] \mathbf{1}_K \le \sup_{m_x \le z \le M_x} F\left[\alpha_f z + \beta_f, g(z)\right] \mathbf{1}_K. \end{aligned} \tag{10.64}$$

If f *is concave, then the opposite inequalities are valid in* (10.64) *with* inf *instead of* sup.

Proof. We only prove the case when f is convex. Then $\delta_f(m,M) \geq 0$ implies $0 \leq \delta_f(m,M)\, m_{\widetilde{x}}\, \mathbf{1}_K \leq \delta_f(m,M)\, \widetilde{x}$. By using (10.58) it follows

$$\int_T \Phi_t(f(x_t))\, d\mu(t) \leq \alpha_f x + \beta_f - \delta_f(m,M)\, \widetilde{x} \leq \alpha_f x + \beta_f - \delta_f(m,M)\, m_{\widetilde{x}} \mathbf{1}_K \leq \alpha_f x + \beta_f.$$

Taking into account operator monotonicity of $F(\cdot,v)$ in the first variable, we obtain (10.64). □

Example 10.4 *We give examples for the matrix cases and $T = \{1,2\}$. We put $F(u,v) = u - v$, $f(t) = t^4$ which is convex but not operator convex, and $g \equiv f$. As a special case of* (10.64)*, we have*

$$\Phi_1(X_1^4) + \Phi_2(X_2^4) \leq \left(\Phi_1(X_1) + \Phi_2(X_2)\right)^4 + \overline{C} I_2 - \delta_f(m,M)\widetilde{X} \leq \left(\Phi_1(X_1) + \Phi_2(X_2)\right)^4 + \overline{C} I_2, \tag{10.65}$$

where

$$\overline{C} = \max_{m_x \leq z \leq M_x} \left\{ \frac{M^4 - m^4}{M - m} z + \frac{Mm^4 - mM^4}{M - m} - z^4 \right\}.$$

Also, we define mappings $\Phi_1, \Phi_2 : M_3(\mathbb{C}) \to M_2(\mathbb{C})$: $\Phi_1((a_{ij})_{1 \leq i,j \leq 3}) = \frac{1}{2}(a_{ij})_{1 \leq i,j \leq 2}$, $\Phi_2 \equiv \Phi_1$.

I) *First, we observe an example without the spectra condition. Then we obtain a refined inequality* (10.65)*, but Jensen's inequality doesn't hold.*

$$\text{If} \quad X_1 = 2\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad X_2 = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{then } X = 2\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and $m_1 = -1.604$, $M_1 = 4.494$, $m_2 = 0$, $M_2 = 2$, $m = -1.604$, $M = 4.494$ (rounded to three decimal places). We have

$$(\Phi_1(X_1) + \Phi_2(X_2))^4 = \begin{pmatrix} 16 & 0 \\ 0 & 0 \end{pmatrix} \not\geq\not\leq \begin{pmatrix} 80 & 40 \\ 40 & 24 \end{pmatrix} = \Phi_1\left(X_1^4\right) + \Phi_2\left(X_2^4\right)$$

and

$$\begin{aligned} \begin{pmatrix} 80 & 40 \\ 40 & 24 \end{pmatrix} &= \Phi_1\left(X_1^4\right) + \Phi_2\left(X_2^4\right) \\ < \begin{pmatrix} 111.742 & 39.327 \\ 39.327 & 142.858 \end{pmatrix} &= \Phi_1\left(X_1^4\right) + \Phi_2\left(X_2^4\right) + \overline{C} I_2 - \delta_f \widetilde{X} \\ < \begin{pmatrix} 243.758 & 0 \\ 0 & 227.758 \end{pmatrix} &= (\Phi_1(X_1) + \Phi_2(X_2))^4 + \overline{C} I_2, \end{aligned}$$

since $\overline{C} = 227.758$, $\delta_f = 405.762$, $\widetilde{X} = \begin{pmatrix} 0.325 & -0.097 \\ -0.097 & 0.2092 \end{pmatrix}$

II) *Next, we observe an example with the spectra condition. Then we obtain a series of inequalities involving the refined Jensen's inequality:*

$$\left(\Phi_1(X_1) + \Phi_2(X_2)\right)^4 \leq \Phi_1(X_1^4) + \Phi_2(X_2^4) - \delta_f(\overline{m}, \overline{M})\overline{X} \leq \Phi_1(X_1^4) + \Phi_2(X_2^4)$$

and its converse (10.65).

$$\textit{If } X_1 = \begin{pmatrix} -4 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{pmatrix} \textit{ and } X_2 = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}, \textit{ then } X = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and $m_1 = -4.866$, $M_1 = -0.345$, $m_2 = 1.345$, $M_2 = 5.866$, $m = -4.866$, $M = 5.866$, $a = -0.345$, $b = 1.345$ *and we put* $\overline{m} = a$, $\overline{M} = b$ *(rounded to three decimal places). We have*

$$\begin{aligned}
&\begin{pmatrix} 0.0625 & 0 \\ 0 & 0 \end{pmatrix} = (\Phi_1(X_1) + \Phi_2(X_2))^4 \\
&< \begin{pmatrix} 639.921 & -255 \\ -255 & 117.856 \end{pmatrix} = \Phi_1\left(X_1^4\right) + \Phi_2\left(X_2^4\right) - \delta_f(a,b)\overline{X} \\
&< \begin{pmatrix} 641.5 & -255 \\ -255 & 118.5 \end{pmatrix} = \Phi_1\left(X_1^4\right) + \Phi_2\left(X_2^4\right) \\
&< \begin{pmatrix} 731.649 & -162.575 \\ -162.575 & 325.15 \end{pmatrix} = (\Phi_1(X_1) + \Phi_2(X_2))^4 + \overline{C}I_2 - \delta_f(m,M)\widetilde{X} \\
&< \begin{pmatrix} 872.471 & 0 \\ 0 & 872.409 \end{pmatrix} = (\Phi_1(X_1) + \Phi_2(X_2))^4 + \overline{C}I_2,
\end{aligned}$$

since $\delta_f(a,b) = 3.158$, $\overline{X} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.204 \end{pmatrix}$, $\delta_f(m,M) = 1744.82$, $\widetilde{X} = \begin{pmatrix} 0.325 & -0.097 \\ -0.097 & 0.2092 \end{pmatrix}$ *and* $\overline{C} = 872.409$.

10.6 Notes

The idea of Jensen's inequality without operator convexity is given by Mićić, Pavić and Pečarić [187]. The application on the power operator means is presented in [188]. Extensions of the previous results are given by the same authors in [190]. A refinement of Jensen's inequality and its converses based on research by Mićić, Pečarić and J. Perić is presented in Section 10.5. The interested reader can find additional results in [186, 194, 195, 196].

Chapter 11

Bohr's Inequality

The classical inequality of Bohr says that $|a+b|^2 \le p|a|^2 + q|b|^2$ for all scalars $a,\ b$ and $p,\ q > 0$ with $1/p + 1/q = 1$. The equality holds if and only if $(p-1)a = b$.

In this chapter, we observe some operator versions of Bohr's inequality. Using a general result involving matrix ordering, we derive several inequalities of Bohr's type. Furthermore, we present an approach to Bohr's inequality based on a generalization of the parallelogram theorem with absolute values of operators. Finally, applying Jensen's operator inequality we get a generalization of Bohr's inequality.

11.1 Bohr's inequalities for operators

Let H be a complex separable Hilbert space and $B(H)$ the algebra of all bounded operators on H. The absolute value of $A \in B(H)$ is denoted by $|A| = (A^*A)^{1/2}$.

The classical inequality of Bohr [24] says that

$$|a+b|^2 \le p|a|^2 + q|b|^2$$

for all scalars $a,\ b$ and $p,\ q > 0$ with $1/p + 1/q = 1$. The equality holds if and only if $(p-1)a = b$.

For this, Hirzallah [145] proposed an operator version of Bohr's inequality:

Theorem 11.1 *If A and B are operators on a Hilbert space, and $q \geq p > 0$ satisfy $1/p + 1/q = 1$, then*

$$|A-B|^2 + |(p-1)A+B|^2 \leq p|A|^2 + q|B|^2.$$

Afterwards, several authors have presented generalizations of Bohr's inequality.

Theorem 11.2 *If $A, B \in B(H)$, $\frac{1}{p}+\frac{1}{q} = 1$, and $1 < p \leq 2$, i.e. $q \geq p > 1$, then*

(i) $|A-B|^2 + |(p-1)A+B|^2 \leq p|A|^2 + q|B|^2$,

(ii) $|A-B|^2 + |A+(q-1)B|^2 \geq p|A|^2 + q|B|^2$.

On the other hand, if $p < 1$, then

(iii) $|A-B|^2 + |(p-1)A+B|^2 \geq p|A|^2 + q|B|^2$.

Theorem 11.3 *If $A, B \in B(H)$ and $|\alpha| \geq |\beta|$, then*

$$|A-B|^2 + \frac{1}{|\alpha|^2}\Big| |\beta|A + |\alpha|B \Big|^2 \leq \left(1+\frac{|\beta|}{|\alpha|}\right)|A|^2 + \left(1+\frac{|\alpha|}{|\beta|}\right)|B|^2.$$

We note that it unifies the following inequalities:

(i) If $\alpha \geq |\beta| = -\beta$, then

$$|A-B|^2 + \left|\frac{|\beta|}{\alpha}A + B\right|^2 \leq \left(1+\frac{|\beta|}{\alpha}\right)|A|^2 + \left(1+\frac{\alpha}{|\beta|}\right)|B|^2.$$

(ii) If $0 < \alpha \leq -\beta$, then

$$|A-B|^2 + \left|\frac{\alpha}{|\beta|}A + B\right|^2 \leq \left(1+\frac{\alpha}{|\beta|}\right)|A|^2 + \left(1+\frac{|\beta|}{\alpha}\right)|B|^2.$$

Next we state Bohr's inequalities for multi-operators.

Theorem 11.4 *Suppose that $A_i \in B(H)$, and $r_i \geq 1$ for $i = 1, 2, \cdots, n$ with $\sum_{i=1}^{n} \frac{1}{r_i} = 1$. Then*

$$\left|\sum_{i=1}^{n} A_i\right|^2 \leq \sum_{i=1}^{n} r_i |A_i|^2.$$

In other words, it says that $K(z) = |z|^2$ satisfies Jensen's (operator) inequality:

$$K\left(\sum_{i=1}^{n} t_i\, A_i\right) \leq \sum_{i=1}^{n} t_i\, K(A_i)$$

holds for $t_1, \cdots, t_n > 0$ with $\sum_{i=1}^{n} t_i = 1$.

11.2 Matrix approach to Bohr's inequalities

In this section, we present an approach to Bohr's inequalities by the use of the matrix order.

For this, we introduce two notations: For $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, we define an $n \times n$ matrix $\Lambda(x) = x^* x = (x_i x_j)$ and $D(x) = \operatorname{diag}(x_1, \cdots, x_n)$.

Theorem 11.5 *If $\Lambda(a) + \Lambda(b) \le D(c)$ for $a, b, c \in \mathbb{R}^n$, then*

$$\Big|\sum_{i=1}^n a_i A_i\Big|^2 + \Big|\sum_{i=1}^n b_i A_i\Big|^2 \le \sum_{i=1}^n c_i |A_i|^2$$

for arbitrary n-tuple (A_i) in $B(H)$. Incidentally, the statement is correct even if the order is replaced by the reverse.

Proof. We define a positive mapping Φ of $B(H)^n$ to $B(H)$ by

$$\Phi(X) = (A_1^* \cdots A_n^*) X^{\,T} (A_1 \cdots A_n),$$

where $\cdot^T$ denotes the transpose operation. Since $\Lambda(a) = (a_1, \cdots, a_n)^T (a_1, \cdots, a_n)$, we have

$$\Phi(\Lambda(a)) = \Big(\sum_{i=1}^n a_i A_i\Big)^* \Big(\sum_{i=1}^n a_i A_i\Big) = \Big|\sum_{i=1}^n a_i A_i\Big|^2,$$

so that

$$\Big|\sum_{i=1}^n a_i A_i\Big|^2 + \Big|\sum_{i=1}^n b_i A_i\Big|^2 = \Phi\Big(\Lambda(a) + \Lambda(b)\Big) \le \Phi(D(c)) = \sum_{i=1}^n c_i |A_i|^2.$$

The additional part is easily shown by the same way. □

The meaning of Theorem 11.5 will be well explained in the following theorem.

Theorem 11.6 *Let $t \in \mathbb{R}$.*

(i) *If* $0 < t \le 1$, *then* $|A \mp B|^2 + |tA \pm B|^2 \le (1+t)|A|^2 + \Big(1 + \frac{1}{t}\Big)|B|^2.$

(ii) *If* $t \ge 1$ *or* $t < 0$, *then* $|A \mp B|^2 + |tA \pm B|^2 \ge (1+t)|A|^2 + \Big(1 + \frac{1}{t}\Big)|B|^2.$

Proof. We apply Theorem 11.5 to $a = (1, \mp 1)$, $b = (t, \pm 1)$ and $c = (1+t, 1+1/t)$. Then we consider the order between corresponding matrices:

$$T = \begin{pmatrix} 1+t & 0 \\ 0 & 1+\frac{1}{t} \end{pmatrix} - \begin{pmatrix} 1 & \mp 1 \\ \mp 1 & 1 \end{pmatrix} - \begin{pmatrix} t^2 & \pm t \\ \pm t & 1 \end{pmatrix} = (1-t)\begin{pmatrix} t & \pm 1 \\ \pm 1 & \frac{1}{t} \end{pmatrix}.$$

Since $\det(T) = 0$, T is positive semidefinite (resp. negative semidefinite) if $0 < t < 1$ (resp. $t > 1$ or $t < 0$). □

Remark 11.1 *It is important that all of theorems cited in Section 11.1 follow from Theorem 11.6 easily. For instance, for* (i) *and* (iii) *of Theorem 11.2, we take* $t = p-1$. *For* (ii), *we take* $t = q-1$ *and permute* A *and* B. *Also, Theorem 11.3 follows taking* $t = |\beta|/|\alpha|$.

As another application of Theorem 11.5, we give a proof of Theorem 11.4.

Proof of Theorem 11.4. We check the order between the corresponding matrices $D = \text{diag}(r_1, \cdots, r_n)$ and $C = (c_{ij})$ where $c_{ij} = 1$. All principal minors of $D - C$ are nonnegative and it follows that $C \le D$. Really, for natural numbers $k \le n$, put $D_k = \text{diag}(r_{i_1}, \cdots, r_{i_k})$, $C_k = (c_{ij})$ with $c_{ij} = 1$, $i, j = 1, \ldots, k$ and $R_k = \sum_{j=1}^{k} 1/r_{i_j}$ where $1 \le r_{i_1} < \cdots < r_{i_k} \le n$. Then

$$\det(D_k - C_k) = (r_{i_1} \cdot \cdots \cdot r_{i_k})(1 - R_k) \ge 0 \quad \text{for arbitrary } k \le n.$$

Hence we have the conclusion by Theorem 11.5. □

In the remainder, we cite additional results obtained by Theorem 11.5.

Corollary 11.1 *If* $a = (a_1, a_2)$, $b = (b_1, b_2)$ *and* $p = (p_1, p_2)$ *satisfy*

$$p_1 \ge a_1^2 + b_1^2, \quad p_2 \ge a_2^2 + b_2^2, \quad (p_1 - (a_1^2 + b_1^2))(p_2 - (a_2^2 + b_2^2)) \ge (a_1 a_2 + b_1 b_2)^2,$$

then

$$|a_1 A + a_2 B|^2 + |b_1 A + b_2 B|^2 \le p_1 |A|^2 + p_2 |B|^2$$

holds all $A, B \in B(H)$.

Proof. Since the assumption of the above is nothing but the matrix inequality $\Lambda(a) + \Lambda(b) \le D(p)$, Theorem 11.5 implies the conclusion. □

Finally, we remark the monotonicity of the operator function $F(a) = \left| \sum_{i=1}^{n} a_i A_i \right|^2$.

Corollary 11.2 *For a fixed n-tuple* (A_i) *in* $B(H)$, *the operator function* $F(a) = \left| \sum_{i=1}^{n} a_i A_i \right|^2$ *for* $a = (a_1, \cdots, a_n)$ *is order preserving, that is,*

$$\text{if} \quad \Lambda(a) \le \Lambda(b), \quad \text{then} \quad F(a) \le F(b).$$

Proof. We prove this putting $F(a) = \Phi(a^* a)$, where Φ is a positive linear mapping as in the proof of Theorem 11.5. □

Corollary 11.3 *If* $a = (a_1, a_2, a_3)$ *and* $b = (b_1, b_2, b_3)$ *satisfy* $|a_i| \le |b_i|$ *for* $i = 1, 2, 3$ *and* $a_i b_j = a_j b_i$ *for* $i \ne j$, *then* $F(a) \le F(b)$.

Proof. It follows from assumptions that if $i \ne l$ and $j \ne k$, then

$$\begin{vmatrix} a_i a_j - b_i b_j & a_i a_k - b_i b_k \\ a_l a_j - b_l b_j & a_l a_k - b_l b_k \end{vmatrix} = a_k b_j (a_i b_l - b_i a_l) + a_j b_k (b_i a_l - a_i b_l) = 0.$$

This means that all 2nd principal minors of $\Lambda(b) - \Lambda(a)$ are zero. It follows that $\det(\Lambda(b) - \Lambda(a)) = 0$. Since the diagonal elements satisfy $|a_i| \le |b_i|$ for $i = 1, 2, 3$, we have the matrix inequality $\Lambda(a) \le \Lambda(b)$. Now it is sufficient to apply Corollary 11.2. □

11.3 Generalized parallelogram law for operators

Next we give another approach to Bohr's inequality. In our frame, the following generalization of the parallelogram law easily implies Theorem 11.6 which covers many previous results as discussed in the preceding section.

Theorem 11.7 *If A and B are operators on a Hilbert space and $t \neq 0$, then*

$$|A+B|^2 + \frac{1}{t}|tA-B|^2 = (1+t)|A|^2 + \left(1+\frac{1}{t}\right)|B|^2.$$

Proof. It is easily checked that

$$\begin{aligned} & |A+B|^2 + \frac{1}{t}|tA-B|^2 \\ = \ & |A|^2 + |B|^2 + A^*B + B^*A + t|A|^2 + \frac{1}{t}|B|^2 - A^*B - B^*A \\ = \ & (1+t)|A|^2 + \left(1+\frac{1}{t}\right)|B|^2. \end{aligned}$$

□

Remark 11.2 *We immediately obtain Theorem 11.6 by noting the condition of t in Theorem 11.7. This means that Theorems 11.1 and 11.2 also follow from Theorem 11.7 .*

Next we extend Theorem 11.7 for several operators.

Theorem 11.8 *Suppose that $A_i \in B(H)$ and $r_i \geq 1$ with $\sum_{i=1}^{n} \frac{1}{r_i} = 1$ for $i = 1, 2, ..., n$. Then*

$$\sum_{i=1}^{n} r_i |A_i|^2 - |\sum_{i=1}^{n} A_i|^2 = \sum_{1 \leq i < j \leq n} \left| \sqrt{\frac{r_i}{r_j}} A_i - \sqrt{\frac{r_j}{r_i}} A_j \right|^2.$$

Proof. We show it by the induction on n. Note that it is true for $n = 2$ by Theorem 11.7, because it is expressed as follows: Let $A_i \in B(H)$ and $r_i \geq 1$ for $i = 1, 2$ satisfying $\frac{1}{r_1} + \frac{1}{r_2} = 1$. Then

$$r_1|A_1|^2 + r_2|A_2|^2 - |A_1 + A_2|^2 = \left| \sqrt{\frac{r_1}{r_2}} A_1 - \sqrt{\frac{r_2}{r_1}} A_2 \right|^2.$$

Now suppose that it is true for $n = k$, then we take $A_1, \cdots, A_{k+1} \in B(H)$ and $r_1, \cdots, r_{k+1} > 1$ satisfying $\sum_{i=1}^{k+1} \frac{1}{r_i} = 1$. Here we put $r_i' = r_i \left(1 - \frac{1}{r_{k+1}}\right)$ for $i = 1, \cdots, k$ and $B = \sum_{i=1}^{k} A_i$ for convenience, then $r_i' > 1$ and $\sum_{i=1}^{k} \frac{1}{r_i'} = 1$. Hence we have

$$
\begin{aligned}
&\sum_{i=1}^{k+1} r_i|A_i|^2 - \left|\sum_{i=1}^{k+1} A_i\right|^2 \\
&= \sum_{i=1}^{k} r_i|A_i|^2 + r_{k+1}|A_{k+1}|^2 - \left|\sum_{i=1}^{k} A_i + A_{k+1}\right|^2 \\
&= \left(1-\frac{1}{r_{k+1}}\right)\sum_{i=1}^{k} r_i|A_i|^2 - |B|^2 + (r_{k+1}-1)|A_{k+1}|^2 + \frac{1}{r_{k+1}}\sum_{i=1}^{k} r_i|A_i|^2 - B^*A_{k+1} - A_{k+1}^*B \\
&= \left(\sum_{i=1}^{k} r_i'|A_i|^2 - |B|^2\right) + \sum_{i=1}^{k} \frac{r_i}{r_{k+1}}|A_i|^2 - B^*A_{k+1} - A_{k+1}^*B + (r_{k+1}-1)|A_{k+1}|^2 \\
&= \sum_{1\le i<j\le k} \left|\sqrt{\frac{r_i}{r_j}}A_i - \sqrt{\frac{r_j}{r_i}}A_j\right|^2 + \sum_{i=1}^{k} \frac{r_i}{r_{k+1}}|A_i|^2 - B^*A_{k+1} - A_{k+1}^*B + \sum_{i=1}^{k} \frac{r_{k+1}}{r_i}|A_{k+1}|^2 \\
&= \sum_{1\le i<j\le k} \left|\sqrt{\frac{r_i}{r_j}}A_i - \sqrt{\frac{r_j}{r_i}}A_j\right|^2 + \sum_{i=1}^{k+1} \left|\sqrt{\frac{r_i}{r_{k+1}}}A_i - \sqrt{\frac{r_{k+1}}{r_i}}A_{k+1}\right|^2 \\
&= \sum_{1\le i<j\le k+1} \left|\sqrt{\frac{r_i}{r_j}}A_i - \sqrt{\frac{r_j}{r_i}}A_j\right|^2 .
\end{aligned}
$$

Therefore, the required equality holds for all $n \in \mathbb{N}$. □

Remark 11.3 *Theorem 11.4 is an easy consequence of Theorem 11.8.*

Incidentally, we note that the condition $r_i \geq 1$ in Theorem 11.8 is not necessary. As a matter of fact, we can show the following.

Corollary 11.4 *Let $A_i \in B(H)$ and $r_i \neq 0$ for $i = 1,2,...,n$ with $\sum_{i=1}^{n} \frac{1}{r_i} = 1$. Then*

$$
\sum_{i=1}^{n} r_i|A_i|^2 - \left|\sum_{i=1}^{n} A_i\right|^2 = \sum_{1\le i\le j\le n} \frac{r_j}{r_i}\left|\frac{r_i}{r_j}A_i - A_j\right|^2 .
$$

11.4 The Dunkl-Williams inequality

Dunkl and Williams showed that

$$
\left\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\right\| \leq \frac{4\|x-y\|}{\|x\|+\|y\|}
$$

for every nonzero element x, y in a normed linear space.

Pečarić and Rajić gave the following refinement: For every nonzero element x, y in a normed linear space X

$$\left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\| \leq \frac{\sqrt{2\|x-y\|^2 + 2(\|x\| - \|y\|)^2}}{\max\{\|x\|, \|y\|\}}.$$

Furthermore they generalized it to an operator inequality as follow:

Theorem 11.9 *Let $A, B \in B(H)$ be operators where $|A|$ and $|B|$ are invertible, and let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

$$\left| A|A|^{-1} - B|B|^{-1} \right|^2 \leq |A|^{-1} \left(p|A-B|^2 + q(|A| - |B|)^2 \right) |A|^{-1}.$$

The equality holds if and only if

$$p\,(A-B)|A|^{-1} = q\,B\left(|A|^{-1} - |B|^{-1}\right).$$

Very recently, Saito-Tominaga improved Theorem 11.9 without the assumption of the invertibility of the absolute value of operators.

Theorem 11.10 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$, and let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\left| (U - V)|A| \right|^2 \leq p|A-B|^2 + q\left(|A| - |B|\right)^2.$$

The equality holds if and only if

$$p\,(A-B) = q\,V\;(|B| - |A|) \quad \text{and} \quad V^*V \geq U^*U.$$

In this section, we consider the Dunkl-Williams inequality for operators as an application of generalized parallelogram law of operators in Theorem 11.7:

$$|A-B|^2 + \frac{1}{t}|tA+B|^2 = (1+t)|A|^2 + \left(1 + \frac{1}{t}\right)|B|^2$$

for any nonzero $t \in \mathbb{R}$.

The following lemma follows from it easily.

Lemma 11.1 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$. Then for each $t > 0$*

$$|A-B|^2 \leq (1+t)|A|^2 + \left(1 + \frac{1}{t}\right)|B|^2.$$

The equality holds for t if and only if $tA + B = 0$.

We prepare another lemma.

Lemma 11.2 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$ and $t > 0$. If $t(A-B) + V(|A| - |B|) = 0$ is satisfied, then*

$$t|A-B|^2 \leq |A|^2 - |B|^2,$$

and so $|A| \geq |B|$ and $U^*U \geq V^*V$.

In addition, if $U^*U = V^*V$, then $t|A-B|^2 = |A|^2 - |B|^2$.

Proof. Since $tA - (t+1)B = -V|A|$ by the assumption, we have

$$\left|tA - (t+1)B\right|^2 = |A|V^*V|A|.$$

Adding $t|A|^2 - (t+1)|B|^2$ to both sides, it follows that

$$t(t+1)|A-B|^2 = |A|V^*V|A| + t|A|^2 - (t+1)|B|^2 \leq (t+1)\left(|A|^2 - |B|^2\right),$$

so that

$$0 \leq t|A-B|^2 \leq |A|^2 - |B|^2.$$

Hence it follows that $|A| \geq |B|$ and $U^*U \geq V^*V$. Moreover, if $U^*U = V^*V$ is assumed, then $V^*V|A| = |A|$ and so

$$t|A-B|^2 = |A|^2 - |B|^2.$$

□

The following theorem is proved by the lemmas cited above, and it changes to Theorem 11.10 by taking $t = p - 1$.

Theorem 11.11 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$, and $t > 0$. Then*

$$\left|(U-V)|A|\right|^2 \leq (t+1)|A-B|^2 + \left(1 + \frac{1}{t}\right)(|A| - |B|)^2.$$

The equality holds if and only if

$$t(A-B) = V(|B| - |A|) \quad \textit{and} \quad V^*V = U^*U.$$

Proof. We replace A and B in Lemma 11.1 by $A - B$ and $V(|A| - |B|)$, respectively. Then we have the required inequality, and the condition for which the equality holds is that

$$t(A-B) = V(|B| - |A|) \quad \text{and} \quad V^*V = U^*U.$$

The latter in above is equivalent to $|A|V^*V|A| = |A|^2$, that is, $V^*V \geq U^*U$. By the help of Lemma 11.2, it becomes $V^*V = U^*U$. □

Lemma 11.3 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$, and $t > 0$. Suppose that $V^*V = U^*U$. Then*

$$t(A-B) = V(|B| - |A|)$$

if and only if

$$|A| = |B| + t|A-B| \text{ and } A - B = -V|A-B|.$$

Proof. Let $t(A-B) = -V(|A| - |B|)$. It follows from Lemma 11.2 that

$$t|A-B| = \Big| |A| - |B| \Big| = |A| - |B|$$

and moreover

$$A - B = \frac{1}{t} V(|B| - |A|) = -\frac{1}{t}\, tV|A-B| = -V|A-B|.$$

Conversely, let $|A| - |B| = t|A-B|$ and $A - B = -V|A-B|$. Then

$$t(A-B) + V(|A| - |B|) = -tV|A-B| + tV|A-B| = 0.$$

□

Lemma 11.4 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$, and $t > 0$. Suppose that $V^*V = U^*U$. If $t(A-B) = V(|B| - |A|)$, then*

$$|B||A-B| + |A-B||B| = (1-t)|A-B|^2.$$

Proof. Put $C = A - B$. The preceding lemma ensures that

$$t|C| = |B+C| - |B| \text{ and } C = -V|C|.$$

Then it follows that

$$|B+C| = |B| + t|C|,$$

and that

$$B^*C = -B^*V|C| = -(|B|V^*V)|C| = -|B||C|.$$

Hence we have

$$|B+C|^2 = (|B| - |C|)^2 \text{ and } |B+C|^2 = (|B| + t|C|)^2,$$

so that

$$(t+1)\big(|B||C| + |C||B|\big) = (1-t^2)|C|^2,$$

which is equivalent to the conclusion. □

Theorem 11.12 *Let $A, B \in B(H)$ be operators with polar decompositions $A = U|A|$ and $B = V|B|$, and $C = A - B = W|C|$ the polar decomposition of C. Assume that the equality*

$$\big|(U-V)|A|\big|^2 = (t+1)|A-B|^2 + \left(1 + \frac{1}{t}\right)\big(|A| - |B|\big)^2.$$

holds for some $t > 0$.

(1) *If* $t \geq 1$, *then* $A = B$.

(2) *If* $0 \leq t \leq 1$, *then*

$$A = B\left(I - \frac{2}{1-t}W^*W\right) \quad \textit{and} \qquad |A| = |B|\left(I + \frac{2t}{1-t}W^*W\right),$$

and the converse is true.

Proof. The preceding lemma leads us the fact that if positive operators S and T satisfy $ST + TS = rS^2$ for some $r \in \mathbb{R}$, then

$$\text{(i)}\ \ S = 0 \quad \text{if } r < 0, \qquad \text{and} \qquad \text{(ii)}\ \ S \text{ and } T \text{ commute} \quad \text{if } r \geq 0.$$

(Since $S^2T = STS - tS^3$ is self-adjoint, S^2 commutes with T and so does S.) Thus we apply it for $S = |A-B||C|$, $T = |B|$ and $r = 1-t$.

(1) Since $r = 1-t \leq 0$, we first suppose that $r < 0$. Then $S = |A-B| = 0$, that is, $A = B$, as desired. Next we suppose $r = 0$. Then $S = |C|$ commutes with $T = |B|$ and so $ST = 0$. Hence we have $|C|V^*V = 0$. Moreover, since $C = -V|C|$ by Lemma 11.3, it follows that $|C|^2 = |C|V^*V|C| = 0$, i.e. $C = 0$.

(2) We apply (ii). Namely we have

$$|B||C| = |C||B| = \frac{1-t}{2}|C|^2,$$

so that

$$B|C| = V|B||C| = \frac{1-t}{2}V|C|^2 = \frac{t-1}{2}C|C| = \frac{t-1}{2}A|C| - \frac{t-1}{2}B|C|.$$

It implies that

$$A|C| = \frac{2}{t-1}\left(1 + \frac{t-1}{2}\right)B|C| = \frac{t+1}{t-1}B|C|,$$

and so

$$AW^*W = \frac{t+1}{t-1}BW^*W.$$

Therefore we have

$$A = AW^*W + A(I - W^*W) = \frac{t+1}{t-1}BW^*W + B(I - W^*W) = B\left(I + \frac{2}{t-1}W^*W\right).$$

For the second equality, it suffices to show that W^*W commutes with $|B|$ because

$$\left|I - \frac{2}{1-t}W^*W\right| = I + \frac{2t}{1-t}W^*W$$

is easily seen. For this commutativity, we note that $C = A - B = \frac{2}{t-1}BW^*W$ by the first equality, $C = -V|C|$ by Lemma 11.3, and $V^*V \geq W^*W$ by $W^*W \leq \sup\{V^*V, U^*U\}$ and $V^*V = U^*U$. So we prove that

$$|B|W^*W = V^*BW^*W = -\frac{1-t}{2}V^*C = \frac{1-t}{2}V^*V|C| = \frac{1-t}{2}|C|.$$

Incidentally the converse implication in (2) is as follows: We first note that the second equality assures the commutativity of $|B|$ and W^*W. Next it follows that

$$|A| - |B| = -\frac{2t}{1-t}|B|W^*W$$

and

$$V|A| - B = V(|A| - |B|) = -\frac{2t}{1-t}BW^*W = -t(A-B)$$

by the first equality. Hence we have

$$(U-V)|A| = A - V|A| = A + t(A-B) - B = (1+t)(A-B)$$

and so

$$\big|(U-V)|A|\big|^2 = (1+t)^2|A-B|^2.$$

On the other hand, since

$$(|A| - |B|)^2 = \left(\frac{2t}{1-t}\right)^2 B^*BW^*W = t^2|A-B|^2$$

we have

$$\begin{aligned} &(1+t)|A-B|^2 + (1+\frac{1}{t})(|A|-|B|)^2 \\ &= \left((1+t) + \left(1+\frac{1}{t}\right)t^2\right)|A-B|^2 = (1+t)^2|A-B|^2, \end{aligned}$$

which completes the proof. □

11.5 From Jensen's inequality to Bohr's inequality

As an application of Jensen's inequality, in this section we consider a generalization of Bohr's inequality. Namely Jensen's inequality implies Bohr's inequality even in the operator case.

For this, we first target the following inequality which is an extension of Bohr's inequality, precisely, it is a multiple version of Bohr's inequality in the case $r = 2$:
If $r > 1$ and $a_1, \cdots, a_n > 0$, then

$$\left|\sum_{i=1}^n z_i\right|^r \leq \left(\sum_{i=1}^n a_i^{\frac{1}{1-r}}\right)^{r-1} \sum_{i=1}^n a_i|z_i|^r$$

for all $z_1, \cdots, z_n \in \mathbb{C}$.

We note that it follows from Hölder inequality. Actually, $p = \frac{r}{r-1}$ and $q = r$ are conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. We here set

$$u_i = a_i^{-\frac{1}{q}}, \quad w_i = u_i^{-1} z_i \qquad (i = 1, 2, \cdots, n)$$

and apply them to Hölder inequality. Then we have

$$\left|\sum_{i=1}^n z_i\right|^r = \left|\sum_{i=1}^n u_i w_i\right|^r \le \left(\sum_{i=1}^n |u_i|^p\right)^{\frac{r}{p}} \left(\sum_{i=1}^n |w_i|^q\right)^{\frac{r}{q}} = \left(\sum_{i=1}^n a_i^{\frac{1}{1-r}}\right)^{r-1} \sum_{i=1}^n a_i |z_i|^r.$$

Now we propose its operator extension. For the sake of convenience, we recall Jensen's inequality (see also Remark 10.1 with conditions on spectra) for our use below:

Let f be an operator convex function on an interval J, let T be a locally compact Hausdorff space with a bounded Radon measure μ, and let $\mathscr{A}$ and $\mathscr{B}$ be unital C^*-algebras. If $(\psi_t)_{t\in T}$ is a unital field of positive linear mappings of $\mathscr{A}$ to $\mathscr{B}$, then

$$f\left(\int_T \psi_t(x_t)\, d\mu(t)\right) \le \int_T \psi_t(f(x_t))\, d\mu(t)$$

holds for bounded continuous fields $(x_t)_{t\in T}$ of self-adjoint elements in $\mathscr{A}$ whose spectra are contained in J.

Theorem 11.13 *Let T be a locally compact Hausdorff space with a bounded Radon measure μ, and let $\mathscr{A}$ and $\mathscr{B}$ be unital C^*-algebras. If $1 < r \le 2$, $a : T \to \mathbb{R}$ is a bounded continuous positive function and $(\phi_t)_{t\in T}$ is a field of positive linear mappings $\phi_t : \mathscr{A} \to \mathscr{B}$ satisfying*

$$\int_T a(t)^{\frac{1}{1-r}} \phi_t(\mathbf{1}) d\mu(t) \le \int_T a(t)^{\frac{1}{1-r}} d\mu(t)\, \mathbf{1}, \tag{11.1}$$

then

$$\left(\int_T \phi_t(x_t) d\mu(t)\right)^r \le \left(\int_T a(t)^{\frac{1}{1-r}} d\mu(t)\right)^{r-1} \int_T a(t) \phi_t(x_t^r) d\mu(t) \tag{11.2}$$

holds for all continuous fields $(x_t)_{t\in T}$ of positive elements in $\mathscr{A}$.

Proof. We set $\psi_t = \dfrac{1}{M} a(t)^{\frac{1}{1-r}} \phi_t$, where $M = \displaystyle\int_T a(t)^{\frac{1}{1-r}} d\mu(t) > 0$. Then we have $\displaystyle\int_T \psi_t(\mathbf{1}) d\mu(t) \le \mathbf{1}$. By a routine way, we may assume that $\displaystyle\int_T \psi_t(\mathbf{1}) d\mu(t) = \mathbf{1}$. Since $f(t) = t^r$ is operator convex for $1 < r \le 2$, then we applying Jensen's inequality cited above and obtain

$$\left(\int_T \frac{1}{M} a(t)^{\frac{1}{1-r}} \phi_t(\tilde{x}_t) d\mu(t)\right)^r \le \int_T \frac{1}{M} a(t)^{\frac{1}{1-r}} \phi_t(\tilde{x}_t^r) d\mu(t)$$

for every bounded continuous fields $(\tilde{x}_t)_{t\in T}$ of positive elements in $\mathscr{A}$. Replacing $\tilde{x}_t$ by $a(t)^{-1/(1-r)} x_t$, the above inequality can be written as

$$\left(\int_T \phi_t(x_t) d\mu(t)\right)^r \le M^{r-1} \int_T a(t) \phi_t(x_t^r) d\mu(t)$$

which is the desired inequality. □

Remark 11.4 *We can obtain the inequality in a broader region for r under conditions on spectra. Let*

$$(m_x, M_x) \cap [a(t)^{-1/(1-r)} m_t, a(t)^{-1/(1-r)} M_t] = \emptyset, \qquad t \in T,$$

where m_x and M_x, $m_x \le M_x$, be the bounds of $x = \int_T \phi_t(x_t)\, d\mu(t)$ and m_t and M_t, $m_t \le M_t$, be the bounds of x_t, $t \in T$. If the condition (11.1) *is valid, then the inequality* (11.2) *holds for every $r \in (-\infty, 0) \cup (1, \infty)$.*

The following corollary is a discrete version of Theorem 11.13.

Corollary 11.5 *If $1 < r \le 2$, $a_1, \cdots, a_n > 0$ and positive linear mappings $\phi_1, \cdots, \phi_n$ on $B(H)$ satisfy*

$$\sum_{i=1}^{n} a_i^{\frac{1}{1-r}} \phi_i(I) \le \sum_{i=1}^{n} a_i^{\frac{1}{1-r}} I,$$

then

$$\left(\sum_{i=1}^{n} \phi_i(A_i)\right)^r \le \left(\sum_{i=1}^{n} m a_i^{\frac{1}{1-r}}\right)^{r-1} \sum_{i=1}^{n} a_i \phi_i(A_i^r)$$

holds for positive operators $A_1, \cdots, A_n \ge 0$ on H.

For a typical positive linear mapping $\phi(A) = X^* AX$ for some X, the preceding corollary is written as follows:

Corollary 11.6 *If $1 < r \le 2$, and $a_1, \cdots, a_n > 0$ and $X_1, \cdots, X_n$ in $B(H)$ satisfy*

$$\sum_{i=1}^{n} a_i^{\frac{1}{1-r}} X_i^* X_i \le \sum_{i=1}^{n} a_i^{\frac{1}{1-r}} I,$$

then

$$\left(\sum_{i=1}^{n} X_i^* A_i X_i\right)^r \le \left(\sum_{i=1}^{n} a_i^{\frac{1}{1-r}}\right)^{r-1} \sum_{i=1}^{n} a_i X_i^* A_i^r X_i$$

holds for positive operators $A_1, \cdots, A_n \ge 0$ on H.

11.6 Notes

The original inequality of Bohr [24] was established for scalars in 1929. Hirzallah [145] posed an operator version of it. Afterwards, Cheung-Pečarić [32], Zhang [295] and several authors have considered extensions of Bohr's inequality for operators. Very recently, such study has been done by Abramovich-Barić-Pečarić [1], and Fujii-Zuo [105], in which matrix order method is proposed.

The Dunkl-Williams inequality in a normed space was established in [47]. Pečarić-Rajić [250] presented an operator version of it, which was generalized by Saito-Tominaga [256]. Moreover it was discussed in [39] from the viewpoint of generalized parallelogram law for operators. Such operator versions are regarded as applications of Bohr operator inequality.

The results in 8.6 depend on [232], in which Jensen's inequality we used is appeared in [135].

Bibliography

[1] S. Abramovich, J. Barić and J. Pečarić, *A new proof of an inequality of Bohr for Hilbert space operators*, Linear Alg. Appl. **430** (2009), 1432–1435.

[2] M. Alić, P.S. Bullen, J. Pečarić and V. Volenec, *On the geometric-arithmetic mean inequality for matrices*, Math. Commun. **2** (1997), 125–128.

[3] M. Alić, B. Mond, J. Pečarić and V. Volenec, *Bounds for the differences of matrix means*, SIAM J. Matrix Anal. Appl. **18** (1997), 119–123.

[4] G. Alpargu, *The Kantorovich inequality with some extensions and with some stastistical applications*, M. Sc. thesis, McGill University, Montréal, Canada (1996), 1–113.

[5] S. Amari, *Differential Geometrical Methods in Statistics*, Lecture Notes in Statist. **28** (1985).

[6] T. Ando, *Topics on Operator Inequalities*, Lecture notes (mimeographed), Hokkaido Univ., Sapporo, 1978.

[7] T. Ando, *Concavity of certain maps on positive definite matrices and applications to Hadamard products*, Linear Algebra Appl. **26** (1979), 203–241.

[8] T. Ando, *On some operator inequality*, Math. Ann. **279** (1987), 157–159.

[9] T. Ando, *Majorization and inequalities in Matrix Theory*, Linear Algebra Appl. **199** (1994), 17–67.

[10] T. Ando, *Matrix Young inequalities*, Operator Theory: Adv. and Appl. **75** (1995), 33–38.

[11] T. Ando, *Quasi-orders on the positive cone of a C^*-algebra*, Linear Multilinear Algebra **41** (1996), 81–94.

[12] T. Ando and F. Hiai, *Log-majorization and complementary Golden-Thompson type inequalities*, Linear Algebra Appl. **197–198** (1994), 113–131.

[13] T. Ando, C.-K. Li and R. Mathias, *Geometric means*, Linear Algebra Appl. **385** (2004), 305–334.

[14] E. Andruchow, G. Corach and D. Stojanoff, *Geometrical significance of the Löwner-Heinz inequality*, Proc. Amer. Math. Soc. **128** (1999), 1031–1037.

[15] J.S. Aujla and H.L. Vasudeva, *Operator inequalities related to means of operators*, Math. Japon. **41** (1995), 383–388.

[16] J.S. Aujla and H.L. Vasudeva, *Inequalities involving Hadamard product and operator means*, Math. Japon. **42** (1995), 265–272.

[17] E. Bach and T. Furuta, *Order preserving operator inequalities*, J. Operator Theory **19** (1988), 341–346.

[18] G. Bachman and L. Narici, *Functional Analysis*, Academic Press Inc., London, 1966.

[19] N. Bebiano, R. Lemos and J. Providência, *Inequalities for quantum relative entropy*, Linear Algebra Appl. 401 (2005), 159–172.

[20] J. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc. **79** (1955), 58–71.

[21] K.V. Bhagwat and R. Subramanian, *Inequalities between means of positive operators*, Math. Proc. Camb. Phil. Soc. **83** (1978), 393–401.

[22] R. Bhatia, *Matrix Anaysis*, Springer, New York, 1997.

[23] R. Bhatia and T. Sano, *Loewner matrices and operator convexity*, Math. Ann. **344** (2009), 703–716.

[24] H. Bohr, *Zur theorie der fastperiodischen funktionmen I*, Acta Math. **45** (1929), 29–127.

[25] J.-C. Bourin, *Reverse inequality to Araki's inequality comparison of* $A^pZ^pA^p$ *and* $(AZA)^p$, Math. Inequal. Appl. **8** (2005), 373–378.

[26] J.-C. Bourin, *Symmetric norms and reverse inequalities to Davis and Hansen-Pedersen characterizations of operator convexity*, Math. Inequal. Appl. **9** (2006), 33–42.

[27] J.-C. Bourin, *Matrix versions of some classical inequalities*, Linear Algebra Appl. **416** (2006), 890–907.

[28] J.-C. Bourin, E.-Y. Lee, M.Fujii and Y.Seo, *A matrix reverse Hölder inequality*, Linear Algebra Appl., **431** (2009), 2154–2159.

[29] J.-C. Bourin and Y. Seo, *Reverse inequality to Golden-Thompson type inequalities: Comparison of* e^{A+B} *and* e^Ae^B, Linear Algebra Appl. **426** (2007), 312–316.

[30] G.T. Cargo and O. Shisha, *Bounds on ratios of means*, J.Res. Natl. Bur. **66B** (1962), 167–170.

[31] N.N. Chan and M.K. Kwong, *Hermitian matrix inequalities and a conjecture*, Amer. Math. Monthly **92** (1985), 533–541.

[32] W.-S. Cheung and J. Pečarić, *Bohr's inequalities for Hilbert space operators*, J. Math. Anal. Appl. **323** (2006), 403–412.

[33] M. Cho, T. Furuta, J.I. Lee and W.Y. Lee, *A folk theorem on Furuta inequality*, Sci. Math. **3** (2000), 229–231.

[34] M.D. Choi, *Positive linear maps on C^*-algebras*, Canad. J. Math. **24** (1972), 520–529.

[35] M.D. Choi, *A Schwarz inequality for positive linear maps on C^*-algebras*, Illinois J. Math. **18** (1974), 565–574 .

[36] M.D. Choi, *Positive linear maps*, Proc. Symp. Pure Math. **38** (1982), Part 2, 583–590.

[37] G. Corach, H. Porta and L. Recht, *Geodesics and operator means in the space of positive operators*, Int. J. Math. **4** (1993), 193–202.

[38] G. Corach, H. Porta and L. Recht, *Convexity of the geodesic distance on spaces of positive operators*, Illinois J. Math. **38** (1994), 87–94.

[39] F. Dadipour, M. Fujii and M.S. Moslehian, *An operator version of the Dunkl-Williams inequality with respect to the p-angular distance*, Nihonkai Math. J. **21** (2010), 11–20.

[40] C. Davis, *A Schwartz inequality for convex operator functions*, Proc. Amer. Math. Soc. **8** (1957), 42–44.

[41] C. Davis, *Operator-valued entropy of quantum mechanical measurement*, Proc. Japan Acad. **37** (1961), 533–538.

[42] C. Davis, *Notions generalizing convexity for functions defined on spaces of matrices*, Proc. Symp. Pure Math. **7** (1963), 187–201.

[43] J.B. Diaz and F.T. Metcalf, *Stronger forms of a class of inequalities of G.Pólya-G.Szegö, and L.V. Kantorovich*, Bull. Amer. Math. Soc. **69** (1963), 415–418.

[44] O. Došlý and Á. Elbert, *Integral characterization of the principal solution of half-linear second order differential equations*, Studia Sci. Math. Hungar. **36** (2000), 455–469.

[45] R.G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press, New York, 1972.

[46] R. Drnovšek and T. Kosem, *Inequalities between $f(\|A\|)$ and $\|f(|A|)\|$*, Math. Ineq. Appl. **8** (2005), 1–6.

[47] C.F. Dunkl and K.S. Williams, *A simple norm inequality*, Amer. Math. Monthly. **71** (1964), 53–54.

[48] Ky Fan, *Some matrix inequalities*, Abh. Math. Sem. Univ. Hamburg **29** (1966), 185–196.

[49] J.I. Fujii, *On geometric and harmonic mean of operators*, Math. Japon. **24** (1979), 203–207.

[50] J.I. Fujii, *Operator means and the relative operator entropy*, Operator Theory: Advances and Applications, Birkhäuser, **59** (1992), 161–172.

[51] J.I. Fujii, *The relative operator entropy via operator means* (in Japanese), Thesis, 1993.

[52] J.I. Fujii, *The Marcus-Khan theorem for Hilbert space operators*, Math. Japon. **41** (1995), 531–535.

[53] J.I. Fujii, *On the relative operator entropy (in Japanese)*, RIMS Kokyuroku **903** (1995), 49–56.

[54] J.I. Fujii, *Interpolational path and quasi-supremum*, RIMS Kokyuroku **1359** (2004), 1–8.

[55] J.I. Fujii, *Structure of Hiai-Petz parametrized geometry for positive definite matrices*, Linear Algebra Appl. **432** (2010), 318–326.

[56] J.I. Fujii., *Path of quasi-means as a geodesic*, Linear Algebra Appl. **434** (2011), 542–558.

[57] J.I. Fujii and M. Fujii, *Jensen's inequalities on any interval for operators*, RIMS Kokyuroku **1359** (2004), 83–102.

[58] J.I. Fujii, M. Fujii, S. Izumino, F. Kubo and R. Nakamoto, *Strang's inequality*, Math. Japon. **37** (1992), 479–486.

[59] J.I. Fujii, M. Fujii and R. Nakamoto, *Riccati equation and positivity of operator matrices*, Kyungpook Math.J. **49** (2009), 595–603.

[60] J.I. Fujii, M. Fujii, M. Nakamura, J.E. Pečarić and Y. Seo, *A reverse of the weighted geometric mean due to Lawson-Lim*, Linear Algebra Appl. **427** (2007), 272–284.

[61] J.I. Fujii, M. Fujii and Y. Seo, *An extension of the Kubo-Ando theory: solidarities*, Math. Japon. **35** (1990), 387–396.

[62] J.I. Fujii, M. Fujii and Y. Seo, *Bounds for interpolational paths of positive operators*, *Bounds for interpolational path of positive operators*, Recent Topics on Operator Inequalities, RIMS Kokyuroku **1359** (2004), 46-57.

[63] J.I. Fujii, M. Fujii and Y. Seo, *The Golden-Thompson-Segal type inequalities related to the weighted geometric mean due to Lawson-Lim*, Math. Inequal. Appl. **3** (2009), 511–518

[64] J.I. Fujii, M. Fujii, Y. Seo and M. Tominaga, *On generalized Kantorovich inequalities*, Proceeding International Sym. Banach and Fuction spaces, October 2-4 (2003), 205–213.

[65] J.I. Fujii, T.Furuta, T. Yamazaki and M. Yanagida, *Simplified proof of characterization of chaotic order via Specht's ratio*, Sci. Math. **2** (1999), 63–64.

[66] J.I. Fujii, S. Izumino and Y. Seo, *Determinant for positive operators and Specht's Theorem*, Sci. Math. **1** (1998), 307–310.

[67] J.I. Fujii and E. Kamei, *Relative operator entropy in noncommutative information theory*, Math. Japon. **34** (1989), 341–348.

[68] J.I. Fujii and E. Kamei, *Uhlmann's interpolational method for operator means*, Math. Japon. **34** (1989), 541–547.

[69] J.I. Fujii and E. Kamei, *Interpolational paths and their derivatives*, Math. Japon. **39** (1994), 557–560.

[70] J.I. Fujii and Y.O. Kim, *A Kantorovich type inequality with a negative parameter*, Sci. Math. Jpn., On line e-2009 **22** (2009), 87–92.

[71] J.I. Fujii, J. Mićić, J. Pečarić and Y. Seo, *Comparison of operator mean geodesics*, J. Math. Inequal. **2** (2008), 287–298.

[72] J.I. Fujii, M. Nakamura, J. Pečarić and Y. Seo, *Bounds for the ratio and difference between parallel sum and series via Mond-Pečarić method*, Math. Inequal. Appl. **9** (2006), 749–759.

[73] J.I. Fujii and Y. Seo, *Determinant for positive operators*, Sci. Math. **1** (1998), 153–156.

[74] J.I. Fujii and Y. Seo, *On parametrized operator means dominated by power ones*, Sci. Math. **1** (1998), 301–306.

[75] J.I. Fujii and Y. Seo, *Characterizations of chaotic order associated with the Mond-Shisha difference*, Math. Inequal. Appl. **5** (2002), 725–734.

[76] J.I. Fujii, Y. Seo and M. Tominaga, *Kantorovich type operator inequalities via the Specht ratio*, Linear Algebra Appl. **337** (2004), 69–81.

[77] J.I. Fujii, Y. Seo and M. Tominaga, *Kantorovich type reverse inequalities for operator norm*, Math. Inequal. Appl. **8** (2005), 529–535.

[78] M. Fujii, *On operator concavity related to means of operators*, Math. Japon. **30** (1985), 283–288.

[79] M. Fujii, *Furuta's inequality and its mean theoretic approach*, J. Operator Theorey **23** (1990), 67–72.

[80] M. Fujii, T. Furuta and E. Kamei, *Furuta's inequality and its application to Ando's theorem*, Linear Algebra Appl. **179** (1993), 161–169.

[81] M. Fujii and F. Kubo, *Around Jensen's inequality II*, Math. Japon. **27** (1982), 495–499.

[82] M. Fujii, T. Furuta and R. Nakamoto, *Norm inequalities in the Corach-Porta-Recht theory and operator means*, Illinois J.Math. **40** (1996), 527–534.

[83] M. Fujii, T. Furuta, R. Nakamoto and S.-E. Takahasi, *Operator inequalities and covariance in noncommutative probability*, Math. Japon. **46** (1997), 317–320.

[84] M. Fujii, M. Hashimoto, Y. Seo and M. Yanagida, *Characterizations of usual and chaotic order via Furuta and Kantorovich inequalities*, Sci. Math. **3** (2000), 405–418.

[85] M. Fujii, S. Izumino, R. Nakamoto and Y. Seo, *Operator inequalities related to Cauchy-Schwarz and Hölder-McCarthy inequalities*, Nihonkai Math. J. **8** (1997), 117–122.

[86] M. Fujii, J.-F. Jiang and E. Kamei, *Characterization of chaotic order and its application to Furuta inequality*, Proc. Amer. Math. Soc. **125** (1997), 3655–3658.

[87] M. Fujii, J.-F. Jiang, E. Kamei and K. Tanahashi, *A characterization of chaotic order and a problem*, J. Inequal. Appl. **2** (1998), 149–156.

[88] M. Fujii and E. Kamei, *Furuta's inequality for the chaotic order*, Math. Japon. **36** (1991), 603–606.

[89] M. Fujii and E. Kamei, *Mean theoretic approach to the grand Furuta inequality*, Proc. Amer. Math. Soc. **124** (1996), 2751–2756.

[90] M. Fujii and E. Kamei, *Ando-Hiai inequality and Furuta inequality*, Linear Algebra Appl. **416** (2006), 541–545.

[91] M. Fujii and E. Kamei, *Variants of Ando-Hiai inequality*, Operator Theory: Adv. Appl. **187** (2008), 169–174.

[92] M. Fujii, E .Kamei and Y. Seo, *Kantorovich type operator inequalities via grand Furuta inequality*, Sci. Math. **3** (2000), 263–272.

[93] M.Fujii, E.-Y.Lee and Y.Seo, *A difference counterpart to a matrix Hölder inequality*, Linear Algebra Appl. **432** (2010), 2565–2571.

[94] M. Fujii, S.H. Lee, Y. Seo and D. Jung, *Reverse inequalities on chaotically geometric mean via Specht ratio*, Math. Inequal. Appl. **6** (2003), 509–519.

[95] M. Fujii, A. Matsumoto and R. Nakamoto, *A short proof of the best possibility for the grand Furuta inequality*, J. Inequal. Appl. **4** (1999), 339–344.

[96] M. Fujii, J. Mićić, J.E. Pečarić and Y. Seo, *Reverse inequalities on chaotically geometric mean via Specht ratio, II*, J. Inequal. Pure and Appl. Math. **4** (2) Article 40 (2003), 1–8.

[97] M. Fujii and R. Nakamoto, *Norm inequalities in the Corach-Porta-Recht theory and operator means*, Illinois J. Math. **40** (1996), 527–534.

[98] M. Fujii and R. Nakamoto, *A geometric mean in the Furuta inequality*, Sci. Math. Jpn. **55** (2002), 615–621.

[99] M. Fujii, R. Nakamoto and M. Nakamura, *Conditional expectation and Hadamard product of operators*, Math. Japon. **42** (1995), 239–244.

[100] M. Fujii, R. Nakamoto and M. Tominaga, *Generalized Bebiano-Lemos-Providência inequalities and their reverses*, Linear Algebra Appl. **426** (2007), 33–39.

[101] M. Fujii, M. Nakamura, J. Pečarić and Y. Seo, *Bounds for the ratio and difference between parallel sum and series via Mond-Pečarić method*, Math. Inequal. Appl. **9** (2006), 749–759.

[102] M. Fujii and Y. Seo, *Reverse inequalities of Araki, Cordes and Löwner-Heinz inequalities*, Nihonkai Math. J. **16** (2005), 145–154.

[103] M. Fujii and Y. Seo, *Several reverse inequalities of operators*, RIMS Kokyuroku **1452** (2005), 30–37.

[104] M. Fujii, Y. Seo and M. Tominaga, *Golden-Thompson type inequalities related to a geometric mean via Specht's ratio*, Math. Inequal. Appl. **5** (2002), 573–582.

[105] M. Fujii and H. Zuo, *Matrix order in Bohr inequality for operators*, Banach J. Math. Anal. **4** (2010), 21–27.

[106] T. Furuta, $A \geq B \geq 0$ *assures* $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$ *for* $r \geq 0,\ p \geq 0,\ q \geq 1$ *with* $(1+2r)q \geq p+2r$, Proc. Amer. Math. Soc. **101** (1987), 85–88.

[107] T. Furuta, *Elementary proof of an order preserving inequality*, Proc. Japan Acad. **65** (1989), 126.

[108] T. Furuta, *Applications of order preserving operator inequalities*, Operator Theorey: Advances and Applications **59** (1992), 180–190.

[109] T. Furuta, *Extension of the Furuta inequality and Ando-Hiai log-majorization*, Linear Algebra Appl. **219** (1995), 139–155.

[110] T. Furuta, *Generalizations of Kosaki trace inequalities and related trace inequalities on chaotic order*, Linear Algebra Appl. **235** (1996), 153–161.

[111] T. Furuta, *Extensions of Hölder-McCarthy and Kantorovich inequalities and their applications*, Proc. Japan Acad., Ser.A **73** (1997), 38–41.

[112] T. Furuta, *Applications of order preserving operator inequalities to a generalized relative operator entropy*, General Inequalities **7**, Birkhäuser, **123** (1997), 65–76.

[113] T. Furuta, *Operator inequalities associated with Hölder-McCarthy and Kantorovich inequalities*, J. Inequal. Appl. **2** (1998), 137–148.

[114] T. Furuta, *Simplified proof of an order preserving operator inequality*, Proc. Japan. Acad. **74** (1998), 114.

[115] T. Furuta, *Simple proof of the concavity of operator entropy* $f(A) = -A\log A$, Math. Inequal. Appl. **3** (2000), 305–306.

[116] T. Furuta, *Results under* $\log A \geq \log B$ *can be derived from ones under* $A \geq B \geq 0$ *by Uchiyama's method - associated with Furuta and Kantorovich type operator inequalities*, Math. Inequal. Appl. **3** (2000), 423–436.

[117] T. Furuta, *Invitation to Linear Operators*, Taylor&Francis, London, 2001.

[118] T. Furuta, $A \geq B \geq 0$ *ensures* $(A^{r/2}A^pA^{r/2})^{1/q} \geq (A^{r/2}B^pA^{r/2})^{1/q}$ *for* $p \geq 0$*,* $q \geq 1$*,* $r \geq 0$ *with* $(1+r)q \geq p+r$ *and its applications*, Sci. Math. Jpn. **53** (2001), 555–602.

[119] T. Furuta, *Operator inequalities associated with* $A\log A$ *via Specht ratio*, Linear Algebra Appl. **375** (2003), 251–273.

[120] T. Furuta, *Basic property of the generalized Kantorovich constant* $K(h,p) = \frac{h^p-h}{(p-1)(h-1)}\left(\frac{p-1}{p}\frac{h^p-1}{h^p-h}\right)^p$ *and its applications*, Acta Aci. Math. (Szeged) **70** (2004), 319–337.

[121] T. Furuta, *Specht ratio* $S(1)$ *can be expressed by Kantorovich constant* $K(p)$*:* $S(1) = \exp K'(1)$ *and its applications*, Math. Inequal. Appl. **6** (2003), 521–530.

[122] T. Furuta and M. Giga, *A complementary result of Kantorovich type order preserving inequalities by Mićić-Pečarić-Seo*, Linear Algebra Appl. **369** (2003), 27–40.

[123] T. Furuta, M. Hashimoto and M. Ito, *Equivalence relation between generalized Furuta inequality and related operator functions*, Sci. Math. **1** (1998), 257–259.

[124] T. Furuta, J. Mićić Hot, J. Pečarić and Y. Seo, *Mond-Pečarić Method in Operator Inequalities*, Monographs in Inequalities **1**, Element, Zagreb, 2005.

[125] T. Furuta and J. Pečarić, *An operator inequality associated with the operator concavity of operator entropy* $A\log A^{-1}$, Math. Inequal. Appl. **6** (2003), 531–543.

[126] T. Furuta and Y. Seo, *An application of generalized Furuta inequality to Kantorovich type inequalities*, Sci. Math. **2** (1999), 393–399.

[127] T. Furuta and D. Wang, *A decreasing operator function assocaited with the Furuta inequality*, Proc. Amer. Math. Soc. **126** (1998), 2427–2432.

[128] T. Furuta, T. Yamazaki and M. Yanagida, *Operator functions implying generalized Furuta inequality*, Math. Inequal. Appl. **1** (1998), 123–130.

[129] M. Giga, *Kantorovich type inequalities for* $1 > p > 0$, J. Inequal. Pure and Appl. Math. **4** (2) Article 105 (2003), 1–6.

[130] M. Giga, *The equivalence among three Kantorovich type inequalities*, RIMS Kokyuroku **1427** (2005), 103–113.

[131] S. Golden, *Lower bounds for Helmholtz function*, Phys. Rev. **137** (1965), B1127–B1128.

[132] W. Greub and W. Rheinboldt, *On a generalization of an inequality of L.V. Kantorovich*, Proc. Amer. Math. Soc. **10** (1959), 407–415.

[133] F. Hansen, *An operator inequality*, Math. Ann. **246** (1980), 249–250.

[134] F. Hansen, J. Pečarić and I. Perić, *A generalization of disrete Jensen's operator inequality and its converses*, Hungrvic-Croatian mathematices, workshop, Debrecen, 2005.

[135] F. Hansen, J. Pečarić and I. Perić, *Jensen's operator inequality and its converse*, Math. Scad. **100** (2007), 61–73.

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

[137] F. Hansen and G.K. Pederson, *Jensen's operator inequality*, Bull. London Math. Soc. **35** (2003), 553–564.

[138] G. Hardy, J.E. Littlewood and G. Pólya, *Inequalities* (Reprinted), Cambridge Univ. Press, Cambridge, 1967.

[139] M. Hashimoto and T. Yamazaki, *Further extensions of characterizations of chaotic order associated with Kantorovich type inequalities*, Sci. Math. **3** (2000), 127–136.

[140] E. Heinz, *Beitrage zur Storungstheorie der Spectralzegung*, Math. Ann. **123** (1951), 415–438.

[141] P. Henrici, *Two remarks on the Kantorovich inequality*, Amer. Math. Monthly **68** (1961), 904–906.

[142] F. Hiai and D. Petz, *The Golden-Thompson trace inequality is complemented*, Linear Algebra Appl. **181** (1993), 153–185.

[143] F. Hiai and K. Yanagi, *Hilbert Spaces and Linear Operators*, Makino shoten (in Japanese), 1995.

[144] H. Hidaka, Solution of 5666, *Extension of Kantorovich's Inequality*, Amer. Math. Monthly **77** (1970), 411–412.

[145] O. Hirzallah, *Non-commutative operator Bohr inequalities*, J. Math. Anal. Appl., **282** (2003), 578–583.

[146] R.A. Horn and C.R. Johnson, *Matric Analysis*, Cambridge University Press, 1985.

[147] S. Izumino and R. Nakamoto, *Functional orders of positive operators induced from Mond-Pečarić convex inequalities*, Sci. Math. **2** (1999), 195–200.

[148] S. Izumino and M. Nakamura, *Kadison's Schwarz and Kantorovich inequalities on correlation operators*, Toyama Math. J. **28** (2005), 133–138.

[149] R.V. Kadison, *A generalized Schwarz inequality and algebraic invariants for operator algebras*, Ann. Math. **56** (1952), 494–503.

[150] D. Kainuma and M. Nakamura, *Around Jensen's inequality*, Math. Japon. **25** (1980), 585–588.

[151] E. Kamei, *A satellite to Furuta's inequality*, Math. Japon. **33** (1988), 883–886.

[152] E. Kamei, *The Furuta inequality and the relative operator entropy*, Thesis, 1993.

[153] E. Kamei, *Paths of operators parametrized by operator means*, Math. Japon. **39** (1994), 395–400.

[154] E. Kamei, *Interpolations between the operator and chaotic orders via Kantorovich type operator inequalities*, Sci. Math. **3** (2000), 257–262.

[155] L.V. Kantorovich, *Functional analysis and applied mathematics* (in Russian), Uspechi Mat. Nauk. **3** (1948), 89–185.

[156] L.V. Kantorovich, *Functional analysis and applied mathematics*, Uspehi Mat. Nauk. **3** (1948), 89–185. Translated from the Russian, National Bureau of Standards, Report 1509, March 7, 1952.

[157] T. Kato, *Notes on some inequalities for linear operators*, Math. Ann. **125** (1952), 208–212.

[158] Y. Kijima, *On some operaror inequalities* (in Japanese), Proc. 1st Symp. Funct. Anal. held at Tokyo Inst. Tech. **1** (1966), 25–29.

[159] Y.O. Kim, *A difference version of Furuta-Giga theorem on Kantorovich type inequality and its application*, Sci. Math. Jpn. **68** (2008), 89–94.

[160] S. Kim and Y. Lim, *A converse inequality of higher-order weighted arithmetic and geometric means of positive definite operators*, Linear Algebra Appl. **426** (2007), 490–496.

[161] S. Kim, H. Lee and Y. Lim, *A sharp converse inequality of three weighted arithmetic and geometric means of positive definite operators*, Math. Inequal. Appl. **12** (2009), pp. 519–523.

[162] K. Kitamura and Y. Seo, *Operator inequalities on Hadamard product associated with Kadison's Schwarz inequalities*, Sci. Math. **1** (1998), 237–241.

[163] M.A. Krasnoselskii and S.G. Krein, *An iteration process with minimal residuals (Russian)*, Mat. Sb. **31** (1952), 315–334.

[164] F. Kubo, *An inequality of Kantorovich* (in Japanese), private note, 1995.

[165] F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann. **246** (1980), 205–224.

[166] S.S. Kutateladze, *The world line of Kantorovich*, Notices from the ISMS, January 2007., 1–7.

[167] E.-Y.Lee, *A matrix reverse Cauchy-Schwarz inequality*, Linear Algebra Appl., **430** (2009), 805–810.

[168] C.-K. Li and R. Mathias, *Matrix inequalities involving a positive linear map*, Linear Multilinear Algebra **41** (1996), 221–231.

[169] C.-S. Lin, *The Furuta inequality and an operator equation for linear operators*, Publ. RIMS, Kyoto Univ. **35** (1999), 309–313.

[170] S. Liu and H. Neudecker, *Several matrix Kantorovich-type inequalities*, J. Math. Anal. Appl. **197** (1996), 23–26.

[171] K. Löwner, *Über monotone Matrix function*, Math. Z. **38** (1934), 177–216.

[172] J. Lawson and Y. Lim, *A general framework for extending means to higher orders*, Colloq. Math. **113** (2008), 191–221.

[173] M. Marcus and N.A. Khan, *A note on the Hadamard product*, Can. Math. Bull. **2** (1959), 81–83.

[174] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1963.

[175] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Mathematics in Science and Engineering, Vol. **143**, Academic Press, 1979.

[176] A. Matković and J. Pečarić, *A variant of Jensen's inequality in Hilbert space*, Toyama M.J. **28** (2005), 93–103.

[177] A. Matković and J. Pečarić, *Chaotic order among power operator means*, Inter. J. Pure and Appl. Math. **34** (2007), 397–406.

[178] A. Matković, J. Pečarić and I. Perić, *A variant of Jensen's inequality of Mercer's type for operators with applications*, Linear Alg. Appl. **418** (2006), 551–564.

[179] A. Matković, J. Pečarić and I. Perić, *Refinements of Jensen's inequality of Mercer's type operator convex functions*, Math. Ineq. Appl. **11** (2008), 113–126.

[180] A. Matsumoto and M. Fujii, *Generalizations of reverse Bebiano-Lemos-Providência inequality*, Linear Algebra Appl. **430** (2009), 1544–1549.

[181] C.A. McCarthy, c_ρ, Israel J. Math. **5** (1967), 249–271.

[182] A. McIntosh, *Heinz inequalities and perturbation of spectral families*, Macquarie Math. Reports, 1979.

[183] A.McD. Mercer, *Improved upper and lower bounds for the difference $A_n - G_n$*, Rocky Mountain J. Math. **31** (2001), 553–560.

[184] A.McD. Mercer, *A monotonicity property of power means*, J. Ineq. Pure and Appl. Math. **3** (3) Article 40 (2002), 1–5.

[185] A.McD. Mercer, *A variant of Jensen's inequality*, J. Ineq. Pure and Appl. Math. **4** (4) Article 73 (2004), 1–5.

[186] J. Mićić and K. Hot, *Inequalities among quasi-arithmetic means for continuous field of operators*, Filomat **26** (2012), 977–991.

[187] J. Mićić, Z. Pavić and J. Pečarić, *Jensen's inequality for operators without operator convexity*, Linear Algebra Appl. **434** (2011), 1228–1237.

[188] J. Mićić, Z. Pavić and J. Pečarić, *Order among power operator means with condition on spectra*, Math. Inequal. Appl. **14** (2011), 709–716.

[189] J. Mićić, Z. Pavić and J. Pečarić, *Some better bounds in converses of the Jensen operator inequality*, Oper. Matrices, **6** (2012), 589–605.

[190] J. Mićić, Z. Pavić and J. Pečarić, *Extension of Jensen's inequality for operators without operator convexity*, Ann. Funct. Anal. **3** (2012), 67–85.

[191] J. Mićić and J. Pečarić, *Order among power means of positive operators*, Sci. Math. Jpn. **61** (2005), 25–46.

[192] J. Mićić and J. Pečarić, *Generalization of the Kantorovich type operator inequalities via grand Furuta inequality*, Math. Inequal. Appl. **9** (2006), 495–510.

[193] J. Mićić and J. Pečarić, *Order among power means of positive operators, II*, Sci. Math. Jpn. **71** (2010), 93–109.; :e2009, 677–693.

[194] J. Mićić and J. Pečarić, *Recent research on Jensen's inequality for operators*, "Linear Algebra - Theorems and Applications", InTech, Rijeka, 2012, 189–214.

[195] J. Mićić, J. Pečarić and J. Perić, *Extension of the refined Jensen's operator inequality with condition on spectra*, Ann. Funct. Anal. **3** (2012), no. 1, 67–85.

[196] J. Mićić, J. Pečarić and J. Perić, *Refined Jensen's operator inequality with condition on spectra*, Oper. Matrices **7** (2013), 293–308.

[197] J. Mićić, J. Pečarić and Y. Seo, *Complementary inequalities to inequalities of Jensen and Ando based on Mond-Pečarić method*, Linear Algebra Appl. **318** (2000), 87–107.

[198] J. Mićić, J. Pečarić and Y. Seo, *Inequalities between operator means based on the Mond-Pečarić method and its applications*, Development of Operator Theory and Problems (Kyoto 2001), RIMS, Kokyuroku **1189** (2001), 1–9.

[199] J. Mićić, J. Pečarić and Y. Seo, *A general form preserving the operator order for convex functions*, Math. Inequal. Appl. **5** (2002), 583–592.

[200] J. Mićić, J. Pečarić and Y. Seo, *Function order of positive operators based on the Mond-Pečarić method*, Linear Algebra Appl. **360** (2003), 15–34.

[201] J. Mićić, J. Pečarić and Y. Seo, *Inequalities between operator means based on the Mond-Pečarić method*, Houston J. Math. **30** (2004), 191–207.

[202] J. Mićić, J. Pečarić and Y. Seo, *Converses of Jensen's operator inequality*, Oper. Matrices **4** (2010), 385–403.

[203] J. Mićić, J. Pečarić and Y. Seo, *An estimate of quasi-arithmetic mean for convex functions*, Sci. Math. Jpn., On line e-2012 **25** (2012), 237–247.

[204] J. Mićić, J. Pečarić and Y. Seo, *Order among quasi-arithmetic means of positive operators*, Math. Rep. **14** (64), (2012), no. 1, 71–86.

[205] J. Mićić, J. Pečarić, Y. Seo and M. Tominaga, *Inequalities for positive linear maps on Hermitian matrices*, Math. Inequal. Appl. **3** (2000), 559–591.

[206] J. Mićić, J. Pečarić and V. Šimić, *Inequalities involving the arithmetic and geometric operator means*, Math. Inequal. Appl. **11** (2008), 415–430.

[207] J. Mićić, Y. Seo, S.-E. Takahasi and M. Tominaga, *Inequalities of Furuta and Mond-Pečarić*, Math. Inequal. Appl. **2** (1999), 83–112.

[208] D.S. Mitrinović, J.E. Pečarić and A.M. Fink, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht, 1993.

[209] B. Mond, *A matrix version of Rennie's generalization of Kantorovich's inequality*, Proc. Amer. Math. Soc. **16** (1965), p. 1131.

[210] B. Mond, *A matrix inequality including that of Kantorovich*, J. Math. Ana. Appl. **13** (1966), 49–52.

[211] B. Mond, *An inequality for operators in a Hilbert space*, Pacif. J. Math. **18** (1966), 161–163.

[212] B. Mond and J. Pečarić, *Converses of Jensen's inequality for linear maps of operators*, Anal. Univ. Timisoara Ser. Mat. Inform. **XXXI** (1993), 223–228.

[213] B. Mond and J. Pečarić, *Convex inequalities in Hilbert space*, Houston J.Math. **19** (1993), 405–420.

[214] B. Mond and J. Pečarić, *Convex inequalities for several self-adjoint operators on a Hilbert space*, Indian J. Math. **35** (1993), 121–135.

[215] B. Mond and J. Pečarić, *Difference and ratio operator inequalities in Hilbert space*, Houston J. Math. **21** (1994), 103–108.

[216] B. Mond and J. Pečarić, *Bounds for Jensen's inequality for several operators*, Houston J. Math. **20** (1994), 645–651.

[217] B. Mond and J. Pečarić, *Converses of Jensen's inequality for several operators*, Rev. Anal. Numér. Théor. Approx. **23** (1994), 179–183.

[218] B. Mond and J. Pečarić, *A matrix version of the Ky Fan generalization of the Kantorovich inequality*, Linear Multilinear Algebra **36** (1994), 217–221.

[219] B. Mond and J. Pečarić, *Jensen's inequality for operator monotone functions*, Rend. Circ. Mat. Palermo **XLIII** (1994), 16–24.

[220] B. Mond and J. Pečarić, *Classical inequalities for matrix functions*, Utilitas Math. **46** (1994), 155–166.

[221] B. Mond and J. Pečarić, *On Jensen's inequality for operator convex functions*, Houston J. Math. **21** (1995), 739–753.

[222] B. Mond and J. Pečarić, *Some matrix inequalities of Ky Fan type*, Tamkang J. Math. **26** (1995), 321–326.

[223] B. Mond and J. Pečarić, *Generalization of a matrix inequality of Ky Fan*, J. Math. Ana. Appl. **190** (1995), 244–247.

[224] B. Mond and J. Pečarić, *Generalized power mean for matrix functions*, Publ. Math. Debrecen **46** (1995), 33–39.

[225] B. Mond and J. Pečarić, *A mixed arithimetic-mean-harmonic-mean matrix inequality*, Linear Algebra Appl. **237/238** (1996), 449–454.

[226] B. Mond and J. Pečarić, *Matrix inequalities for convex functions*, J. Math. Anal. Appl. **209** (1997), 147–153.

[227] B. Mond and J. Pečarić, *Inequalities for the Hadamard product of matrices*, SIAM J. Matrix Anal. Appl. **19** (1998), 66–70.

[228] B. Mond and J. Pečarić, *Operator convex functions of several variables*, Soochow J. Math. **24** (1998), 239–254.

[229] B. Mond, J. Pečarić, J. Šunde and S. Varošanec, *Operator versions of classical inequalities*, Linear Algebra Appl. **264** (1997), 117–126.

[230] B. Mond and O. Shisha, *A difference inequality for operators in Hilbert space*, Blanch Aniversary Volumo Acrospace Res. Lab. U.S.A.F., (1967), 269–275.

[231] B. Mond and O. Shisha, *Difference and ratio inequalities in Hilbert space*, "Inequalities II", (O. Shisha, ed.), Academic Press, New York, 1970, 241–249.

[232] M.S. Moslehian, J. Pečarić and I. Perić, *An operator extension of Bohr's inequality*, Bull. Iranian Math. Soc. **35** (2009), 77–84.

[233] R. Nakamoto, *On the operator equation* $THT = K$, Math. Japon. **18** (1973), 251–252.

[234] R. Nakamoto and M. Nakamura, *Operator mean and Kantorovich inequality*, Math. Japon. **44** (1996), 495–498.

[235] R. Nakamoto and Y. Seo, *Norm inequalities for the geometric mean and its reverse*, Sci. Math. Jpn. **65** (2007), 281–286.

[236] R. Nakamoto and Y. Seo, *A complement of the Ando-Hiai inequality and norm inequalities for the geometric mean*, Nihonkai Math. J. **18** (2007), 43–50.

[237] M. Nakamura, *A remark on a paper of Greub and Rheinboldt*, Proc. Japon. Acad. **36** (1960), 198–199.

[238] M. Nakamura and H. Umegaki, *A note on the entropy for operator algebras*, Proc. Jap. Acad. **37** (1961), 149–154.

[239] C. P. Niculescu, L. E. Persson, *Convex Functions and Their Applications*, Canadian Mathematical Society, 2006.

[240] R.D. Nussbaum, *Hilbert's projective metric and iterated nonlinear maps*, Mem. Amer. Math. Soc. **75** (1988), no. 391, iv+137 pp.

[241] N. Ozeki and K. Ozeki, *Introduction to Inequalities* (in Japanese), Kindai Kagakusya, Tokyo, 1987.

[242] V. Paulsen, *Completely Bounded Maps and Dilations*, Pitman Research Notes in Mathematics Series **146.** Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1986.

[243] J. Pečarić, *Power matrix means and related inequalities*, Math. Commun. **1** (1996), 91–110.

[244] J. Pečarić and J. Mićić, *Chaotic order among of positive operators*, Sci. Math. Jpn. **57** (2003), 139–148.

[245] J. Pečarić and J. Mićić, *Some functions reversing the order of positive operators*, Linear Algebra Appl. **396** (2005), 175–187.

[246] J. Pečarić, J. Mićić and Y. Seo, *Inequalities between operator means based on the Mond-Pečarić method* , Houston J. Math. **30** (2004), 191–207.

[247] J. Pečarić and B. Mond, *A matrix inequality including that of Kantorovich-Hermite, II*, J. Math. Ana. Appl. **168** (1992), 381–384.

[248] J. Pečarić and B. Mond, *The arithmetic mean - the geometric mean and related matrix inequalities*, Internat. Ser. Numer. Math. **123** (1997), 76–91.

[249] J. Pečarić, F. Prochan and Y. Tong, *Convex Functions, Partial Orderings and Statistical Applications*, Academic Press, San Diego, 1992.

[250] J. Pečarić and R. Rajić, *Inequalities of the Dunkl-Williams type for absolute value operators*, J. Math. Inequal. **4** (2010), 1–10.

[251] G.K. Pedersen, *Some operator monotone functions*, Proc. Amer. Math. Soc. **36** (1972), 309–310.

[252] G.K. Pedersen and M. Takesaki, *The operator equation $THT = K$*, Proc. Amer. Math. Soc. **36** (1972), 311-312.

[253] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze der Analysis I*, Springer, Berlin, 1925.

[254] V. Ptak, *The Kantorovich inequality*, Amer. Math. Monthly **102** (1995), 820–821.

[255] B.C. Rennie, *An inequality which includes that of Kantorovich*, Amer. Math. Monthly **70** (1963), 982.

[256] K.-S. Saito and M. Tominaga, *The Dunkl-Williams type inequality for absolute value operators*, Linear Algebra Appl. **432** (2010), 3258–3264.

[257] A.H. Schopf, *On the Kantorovich inequality*, Numer. Math. **2** (1960), 344-346.

[258] P. Schweitzer, *Egy egyenlotlensig az arithmetikai kozepertekrol (An inequality concerning the arithmetic mean)*, Mat. Fyz. Lapok **23** (1914), 257–261.

[259] I. Segal, *Notes toward the construction of nonlinear relativistic quantum fields, III*, Bull. Amer. Math. Soc. **75** (1969), 1390–1395.

[260] Y. Seo, *A characterization of operator order via grand Furuta inequality*, J. Inequal. Appl. **6** (2001), 473–481.

[261] Y. Seo, *Bounds for the ratio and difference between parallel sum and series and noncommutative Kantorovich inequalities*, (Kyoto, 2006), RIMS Kokyuroku **1478** (2006), 71–80.

[262] Y. Seo, *Kantorovich type operator inequalities for Furuta inequality*, Oper. Matrices **1** (2007), 143–152.

[263] Y. Seo, *Norm inequalities for the chaotically geometric mean and its reverse*, J. Math.Inequal. **1** (2007), 39–43.

[264] Y. Seo, *Reverses of the Golden-Thompson type inequalities due to Ando-Hiai-Petz*, Banach J. Math. Anal. **2** (2008), 140–149.

[265] Y. Seo, *Golden-Thompson type inequalities and its reverse*, (Kyoto, 2008), RIMS Kokyuroku **1596** (2008), 42–50.

[266] Y. Seo, *Reverses of Ando's inequality for positive linear maps*, Math. Inequal. Appl. **14** (2011), 905-910.

[267] Y. Seo, S.-E. Takahasi, J. Pečarić and J. Mićić, *Inequalities of Furuta and Mond-Pečarić on the Hadamard product*, J. Inequal. Appl. **5** (2000), 263–285.

[268] Y. Seo and M. Tominaga, *A complement of the Ando-Hiai inequality*, Linear Algebra Appl. **429** (2008), 1546–1554.

[269] O. Shisha and B. Mond, *Bounds on difference of means*, "Inequalities", (O. Shisha, ed.). Academic Press, New York, (1967), 293–308.

[270] P.G. Spain, *Operator versions of the Kantorovich inequality*, Proc. Amer. Math. Soc. **124** (1996), 2813–2819.

[271] W. Specht, *Zur Theorie der elementaren Mittel*, Math. Z. **74** (1960), 91–98.

[272] W.G. Strang, *On the Kantorovich inequality*, Proc. Amer. Math. Soc. **11** (1960), 468.

[273] W. F. Stinespring, *Positive functions on C*-algebras*, Proc. Amer. Math. Soc. **6** (1955), 211–216.

[274] K. Symanzik, *Proof and refinements of an inequality of Feynman*, J.Math. Phys. **6** (1965), 1155–1156.

[275] S.-E. Takahasi, M. Tsukada, K. Tanahashi and T. Ogiwara, *An inverse type of Jensen's inequality*, Math. Japon. **50** (1999), 85–91.

[276] O. Takenouchi, *Functional Analysis*, Asakura Shoten, Tokyo, 1968.

[277] K. Tanahashi, *Best possibility of the Furuta inequality*, Proc. Amer. Math. Soc. **124** (1996), 141–146.

[278] K. Tanahashi, *The best possibility of the grand Furuta inequality*, Proc. Amer. Math. Soc. **128** (2000), 511–519.

[279] C.J. Thompson, *Inequality with applications in statistical mechanics*, J. Math. Phys. **6** (1965), 469–480.

[280] M. Tominaga, *Specht's ratio in the Young inequality*, Sci. Math. Jpn. **55** (2002), 585–588.

[281] M. Tominaga, *The upper bound of a reverse Hölder's type operator inequality and its applications*, J. Inequal. Appl. **7** (2002), 633–645.

[282] H. Toyama, *Theory of Matrices* (in Japanese), Kyoritsu-Shuppan, Tokyo, 1952.

[283] G.E. Trapp, *Hermitian semidefinite matrix means and related matrix inequalities - An introduction*, Linear Multilinear Algebra **16** (1984), 113–123.

[284] M. Tsukada and S.-E. Takahasi, *The best possibility of the bound for the Kantorovich inequality and some remarks*, J. Inequal. Appl. **1** (4) (1997), 327–334.

[285] A.M. Turing, *Rounding off-errors in matrix processes*, Quart. J. Mech. Appl. Math. **1** (1948), 287–308.

[286] M. Uchiyama, *Some exponential operator inequalities*, Math. Inequal. Appl. **2** (1999), 469–471.

[287] M. Uchiyama, *A new majorization between functions, polynomials, and operator inequalities*, J. Funct. Anal. **231** (2006), 231–244.

[288] G.S. Watson, G. Alpargu and G.P.H. Styan, *Some comments on six inequalities associated with the inefficiency of ordinary least squares with one regressor*, Linear Algebra Appl. **264** (1997), 13–53

[289] T. Yamazaki, *Simplified proof of Tanahashi's result on the best possibility of generalized Furuta inequality*, Math. Inequal. Appl. **2** (1999), 473–477.

[290] T. Yamazaki, *An extension of Specht's theorem via Kantorovich inequality and related results*, Math. Inequal. Appl. **3** (2000), 89–96.

[291] T. Yamazaki, *Further characterizations of chaotic order via Specht's ratio*, Math. Inequal. Appl. **3** (2000), 259–268.

[292] T. Yamazaki, *An extension of Kantorovich inequality to n-operators via the geometric mean by Ando-Li-Mathias*, Linear Algebra Appl. **416** (2006), 688–695.

[293] T. Yamazaki and M. Yanagida, *Characterizations of chaotic order associated with Kantorovich inequality*, Sci. Math. **2** (1999), 37–50.

[294] M. Yanagida, *Some applications of Tanahashi's result on the best possibility of Furuta inequality*, Math. Inequal. Appl. **2** (1999), 297–305.

[295] F. Zhang, *On the Bohr inequality of operator*, J. Math. Anal. Appl. **333** (2007), 1264–1271.

Index

Zagreb, March 2013