Linear Algebra and its Applications 433 (2010) 499–510

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

An asymmetric Kadison's inequality

Jean-Christophe Bourin *,[1], Éric Ricard[1]

Laboratoire de Mathématiques, Université de Franche-Comté, 16 route de Gray, 25030 Besançon, France

ARTICLE INFO

Article history:
Received 26 November 2009
Accepted 11 March 2010
Available online 8 April 2010

Submitted by T. Ando

AMS classification:
Primary: 15A60, 47A30, 47A60

Keywords:
Operator inequalities
Matrix geometric mean
Positive linear maps

ABSTRACT

Some inequalities for positive linear maps on matrix algebras are given, especially asymmetric extensions of Kadison's inequality and several operator versions of Chebyshev's inequality. We also discuss well-known results around the matrix geometric mean and connect it with complex interpolation.

0. Introduction

This note lies in the scope of matricial inequalities. The main motivation of this theory is to extend some classical inequalities for reals to self-adjoint matrices. Of course, the non-commutativity of $\mathbb{M}_n$ (the space of $n \times n$ complex matrices) enters into the game, making things much more complicated. The book [4] is a very good introduction to this subject. Many techniques have been developed, such as the theory of operator monotone/convex functions and their links with completely positive maps. Nevertheless the proofs very often rely on quite clever but simple arguments. As an illustration of the available tools, a very classical result is Kadison's inequality [13] saying that if $\Phi : \mathcal{A} \to \mathcal{M}$ is a unital positive (linear) map between C^*-algebras, then for a self-adjoint element A in $\mathcal{A}$,

$$\Phi(A)^2 \leqslant \Phi(A^2).$$

Taking $\Phi : \mathbb{M}_n \oplus \mathbb{M}_n \to \mathbb{M}_n$, $\Phi(A, B) = (A + B)/2$, this reflects the operator convexity of $t \mapsto t^2$. One can think of it as a kind of Jensen's or Cauchy–Schwarz's inequality. The main motivation of this

* Corresponding author.
E-mail addresses: jcbourin@univ-fcomte.fr (J.-C. Bourin), eric.ricard@univ-fcomte.fr (É. Ricard).

1 Research of both authors was supported by ANR 06-BLAN-0015.

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.03.015

paper is to try to get comparison relations between the images of the powers of A. At a first glance, one does not expect to have many positive results beyond operator convexity. But surprisingly, we notice here that $|\Phi(A^p)\Phi(A^q)| \leqslant \Phi(A^{p+q})$ provided that $0 \leqslant p \leqslant q$ and $A \geqslant 0$. This and some variations are our concern of the first section.

The second section deals more generally with monotone pairs, in place of pairs (A^p, A^q). These are pairs (A, B) of positive operators, characterized by joint relations $A = f(C)$ and $B = g(C)$ for some $C \geqslant 0$ in $\mathbb{M}_n$ and two non-decreasing, non-negative functions $f(t)$ and $g(t)$ on $[0, \infty)$. Comparing $\Phi(A)\Phi(B)$ with $\Phi(AB)$ is non-commutative versions of the classical Chebyshev's inequality,

$$\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right) \cdot \left(\frac{1}{n}\sum_{i=1}^{n} b_i\right) \leqslant \frac{1}{n}\sum_{i=1}^{n} a_i b_i$$

for non-negative increasing sequences $\{a_i\}$ and $\{b_i\}$ (here Φ is just a state).

In the last part, we point out the links between complex interpolation and power means. Along with some very classical approach and an idea of [11], this is used to furnish a simple proof of Furuta's inequality, which is the main tool in Section 1.

We assume that the reader is familiar with basic notions in operator and matricial inequalities theories. When possible, we state the results in their general context, that is, for von Neumann or C^*-algebras. But matrix inequalities for positive linear maps are essentially finite dimensional results, especially when it comes to unitary congruences. So, the reader may like to think of the algebras as $\mathbb{M}_n$.

1. Kadison's asymmetric type inequalities

In this section, we deal with positive linear maps $\Phi : \mathcal{A} \to \mathcal{M}$ between two unital C^*-algebra $\mathcal{A}$ and $\mathcal{M}$ with units denoted by I. In fact, we may assume that $\mathcal{A}$ is the unital C^*-algebra generated by a single positive operator A; hence, by a classical dilation theorem of Naimark (see [14, Theorem 3.10]), our maps Φ will be automatically completely positive. We will also always assume that these maps are unital, $\Phi(I) = I$, or more generally sub-unital, $\Phi(I) \leqslant I$.

Kadison's inequality is one of the most basic and useful results for such sub-unital maps; it states that for any $A \in \mathcal{A}^{sa}$ (the self-adjoint elements in $\mathcal{A}$),

$$\Phi(A)^2 \leqslant \Phi(A^2).$$

More generally, if f is operator convex on an interval containing 0 and $f(0) \leq 0$, then one has

$$f(\Phi(A)) \leqslant \Phi(f(A))$$

for all $A \in \mathcal{A}^{sa}$ with spectrum in the domain of f. If we drop the condition that 0 is in the domain of f and $f(0) \leq 0$, this Jensen's inequality remains true for unital maps. When Φ is the compression map to a subspace, it is then a basic characterization of operator convexity due to Davis [8]. The general case was noted in an influential paper of Choi [6]; nowadays everything is very clear using Stinespring's theorem (see [14, Theorem 4.1]) for completely positive maps.

First examples of operator convex/concave functions on $\mathbb{R}^+$ are given by powers, we refer to the corresponding Jensen inequalities as Choi's inequality; for $A \in \mathcal{A}^+$ (the positive cone of $\mathcal{A}$),

$$\Phi(A^p) \leqslant \Phi(A)^p, \quad 0 \leqslant p \leqslant 1,$$

and

$$\Phi(A)^p \leqslant \Phi(A^p), \quad 1 \leqslant p \leqslant 2.$$

In the spirit of operator convexity, one can naturally think of looking for more comparison relations between powers of A.

We start with an asymmetric extension of Kadison's inequality:

Theorem 1.1. *Let $A \in \mathcal{A}^+$ and, $0 \leqslant p \leqslant q$. Then,*

$$|\Phi(A^p)\Phi(A^q)| \leqslant \Phi(A^{p+q}).$$

Proof. We will derive this result from Furuta's inequality that we recall as follows:
Let $X \geqslant Y \geqslant 0$ in some $\mathbb{B}(\mathcal{H})$, let $\alpha \geqslant 1$ and $\beta \geqslant 0$. Then, for $\gamma \geqslant (\alpha+2\beta)/(1+2\beta)$,

$$X^{(\alpha+2\beta)/\gamma} \geqslant (X^{\beta}Y^{\alpha}X^{\beta})^{1/\gamma},$$

with equality if and only if $X = Y$.

We will discuss about it in Section 3. Now, set

$$X = \Phi(A^q)^{\frac{p}{q}}, \quad Y = \Phi(A^p).$$

By Choi's inequality, $X \geqslant Y$. Then we apply Furuta's inequality to X and Y with

$$\alpha = 2, \quad \beta = \frac{q}{p}, \quad \gamma = 2.$$

Note that

$$\gamma = 2 \geqslant \frac{2+2(q/p)}{1+2(q/p)} = \frac{\alpha+2\beta}{1+2\beta}$$

so that assumptions of Furuta's inequality are satisfied. Thus we obtain

$$\{\Phi(A^q)^{\frac{p}{q}}\}^{\frac{2+2(q/p)}{2}} \geqslant (\{\Phi(A^q)^{\frac{p}{q}}\}^{\frac{q}{p}}\{\Phi(A^p)\}^2\{\Phi(A^q)^{\frac{p}{q}}\}^{\frac{q}{p}})^{1/2},$$

equivalently

$$\Phi(A^q)^{1+p/q} \geqslant |\Phi(A^p)\Phi(A^q)|. \tag{1.1}$$

Since $1 \leqslant 1+p/q \leqslant 2$, using once again Choi's inequality for operator convex functions,

$$\Phi(A^{p+q}) \geqslant \Phi(A^q)^{1+p/q}. \tag{1.2}$$

Combining (1.1) and (1.2) completes the proof. □

Remark. Actually, we have shown the stronger inequality (1.1) that can be restated as follows: For $0 \leqslant \alpha \leqslant 1$, $\Phi(A)^{1+\alpha} \geqslant |\Phi(A^{\alpha})\Phi(A)|$.

Remark. For $0 < p \leqslant q$, the equality case in Theorem 1.1 entails the equality case in Choi's inequality, so that $\Phi(A^t) = \Phi(A)^t$ for all $t > 0$, in other words A is in the multiplicative domain of Φ.

Corollary 1.2. *Assume that moreover $\mathcal{M}$ is a von Neumann algebra, then for $A \in \mathcal{A}^+$ and $p, q \geqslant 0$, there is a partial isometry $V \in \mathcal{M}$ such that*

$$|\Phi(A^p)\Phi(A^q)| \leqslant V\Phi(A^{p+q})V^*.$$

If $\mathcal{M}$ is finite, then V can be chosen to be unitary.

This follows from Theorem 1.1 and the polar decomposition. Indeed, for any $Z \in \mathcal{M}$, there is a partial isometry V so that $Z = V|Z|$ and moreover $|Z^*| = V|Z|V^*$ and $|Z| = V^*|Z^*|V$. If $\mathcal{M}$ is finite, then V can also be chosen unitary. But in general, one cannot assume V to be unitary.

Remark. Fix $0 < q < p$. Let

$$A = \begin{bmatrix} 1+\varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 2+\varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 3+\varepsilon \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

and let Φ be the Schur product with B. Then for ε small enough, it follows from tedious computations on derivatives, that we cannot get rid of V in Corollary 1.2 like in Theorem 1.1.

From now on, we come back to the setting of matrix inequalities where $\mathcal{M} = \mathbb{M}_n$ for some positive integer n.

The next two results are variations of Corollary 1.2. We rely on an easy consequence of the min-max principle; if $A \geqslant B \geqslant 0$ in $\mathbb{M}_n$ and $f(t)$ is non-decreasing on $[0, \infty[$, then $f(A) \geqslant Vf(B)V^*$ for some unitary $V \in \mathbb{M}_n$.

Proposition 1.3. *Let $A \in \mathcal{A}^+$ and $p, q, r \geqslant 0$ such that $\min\{p, r\} \leqslant q/2$ and $\max\{p, r\} \leqslant q$. Then, for some unitary $V \in \mathbb{M}_n$,*

$$|\Phi(A^p)\Phi(A^q)\Phi(A^r)| \leqslant V\Phi(A^{p+q+r})V^*.$$

Proof. We may assume $q = 1$ and $r \leqslant 1/2$. We then have

$$\Phi(A^r)^2 \leqslant \Phi(A^{2r}) \leqslant \Phi(A)^{2r}$$

so that $\Phi(A^r) = \Phi(A)^r K$ for a contraction K. Hence, for some unitary U,

$$|\Phi(A^p)\Phi(A)\Phi(A^r)| \leqslant U|\Phi(A^p)\Phi(A)^{1+r}|U^*$$

so

$$|\Phi(A^p)\Phi(A)\Phi(A^r)| \leqslant U\left(\{\Phi(A)\}^{1+r}\{\Phi(A^p)\}^2\{\Phi(A)\}^{1+r}\right)^{\frac{1}{2}} U^*. \tag{1.3}$$

Now, observe that a byproduct of Furuta's inequality is:

If $X \geqslant Y \geqslant 0$ and, $\alpha, \ \beta \geqslant 0$, then for some unitary W,

$$X^{\alpha+2\beta} \geqslant W(X^\beta Y^\alpha X^\beta)W^*.$$

Applying this inequality to $X = \Phi(A)^p$ and $Y = \Phi(A^p)$ with $\alpha = 2, \ \beta = (1 + r)/p$ and combining with (1.3) yields

$$|\Phi(A^p)\Phi(A)\Phi(A^r)| \leqslant V_0(\Phi(A)^{1+p+r})V_0^*$$

for some unitary V_0. Since, by a byproduct of Choi's inequality, we also have some unitary V_1 such that

$$\Phi(A)^{1+p+r} \leqslant V_1\Phi(A^{1+p+r})V_1^*,$$

we get the conclusion. $\square$

At the cost of one more unitary congruence, assumptions of Proposition 1.3 can be relaxed. We will use an inequality of Bhatia and Kittaneh (see [4] for an elementary proof): For all $A, \ B$ in some finite von Neumann algebra $\mathcal{M}$, there is some unitary $U \in \mathcal{M}$ such that

$$|AB^*| \leqslant U\frac{|A|^2 + |B|^2}{2}U^*.$$

Proposition 1.4. *Let $A \geqslant 0$ in $\mathcal{A}$ and let $p, q, r \geqslant 0$ with $q \geqslant p, r$. Then, for some unitaries $U, \ V$ in $\mathbb{M}_n$,*

$$|\Phi(A^p)\Phi(A^q)\Phi(A^r)| \leqslant \frac{U\Phi(A^{p+q+r})U^* + V\Phi(A^{p+q+r})V^*}{2}.$$

Proof. We may assume $q = 1$. Let $\alpha \in [0, 1]$ and note that by Bhatia–Kittaneh's inequality,

$$\begin{aligned} |\Phi(A^p)\Phi(A)\Phi(A^r)| &= |\Phi(A^p)\Phi(A)^\alpha \cdot \Phi(A)^{1-\alpha}\Phi(A^r)| \\ &\leqslant W\frac{|\Phi(A^p)\Phi(A)^\alpha|^2 + |\Phi(A^r)\Phi(A)^{1-\alpha}|^2}{2}W^* \end{aligned} \tag{1.4}$$

for some unitary W. Then set $\alpha = (r - p + 1)/2$ (hence $0 \leqslant \alpha \leqslant 1$). We may estimate each summand in (1.4) via Furuta's inequality, since $\Phi(A^p) \leqslant \Phi(A)^p$ and $\Phi(A^r) \leqslant \Phi(A)^r$. For the first summand, there are some unitaries W_0 and W_1 such that

$$\begin{aligned} |\Phi(A^p)\Phi(A)^{\alpha}|^2 &= \{\Phi(A)^p\}^{\frac{r-p+1}{2p}}\{\Phi(A^p)\}^2\{\Phi(A)^p\}^{\frac{r-p+1}{2p}} \\ &\leqslant W_0\Phi(A)^{1+p+r}W_0^* \\ &\leqslant W_1\Phi(A^{1+p+r})W_1^*, \end{aligned} \tag{1.5}$$

where the last step follows from Choi's inequality. We also have a unitary W_2 such that

$$|\Phi(A^r)\Phi(A)^{1-\alpha}|^2 \leqslant W_2\Phi(A^{1+p+r})W_2^* \tag{1.6}$$

and combining (1.4)–(1.6) completes the proof. □

2. Matrix monotony inequalities

Here we try to understand the results of the first section, using the more general notion of a monotone pair. Recall that (A, B) is said to be a monotone pair in $\mathbb{M}_n$ if there exist a positive element $C \in \mathbb{M}_n$ and two non-negative, non-decreasing functions f and g so that $A = f(C)$ and $B = g(C)$. A typical example is (A^p, A^q) for $A \geqslant 0$ and $p, q \geqslant 0$.

For technical reasons, we have to stick to $\mathbb{M}_n$, as many arguments rely on the min–max principle. For instance, we use the following result of [5] which compares the singular values of AEB and ABE for some projections E.

Theorem 2.1. *Let (A, B) be a monotone pair and let E be a self-adjoint projection. Then, for some unitary V,*

$$|AEB| \leqslant V|ABE|V^*.$$

As consequences, we have the following Chebyshev's type eigenvalue inequalities for compressions [5],

$$\lambda_j[(EAE)(EBE)] \leqslant \lambda_j[EABE]$$

and

$$\lambda_j[(EAE)(EBE)(EAE)] \leqslant \lambda_j[EABAE], \tag{2.1}$$

where $\lambda_j[\cdot]$ stands for the list of eigenvalues arranged in decreasing order with their multiplicities. Let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a unital completely positive (linear) map. It is well known (Stinespring) that Φ can be decomposed as $\Phi(A) = E\pi(A)E$, where $\pi : \mathbb{M}_n \to \mathbb{M}_m$ is a $*$-representation (with $m \leqslant n^2 d$) and $E \in \mathbb{M}_m$ is a rank d projection (and identifying $E\mathbb{M}_m E$ with $\mathbb{M}_d$). Taking into account that we start from a commutative C^*-algebra, (2.1) is then equivalent to:

Corollary 2.2. *Let (A, B) be a monotone pair in $\mathbb{M}_n$ and let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a unital positive map. Then, for some unitary $V \in \mathbb{M}_d$,*

$$\Phi(A)\Phi(B)\Phi(A) \leqslant V\Phi(ABA)V^*.$$

In the case of pairs of positive powers (A^p, A^q), such results are easy consequences of Furuta's inequality. To apply Corollary 2.2 we define a special class of monotone pairs (of positive operators).

Definition. A monotone pair (A, B) is concave if $A = h(B)$ for some concave function $h : [0, \infty) \to [0, \infty)$.

This class contains pairs of powers (A^p, A^q) with $0 \leqslant p \leqslant q$ and we note that Corollaries 2.4 and 2.5 below are variations of Theorem 1.1. We first state a factorization result.

Theorem 2.3. *Let (A, B) be a concave monotone pair in $\mathbb{M}_n$ and let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a unital positive map. Then, for some contraction K and unitary U in $\mathbb{M}_d$,*

$$\Phi(B)\Phi(A) = \sqrt{\Phi(AB)}K\sqrt{\Phi(AB)}U.$$

Proof. By a continuity argument we may assume that A is invertible, hence

$$\begin{pmatrix} AB & B \\ B & BA^{-1} \end{pmatrix} \geqslant 0.$$

Replacing Φ by $\Phi \circ \mathbb{E}$, where $\mathbb{E}$ is the conditional expectation onto the C^*-algebra generated by A and B, we can assume that Φ is completely positive so that we get

$$\begin{pmatrix} \Phi(AB) & \Phi(B) \\ \Phi(B) & \Phi(BA^{-1}) \end{pmatrix} \geqslant 0,$$

equivalently,

$$\begin{pmatrix} \Phi(AB) & \Phi(B)\Phi(A) \\ \Phi(A)\Phi(B) & \Phi(A)\Phi(BA^{-1})\Phi(A) \end{pmatrix} \geqslant 0.$$

The concavity assumption on (A, B) implies that (A, BA^{-1}) is a monotone pair, indeed both $h(t)$ and $t/h(t)$ are non-decreasing. By Corollary 2.2, we then have a unitary U such that

$$\begin{pmatrix} \Phi(AB) & \Phi(B)\Phi(A) \\ \Phi(A)\Phi(B) & U^*\Phi(AB)U \end{pmatrix} \geqslant 0, \tag{2.2}$$

equivalently,

$$\Phi(B)\Phi(A) = \sqrt{\Phi(AB)}LU^*\sqrt{\Phi(AB)}U$$

for some contraction L. $\square$

Theorem 2.3 is equivalent to positivity of the block-matrix (2.2). Considering the polar decomposition $\Phi(A)\Phi(B) = W|\Phi(A)\Phi(B)|$ we infer

$$\begin{pmatrix} I & -W^* \end{pmatrix} \begin{pmatrix} \Phi(AB) & \Phi(B)\Phi(A) \\ \Phi(A)\Phi(B) & U^*\Phi(AB)U \end{pmatrix} \begin{pmatrix} I \\ -W \end{pmatrix} \geqslant 0$$

and thus obtain:

Corollary 2.4. *Let (A, B) be a concave monotone pair in $\mathbb{M}_n$ and let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a unital positive map. Then, for some unitary $V \in \mathbb{M}_d$,*

$$|\Phi(A)\Phi(B)| \leqslant \frac{\Phi(AB) + V\Phi(AB)V^*}{2}.$$

Recall that a norm is said symmetric whenever $\|UAV\| = \|A\|$ for all A and all unitaries U, V. Corollary 2.4 yields for *concave* monotone pairs some Chebyshev's type inequalities for symmetric norms,

$$\|\Phi(A)\Phi(B)\| \leqslant \|\Phi(AB)\|.$$

It is not clear that this can be extended to all monotone pairs. In fact, for concave monotone pairs, Theorem 2.3 entails a stronger statement. Given $X, Y \geqslant 0$, recall that the weak log-majorization relation $X \prec_{\rm wlog} Y$ means

$$\prod_{j \leqslant k} \lambda_j[X] \leqslant \prod_{j \leqslant k} \lambda_j[Y]$$

for all $k = 1, 2, \ldots$ This entails $\|X\| \leqslant \|Y\|$ for all symmetric norms. Theorem 2.3 and Horn's inequality yield:

Corollary 2.5. *Let (A, B) be a concave monotone pair in $\mathbb{M}_n$ and let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a unital, positive linear map. Then,*

$$|\Phi(A)\Phi(B)| \prec_{\text{wlog}} \Phi(AB).$$

In case of pairs (A^p, A^q) we have more:

Proposition 2.6. *Let $A \geqslant 0$ in $\mathbb{M}_n$, let $p, q \geqslant 0$ and let Φ as above. Then, for all eigenvalues,*

$$\lambda_j[\Phi(A^p)]\,\lambda_j[\Phi(A^q)] \leqslant \lambda_j[\Phi(A^{p+q})].$$

Proof. We outline an elementary proof. It suffices to show that for a given projection E,

$$\lambda_j[EA^pE]\lambda_j[EA^qE] \leqslant \lambda_j[EA^{p+q}E]. \tag{2.3}$$

In case of the first eigenvalue, this can be written via the operator norm $\|\cdot\|_\infty$ as

$$\|EA^pE\|_\infty\|EA^qE\|_\infty \leqslant \|EA^{p+q}E\|_\infty. \tag{2.4}$$

The proof of (2.4) follows from Young's and Jensen's inequalities (always true for the operator norm),

$$\begin{aligned}\|EA^pE\|_\infty\|EA^qE\|_\infty &\leqslant \frac{p}{p+q}\|EA^pE\|_\infty^{\frac{p+q}{p}} + \frac{q}{p+q}\|EA^qE\|_\infty^{\frac{p+q}{q}}\\ &\leqslant \|EA^{p+q}E\|_\infty.\end{aligned}$$

The min–max characterization of eigenvalues combined with (2.4) implies the proposition; indeed simply take Q a projection commuting with E of corank $j-1$ so that $\|QEA^{p+q}EQ\|_\infty = \lambda_j[EA^{p+q}E]$ and apply (2.4) with QE instead of E. $\square$

Results of this section follow from (2.1), equivalently from Corollary 2.2, and hence have been stated for unital positive maps. In fact these results can be stated to all sub-unital positive maps. In particular the key Corollary 2.2 becomes:

Corollary 2.2a. *Let (A, B) be a monotone pair in $\mathbb{M}_n$ and let $\Phi : \mathbb{M}_n \to \mathbb{M}_d$ be a sub-unital positive map. Then, for some unitary $V \in \mathbb{M}_d$,*

$$\Phi(A)\Phi(B)\Phi(A) \leqslant V\Phi(ABA)V^*.$$

Proof. Let $\mathcal{A}$ be the unital $*$-algebra generated by A and B. Restricting Φ to $\mathcal{A}$, it follows from Stinespring's theorem (or from Naimark's theorem) that Φ can be decomposed as $\Phi(A) = Z\pi(A)Z$, where $\pi : \mathcal{A} \to \mathbb{M}_m$ is a $*$-representation (with $m \leqslant nd$) and $Z \in \mathbb{M}_m$ is a positive contraction (and identifying $E\mathbb{M}_mE$ with $\mathbb{M}_d$ for some projection $E \geqslant Z$).

Since $(\pi(A), \pi(B))$ is monotone, it then suffices to prove the result for congruence maps of $\mathbb{M}_n$ of type $\Phi(X) = ZXZ$ where Z is a positive contraction. We may then derive the result from (2.1) and a two-by-two trick: Note that

$$A_0 = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B_0 = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$$

form a monotone pair. Note also that

$$E = \begin{pmatrix} Z & (Z(I-Z))^{1/2} \\ (Z(I-Z))^{1/2} & I-Z \end{pmatrix}$$

is a projection. By (2.1),

$$\lambda_j[(EA_0E)(EB_0E)(EA_0E)] \leqslant \lambda_j[EA_0B_0A_0E],$$

equivalently,

$$\lambda_j[|EA_0EB_0^{1/2}|^2] \leqslant \lambda_j[(A_0B_0A_0)^{1/2}E(A_0B_0A_0)^{1/2}]. \tag{2.5}$$

Observe that

$$|EA_0EB_0^{1/2}|^2 = \begin{pmatrix} B^{1/2}ZAZAZB^{1/2} & 0 \\ 0 & 0 \end{pmatrix} \simeq \begin{pmatrix} Z^{1/2}AZBZAZ^{1/2} & 0 \\ 0 & 0 \end{pmatrix} \tag{2.6}$$

where $\simeq$ means unitary equivalence, and similarly,

$$(A_0B_0A_0)^{1/2}E(A_0B_0A_0)^{1/2} \simeq \begin{pmatrix} Z^{1/2}ABAZ^{1/2} & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.7}$$

Combining (2.6) and (2.7) with (2.5) and replacing $Z^{1/2}$ by Z yields

$$ZAZ \cdot ZBZ \cdot ZAZ \leqslant V(ZABAZ)V^*$$

for some unitary V. $\square$

We end Section 2 with a remark about the two-by-two trick used to derive Corollary 2.4. This can be used to get some triangle type matrix inequalities. For instance, given two operators A and B in some von Neumann algebra $\mathcal{M}$, there exists a partial isometry V such that:

$$|A+B| \leqslant \frac{|A| + |B| + V^*(|A^*| + |B^*|)V}{2}. \tag{2.8}$$

To check it, note that, since for all X,

$$\begin{pmatrix} |X^*| & X \\ X^* & |X| \end{pmatrix} \geqslant 0,$$

we thus have for all V,

$$(-V^* I)\begin{pmatrix} |A^*| + |B^*| & A+B \\ A^* + B^* & |A| + |B| \end{pmatrix}\begin{pmatrix} -V \\ I \end{pmatrix} \geqslant 0,$$

and taking V the partial isometry in the polar decomposition of $A+B$ yields (2.8). This can be used to give a very short proof of the triangle inequality for the trace norm in semi-finite von Neumann algebras.

3. Means and order preserving relations

Furuta's inequality was used as key tool in the first section; here we present a possible proof for completeness. For that purpose we use the geometric mean of positive definite matrices and Ando–Hiai's inequality. We do not pretend to originality and we closely follow an approach due to Ando, Hiai, Fujii and Kamei. However, we point out an interesting observation connecting the geometric mean to complex interpolation. In fact this observation is rather old: Identifying positive operators with quadratic forms, it is worth noting that Donoghue's construction with complex interpolation [9] seems to be the first appearance of the matrix geometric mean.

In the whole section we consider $\mathbb{B}(\mathcal{H})$, the set of all bounded operators on a Hilbert space $\mathcal{H}$, and its positive invertible part, $\mathbb{B}^+$.

For details and some important results around the geometric mean we refer to [1,2], references herein, and [4] for a nice survey of other features of the weighted geometric means, especially as geodesics on the cone of positive operators.

3.1. Means and interpolation

Let $\alpha \in [0,1]$ and consider a map

$$\mathbb{B}^+ \times \mathbb{B}^+ \to \mathbb{B}^+$$
$$(A,B) \mapsto A\sharp_\alpha B$$

satisfying the two natural requirements for an α-geometrical mean

1. If $AB = BA$ then $A\sharp_\alpha B = A^{1-\alpha}B^\alpha$.
2. $(X^*AX)\sharp_\alpha(X^*BX) = X^*(A\sharp_\alpha B)X$ for any invertible X.

Choosing the appropriate X, we necessarily have

$$A\sharp_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}.$$

So there is a unique extension of the α-geometrical mean for commuting operators which is invariant under congruence, that is called the α-geometrical mean.

Matrix geometric means have their roots in the work of Pusz and Woronowicz [15] about functional calculus for sesquilinear forms. Their construction is closely related to complex interpolation. Coming back to means, these links are even clearer.

We briefly recall the complex interpolation method of Calderon, see [3] for a complete exposition.

Two Banach spaces A_0 and A_1 are said to be an interpolation couple if there is another Banach space V and continuous embeddings $A_i \to V$. So we have a way to identify elements and it makes sense to speak of $A_0 \cap A_1$ and $A_0 + A_1$ (which are also Banach spaces with the usual norms). The idea of interpolation is to assign for each $\alpha \in [0, 1]$ a space that is intermediate between the A_i. The construction is a bit technical.

Let $\Delta = \{z \in \mathbb{C} | 0 < \mathrm{Re}\, z < 1\}$, $\delta_i = \{z \in \mathbb{C} | \mathrm{Re}\, z = i\}$ for $i = 0, 1$. Define $\mathcal{F}(A_0, A_1)$ as the set of maps $f : \overline{\Delta} \to A_0 + A_1$, such that

(i) f is analytic in Δ.
(ii) for $i = 0, 1, f(\delta_i) \subset A_i$ and $f : \delta_i \to A_i$ is bounded and continuous.
(iii) for $i = 0, 1, \lim_{t\in\mathbb{R}\to\pm\infty} \|f(i + it)\|_{A_i} = 0$.

Equipped with the norm

$$\|f\| = \max_{i=0,1}\{\sup_{z\in\delta_i} \|f(z)\|_{A_i}\}$$

$\mathcal{F}(A_0, A_1)$ becomes a Banach space. Finally for $\alpha \in [0, 1]$,

$$(A_0, A_1)_\alpha = \{x \in A_0 + A_1 : \exists f \in \mathcal{F}(A_0, A_1) \text{ so that } f(\alpha) = x\}$$

with the quotient norm

$$\|x\|_{(A_0,A_1)_\alpha} = \inf\{\|f\| : f(\alpha) = x\}.$$

This functor has many nice properties. The most common is the interpolation principle; consider two interpolation couples (A_0, A_1) and (B_0, B_1) and bounded maps $T_i : A_i \to B_i$ so that T_1 and T_2 coincide on $A_0 \cap A_1$, then one can define a map $T_\alpha : (A_0, A_1)_\alpha \to (B_0, B_1)_\alpha$ which extends T_i on $A_0 \cap A_1$, and moreover one has $\|T_\alpha\| \leqslant \|T_0\|^{1-\alpha}\|T_1\|^\alpha$.

There are concrete examples where these interpolated norms can be computed. If $A_0 = L_\infty([0, 1])$ and $A_1 = L_1([0, 1])$, one has $(A_0, A_1)_\alpha = L_{1/\alpha}([0, 1])$.

Using basic properties of the interpolation, it is easy to see that the interpolation of two compatible Hilbert spaces is still a Hilbert space. Indeed, by [3] Theorem 5.1.2, the parallelogram identity is preserved by the complex interpolation method.

Let $A \in \mathbb{B}^+$, then it defines an equivalent hilbertian norm on $\mathcal{H}$ by $\|h\|_A = \|A^{1/2}h\|_\mathcal{H}$. And conversely any equivalent hilbertian norm on $\mathcal{H}$ arises from some $A \in \mathbb{B}^+$. We denote by $\mathcal{H}_A$ the Hilbert space coming from A.

Now take $A_i \in \mathbb{B}^+$, $(\mathcal{H}_{A_0}, \mathcal{H}_{A_1})$ forms an interpolation couple of Hilbert space (with the obvious identification). The resulting interpolated space for $\alpha \in [0, 1]$ will also give an equivalent norm on $\mathcal{H}$, associated to an operator that we call A_α. Let us have a look at the properties of $(A_0, A_1) \mapsto I_\alpha(A_0, A_1) = A_\alpha$.

First, it is an easy exercise to check that if A_0 and A_1 commute then $I_\alpha(A_0, A_1) = A_\alpha = A_0^{1-\alpha}A_1^\alpha$.

Secondly, let $X \in \mathbb{B}(\mathcal{H})$ be invertible. With $B_i = X^*A_iX$, it is clear that $X : \mathcal{H}_{B_i} \to \mathcal{H}_{A_i}$ is a unitary for $i = 0, 1$. From the interpolation principle, X will also be unitary for the interpolated norms. Coming back to operators, this says that $I_\alpha(X^*A_0X, X^*A_1X) = X^*I_\alpha(A_0, A_1)X$.

So we can conclude that the α-geometric mean is the interpolation functor of index α. With this in mind, all properties of the means come from basic results in the complex interpolation theory.

Take (A_0, A_1) and (B_0, B_1) in $(\mathbb{B}^+)^2$ and assume that $B_i \leqslant A_i$. This means that the identity of $\mathcal{H}$ is a contraction from $\mathcal{H}_{A_i}$ to $\mathcal{H}_{B_i}$. By the interpolation principle, the same holds for the interpolated norms. So we can conclude that the α-mean is monotone. Note that this gives another proof of the monotony of $A \mapsto A^\alpha$ for $0 \leqslant \alpha \leqslant 1$.

To get concavity of the mean is also easy for people familiar with interpolation. Take A_i and B_i in $\mathbb{B}^+$, and $0 < \lambda < 1$ and notice that the map $\mathcal{H}_{\lambda A_i + (1-\lambda)B_i} \to \mathcal{H}_{\lambda A_i} \oplus_2 \mathcal{H}_{(1-\lambda)B_i}$, $h \mapsto (h, h)$ is an isometry. From properties of the interpolation functor, we deduce that the same map $\mathcal{H}_{(\lambda A_0 + (1-\lambda)B_0)\sharp_\alpha(\lambda A_1 + (1-\lambda)B_1)} \to \mathcal{H}_{\lambda(A_0\sharp_\alpha A_1)} \oplus_2 \mathcal{H}_{(1-\lambda)(B_0\sharp_\alpha B_1)}$ is a contraction. Coming back to an inequality on operators gives the concavity. This illustrates the well-known fact that taking subspaces and interpolation do not commute.

Another useful result is the reiteration theorem: For any $\alpha, \beta, \gamma \in [0, 1]$, we have (provided that $A_0 \cap A_1$ is dense in both A_0 and A_1)

$$((A_0, A_1)_\alpha, (A_0, A_1)_\beta)_\gamma = (A_0, A_1)_{(1-\gamma)\alpha+\gamma\beta}.$$

This means that for any $x, y, z \in [0, 1]$ and $A, B \in \mathbb{B}^+$,

$$(A\sharp_x B)\sharp_z(A\sharp_y B) = A\sharp_{x(1-z)+yz}B.$$

Of course this can also be checked directly from the formulae defining $\sharp$.

The next theorem is the Ando–Hiai inequality. We only use the language of operator mean, but this is really a proof in the spirit of the interpolation theory.

Theorem 3.1. *Let $A, B \in \mathbb{B}^+$ and $0 < s < 1$. Then,*

$$\|(A\sharp_\alpha B)^s\|_\infty \leqslant \|A^s\sharp_\alpha B^s\|_\infty.$$

Proof. By homogeneity we may assume $\|A\sharp_\alpha B\|_\infty = 1$. Hence we have $A\sharp_\alpha B \leqslant I$. By using monotony of geometric means and the reiteration principle we then get

$$\begin{aligned} A^s\sharp_\alpha B^s &= (I\sharp_s A)\sharp_\alpha(I\sharp_s B) \\ &\geqslant ((A\sharp_\alpha B)\sharp_s A)\sharp_\alpha((A\sharp_\alpha B)\sharp_s B) \\ &= ((A\sharp_\alpha B)\sharp_s(A\sharp_0 B))\sharp_\alpha((A\sharp_\alpha B)\sharp_s(A\sharp_1 B)) \\ &= (A\sharp_{\alpha(1-s)}B)\sharp_\alpha(A\sharp_{\alpha(1-s)+s}B) = A\sharp_\alpha B. \end{aligned}$$

Thus $\|A^s\sharp_\alpha B^s\|_\infty \geqslant 1$ and this proves the theorem. $\square$

Remark. A theory of complex interpolation for families of Banach spaces has been developed in [7]. The family may be indexed by the unit circle in $\mathbb{C}$, say $A(z)$, with some technical assumptions. The interpolation then provides a family of spaces $A(z)$ for $|z| < 1$. This can be used define a mean of several operators. For instance, in the case of n operators, one may pick a partition of the unit circle in n sets E_i with Lebesgue measure α_i, and choose the family $A(z) = \mathcal{H}_{A_i}$ if $z \in E_i$. Then the interpolated space at 0 is of the form $A(0) = \mathcal{H}_A$, and one may think of A as a (α_i)-mean of the A_i's. Unfortunately this definition depends on the choice of the E_i (unless $n = 2$). This kind of approach for interpolation of a finite family of spaces can also be found in [10].

3.2. From means to order relations

Next we explain how to go from Ando–Hiai's inequality to Furuta's theorem (their equivalence was pointed out in [11]).

Let $A, B \in \mathbb{B}^+$ with $A \geqslant B$. Then, $A^{-1}\sharp_{1/2}B \leqslant I$, so by Ando–Hiai's inequality, $A^{-p}\sharp_{1/2}B^p \leqslant I$ whenever $p \geqslant 1$. Equivalently we have an order preserving relation for $f(t) = t^p$ with $p \geqslant 1$,

$$A^p \geqslant (A^{p/2}B^pA^{p/2})^{1/2}, \quad p \geqslant 1.$$

Such inequalities suggest to look for the best exponents $p,\ r,\ w$ for which

$$A \geqslant B \geqslant 0 \Rightarrow A^{(p+r)w} \geqslant (A^{r/2}B^pA^{r/2})^w \tag{3.1}$$

and consequently to get interesting substitutes to the lack of operator monotony of $f(t) = t^p, p \geqslant 1$.

To do so, it seems natural to find relations for weighted geometric means of the form

$$A \geqslant B \geqslant 0 \Rightarrow A^{-r}\sharp_\alpha B^p \leqslant I. \tag{3.2}$$

Because of homogeneity, this can hold only for $\alpha = \frac{r}{p+r}$. If $p \leqslant 1$, this inequality is obvious by the monotony of the mean. For $p > 1$, as above thanks to Ando–Hiai's inequality, one only need to find $s \leqslant 1$ so that $A^{-sr}\sharp_\alpha B^{sp} \leqslant I$; we have just said that $s = 1/p$ works. We have proved:

Lemma 3.2. *Let $A, B \in \mathbb{B}^+$ with $A \geqslant B$ and $p, r > 0$. Then,*

$$A^{-r}\sharp_{\frac{r}{p+r}} B^p \leqslant I.$$

Taking another mean with B^p, we obtain the optimal form of (3.2):

Lemma 3.3. *Let $A, B \in \mathbb{B}^+$ with $A \geqslant B$ and $r > 0, p \geqslant 1$. Then,*

$$A^{-r}\sharp_{\frac{1+r}{p+r}} B^p \leqslant B \leqslant A.$$

Hence we have recaptured quite easily two lemmas due to Fujii and Kamei [11].

We come back to relations of the form (3.1), the last lemma says:

$$A \geqslant B \geqslant 0 \Rightarrow A^{1+r} \geqslant \left(A^{\frac{r}{2}}B^pA^{\frac{r}{2}}\right)^{\frac{1+r}{p+r}}, \quad r > 0,\ p \geqslant 1. \tag{3.3}$$

Equivalently,

$$A^{(p+r)w} \geqslant (A^{\frac{r}{2}}B^pA^{\frac{r}{2}})^w, \quad r > 0,\ p \geqslant 1,$$

where $w = \frac{1+r}{p+r}$. This is still valid for $w \leqslant \frac{1+r}{p+r}$ by the operator monotony of $t \mapsto t^\alpha$, $0 \leqslant \alpha \leqslant 1$. We obtain Furuta's theorem:

Theorem 3.4. *Let $A, B \geqslant 0$ in $\mathbb{B}(\mathcal{H})$ and $r \geqslant 0, p \geqslant 1$. If $q \geqslant (p + r)/(1 + r)$, then*

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

The general statement follows from the case $B \in \mathbb{B}^+$ by continuity.

3.2.1. Comments

Around 1985 it was conjectured by Kwong that $A \geqslant B \geqslant 0$ entails $A^2 \geqslant (AB^2A)^{1/2}$, equivalently $A^2 \geqslant |BA|$. In 1987, Furuta [12] proved his inequality. Some numerical experiments lead him to know the condition on the exponents and he obtained a direct, quite ingenious proof. However the natural conjecture of Kwong may be written via geometric means and is nicely answered by a basic case of Ando–Hiai's inequality (1994). Hence order preserving relations may be obtained from a study of weighted geometric means. We have followed this idea, mainly developed by Ando, Hiai, Fujii and Kamei.

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl. 26 (1979) 203–241.

[2] T. Ando, F. Hiai, Log majorization and complementary Golden–Thompson type inequality, Linear Algebra Appl. 197 (1994) 113–131.

[3] J. Bergh, J. Löfström, Interpolation Spaces. An Introduction, Grundlehren der Mathematischen Wissenschaften, vol. 223, Springer-Verlag, Berlin/New York, 1976.
[4] R. Bhatia, Positive Definite Matrices, Princeton University Press, Princeton, NJ, 2007.
[5] J.-C. Bourin, Singular values of compressions, restrictions and dilations, Linear Algebra Appl. 360 (2003) 259–272.
[6] M.-D. Choi, A Schwarz inequality for positive linear maps on C^*-algebras, Illinois J. Math. 18 (1974) 565–574.
[7] R.R. Coifman, M. Cwikel, R. Rochberg, Y. Sagher, G. Weiss, A theory of complex interpolation for families of Banach spaces, Adv. Math. 43 (3) (1982) 203–229.
[8] C. Davis, A Schwarz inequality for convex operator functions, Proc. Amer. Math. Soc. 8 (1957) 42–44.
[9] W.F. Donoghue, The interpolation of quadratic norms, Acta. Math. 118 (1967) 251–270.
[10] A. Favini, Su una estensione del metodo d'interpolazione complesso, Rend. Sem. Mat. Univ. Padova 47 (1972) 243–298.
[11] M. Fujii, E. Kamei, Ando–Hiai inequality and Furuta inequality, Linear Algebra Appl. 416 (2006) 541–545.
[12] T. Furuta, $A \geqslant B \geqslant 0$ assures $(B^rA^pB^r)^{1/q} \geqslant B^{(p+2r)/q}$ for $r \geqslant 0$, $p \geqslant 0$, $q \geqslant 1$ with $(1+2r)q \geqslant p+2r$, Proc. Amer. Math. Soc. 101 (1987) 85–88.
[13] R. Kadison, A generalized Schwarz inequality and algebraic invariants for operator algebras, Ann. Math. 56 (2) (1952) 494–503.
[14] V.I. Paulsen, Completely Bounded Maps and Dilations, Pitman Research Notes in Mathematics Series, vol. 146, Longman Scientific & Technical/John Wiley & Sons, Inc., Harlow/New York, 1986.
[15] W. Pusz, S.L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, Rep. Math. Phys. 8 (2) (1975) 159–170.