Introduction

We will briefly present the main results of the five chapters of this work. The chapters can be read independently. Throughout the presentation of Chapter 1 to 4, by operator we mean a linear operator on a **finite** dimensional, real or complex Hilbert space $\mathcal{H}$. Chapter 5 is specific to infinite dimensional spaces.

Given an operator A on $\mathcal{H}$ and a (ortho)projection E onto a subspace $\mathcal{E}$, we recall that the compression of A onto $\mathcal{E}$, denoted by $A_{\mathcal{E}}$, is the restriction of EAE to $\mathcal{E}$.

Chapter 1 ([6])

Recall that a continuous function $f : (a, b) \longrightarrow \mathrm{R}$ is operator convex if

$$f((A + B)/2) \leq (f(A) + f(B))/2$$

for every Hermitians A, B with spectra in (a, b). The simplest nontrivial examples of operator convex functions are $t \longrightarrow t^2$ on the whole real line and $t \longrightarrow t^{-1}$ on the positive half-line. It is obvious that $(A_{\mathcal{E}})^2 \leq (A^2)_{\mathcal{E}}$ for every Hermitian operator A and every subspace $\mathcal{E}$. Moreover, if $A \geq 0$, a basic result in Matrix Theory states that

$$(A_{\mathcal{E}})^{-1} \leq (A^{-1})_{\mathcal{E}}. \tag{B}$$

More generally, a remarkable fact, due to C. Davis, states that a function f on (a, b) is operator convex if and only if for every subspace $\mathcal{E}$ and every Hermitian operator A whose spectrum lies in (a, b) one has

$$f(A_{\mathcal{E}}) \leq f(A)_{\mathcal{E}}. \tag{D}$$

F. Hansen and G.K. Pedersen showed that davis' characterization is equivalent to the noncommutative Jensen's Inequality

$$f(\sum_i Z_i^* A_i Z_i) \leq \sum_i Z_i^* f(A_i) Z_i. \tag{J}$$

for all Hermitians $\{A_i\}_{i=1}^m$ with spectra in $[a, b]$ and all isometric column $\{Z_i\}_{i=1}^m$ Here, isometric column means that $\sum_{i=1}^m Z_i^* Z_i = I$. If $0 \in [a, b]$, $f(0) \leq 0$, and A

is a Hermitian with spectrum in $[a,b)]$, then (J) entails its contractive version

$$f(Z^*AZ) \leq Z^*f(A)Z. \tag{C}$$

for all contractions Z. In fact, Hansen-Pedersen first showed (C).

What can be said about convex, not operator convex functions ? An immediate application of Jensen's inequality shows that

$$f(\langle h, Ah\rangle) \leq \langle h, f(A)h\rangle$$

for all norm vectors h. One may deduce Berezin's inequality

$$\mathrm{Tr}\, f(A_{\mathcal{E}}) \leq \mathrm{Tr}\, f(A)_{\mathcal{E}},$$

that is Davis' inequality remains valid inside the trace. Similarly (J) remains valid inside the trace (Hansen-Pedersen) as well as (C) (Brown-Kosaki). We will prove that, under simple additional assumptions, these trace inequalities are strenghtened as eigenvalues inequalities. Let g be operator convex on $[a,b]$ and let ϕ be a nondecreasing, convex function on $g([a,b])$. Then, $f = \phi \circ g$ is convex and we say that f is *unitary convex* on $[a,b]$. Since $t \longrightarrow -t$ is trivially operator convex, we note that *the class of unitary convex functions contains the class of monotone convex functions.* The following result holds:

Theorem 1.1. *Let f be a monotone, or more generally unitary, convex function on $[a,b]$, let A be a Hermitian whose spectrum lies in $[a,b]$ and let $\mathcal{E}$ be a subspace. Then, there exists a unitary U on $\mathcal{E}$ such that*

$$f(A_{\mathcal{E}}) \leq Uf(A)_{\mathcal{E}}U^*.$$

If A acts on an infinite dimensional space and $\mathcal{E}$ is an infinite dimensional subspace, then the right hand side needs an additional rI term, where I is the identity and $r > 0$ is arbitrarily small. We do not kwnow wether such an additional rI term is necessary. The same remark holds for the following related result. Recall that an isometric column $\{Z_i\}_{i=1}^m$ means that $\sum_{i=1}^m Z_i^*Z_i = I$.

Theorem 1.2. *Let f be a monotone, or more generally unitary, convex function on $[a,b]$ and let $\{A_i\}_{i=1}^m$ be Hermitians with spectra in $[a,b]$. If $\{Z_i\}_{i=1}^m$ is an isometric column, then there exists a unitary U such that*

$$f(\sum_i Z_i^*A_iZ_i) \leq U\{\sum_i Z_i^*f(A_i)Z_i\}U^*.$$

In particular, if A, B are two Hermitians with spectrum in $[a,b]$, the above result says

$$f\left(\frac{A+B}{2}\right) \leq U\cdot\frac{f(A)+f(B)}{2}\cdot U^*.$$

Such an inequality does not extend to all convex functions, a simple counterexample being the absolute value.

Corollary 1.3. *Let f be a monotone, or more generally unitary, convex function on $[a,b]$ and let A be a Hermitian with spectrum in $[a,b]$. If Z is a contraction, $0 \in [a,b]$ and $f(0) \leq 0$, then there exists a unitary U such that*

$$f(Z^*AZ) \leq UZ^*f(A)ZU^*.$$

The preceding results can be rephrased as eigenvalues inequalities; for instance, Corollary 1.3 claims that $\lambda_k(f(Z^*AZ) \leq \lambda_k(Z^*f(A)Z)$ (where $\{\lambda_k(\cdot)\}$ stands for the sequence of eigenvalues arranged in decreasing order and counted with their multiplicities). Since a general convex function f can be approached by the sum of an affine function and a monotone convex function, we get the Brown-Kosaki Inequality

$$\operatorname{Tr} f(Z^*AZ) \leq \operatorname{Tr} Z^*f(A)Z. \tag{1}$$

By definition, Z is an expansive operator if Z^*Z is greater than or equal to the identity. We will prove that (1) admits the following companion result:

Theorem 1.4. *Let A be a positive operator and let Z be an expansive operator. Let f be a convex function defined on $[0,b]$, $b \geq \|Z^*AZ\|_\infty$, with $f(0) \leq 0$. Then,*

$$\operatorname{Tr} f(Z^*AZ) \geq \operatorname{Tr} Z^*f(A)Z. \tag{2}$$

We will see that (2), contrary to (1), can not be extended to inequalities between eigenvalues. We will also see that the assumption $A \geq 0$ can not be dropped.

Of course if f is concave and $f(0) \geq 0$, then inequalities such as (1) and (2) are reversed. Theorem 2.3 will imply the following

Theorem 1.5. *Let A be a positive operator, let Z be an expansive operator and let $f : [0,\infty) \longrightarrow [0,\infty)$ be a nondecreasing concave function. Then,*

$$\|f(Z^*AZ)\|_\infty \leq \|Z^*f(A)Z\|_\infty.$$

Chapter 2 ([7])

A norm $\|\cdot\|$ on operators on $\mathcal{H}$ is said symmetric if $\|UAV\| = \|A\|$ for all operators A and all unitaries U, V. A basic inequality for symmetric norm is

$$\|AB\| \leq \|BA\|$$

for all operators A and B with a normal product AB. When AB is positive we will establish the more general inequality:

Theorem 2.1. *Let A, B be operators with $AB \geq 0$ and let Z be a strictly positive operator with extremal eigenvalues a and b. Then, for every symmetric norm, the following sharp inequality holds,*

$$\|ZAB\| \leq \frac{a+b}{2\sqrt{ab}}\,\|BZA\|.$$

The above inequality is sharp because, Z being fixed, there is a rank one projection E for which $A = B = E$ entails equality.

Theorem 2.1 has the following two corollaries where the inequalities are sharp.

Corollary 2.2. *Let A be a positive operator and let Z be a strictly positive operator with extremal eigenvalues a and b. Then,*

$$\|AZ\|_\infty \leq \frac{a+b}{2\sqrt{ab}}\,\rho(AZ)$$

and

$$\|AZ\|_1 \leq \frac{a+b}{2\sqrt{ab}}\,\mathrm{Tr}\,AZ.$$

Corollary 2.3. *Let A be a positive contraction and let Z be a strictly positive operator with extremal eigenvalues a and b. Then,*

$$AZA \leq \frac{(a+b)^2}{4ab}\,Z.$$

From Corollary 2.3 one may derive a (sharp) reverse inequality to (B), first proved by B. Mond and J.E. Pecaric:

$$(Z_{\mathcal{E}})^{-1} \geq \frac{4ab}{(a+b)^2}\,(Z^{-1})_{\mathcal{E}}$$

for every subspace $\mathcal{E}$. A companion result is

Theorem 2.4. *Let $f : [0,\infty) \longrightarrow [0,\infty)$ be operator convex and let Z be a strictly positive operator with extremal eigenvalues a and b. Then, for every subspace $\mathcal{E}$,*

$$f(Z_{\mathcal{E}}) \geq \frac{4ab}{(a+b)^2}\,(f(Z))_{\mathcal{E}}.$$

This result is a reverse inequality to Davis inequality (D), and similarly we will give reverse inequalities to Hansen-Pedersen inequalities (J) and (C). For instance, let us give a special case of the reverse inequality to (J). If A, B are positive operators with spectra in $[r, 2r]$, $r > 0$, then

$$f(\alpha A + \beta B) \geq \frac{8}{9} \cdot \{\alpha f(A) + \beta f(B)\}$$

for all operator convex functions $f : [0,\infty) \longrightarrow [0,\infty)$ and all $\alpha,\beta > 0$ with $\alpha+\beta = 1$. For $f(t) = t^2$, the constant $8/9$ is optimal.

Chapter 3 ([4],[5])

An operator Z on $\mathcal{H}\oplus\mathcal{H}$ or $\oplus^k\mathcal{H}$ is a dilation of the operator A on $\mathcal{H}$ if

$$Z = \begin{pmatrix} A & \star \\ \star & \star \end{pmatrix}, \text{ or } Z = \begin{pmatrix} A & \star & \dots \\ \star & \star & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}.$$

In the above dilations there is a lack of symmetry between the summands in $\mathcal{H}\oplus\mathcal{H}$ or $\oplus^k\mathcal{H}$. We then introduce the following natural notion: An operator Z on $\oplus^k\mathcal{H}$ is said to be a *total dilation* of the operator A on $\mathcal{H}$ if the operator diagonal of Z consists of a repetition of A,

$$Z = \begin{pmatrix} A & \star & \dots \\ \star & A & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}.$$

We will establish the following result (the stars hold for unspecified entries):

Theorem 3.1. *Let A, B be strictly positive operators on $\mathcal{H}$. Then, the condition $A^{-1} \le B$ is equivalent to the existence of a strictly positive operator Z on $\mathcal{H}\oplus\mathcal{H}$ such that*

$$Z = \begin{pmatrix} A & \star \\ \star & A \end{pmatrix} \qquad \textit{and} \qquad Z^{-1} = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

By (B) or (D), it is well-known that the existence of Z entails the relation $A^{-1} \le B$. The claim of the theorem is the converse implication.

Theorem 3.1 and some other results will lead us to state two conjectures:

Conjecture 3.2. *Let A, B be strictly positive operators on $\mathcal{H}$ and let $f : (0,\infty) \longrightarrow (0,\infty)$ be onto, nonlinear and operator convex. Then, the following statements are equivalent:*

(1) $A \le B$.

(2) *There exists a strictly positive operator Z on $\mathcal{H}\oplus\mathcal{H}$ such that*

$$A = f(Z_{\mathcal{H}}) \qquad \textit{and} \qquad B = f(Z)_{\mathcal{H}}.$$

We recall that the above implication (2) $\Rightarrow$ (1) is Davis' characterization (D) of operator convexity. Thus, we conjecture the converse implication. Our second conjecture is

Conjecture 3.3. *Let A, B be strictly positive operators on a finite dimensional space $\mathcal{H}$ and let $f : (0,\infty) \to (-\infty,\infty)$ be strongly convex. Then, the condition $f(A) < B$ ensures the existence of a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = f(Z)_{\mathcal{H}}$.*

Here, f is said to be strongly convex if its epigraph equals the convex hull of its graph.

If $\mathcal{H}$ is a space with an even finite dimension, we then say that the orthonormal decomposition $\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^{\perp}$ is a *halving* decomposition whenever $\dim \mathcal{F} = (1/2) \dim \mathcal{H}$. A key lemma for a standard proof of the Hausdorff-Toeplitz Theorem states that any operator A on a two-dimensional space can be written

$$A = \begin{pmatrix} b & \star \\ \star & b \end{pmatrix}.$$

with respect to some orthonormal basis. The following more general fact holds:

Theorem 3.4. *Let A be an operator on a space $\mathcal{H}$ with an even finite dimension. Then there exists a halving decomposition $\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^{\perp}$ for which we have a total dilation*

$$A = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

To be very precise, this theorem says that $A_{\mathcal{F}}$ and $A_{\mathcal{F}}^{\perp}$ are unitarily congruent. Its proof is not an adaption of the two dimensional case. This result raises several questions about the set of all operators B which can be totally dilated into A.

Chapter 4 ([1] [2])

Many inequalities involving operators occur from the various ways of arranging the terms of a product. One of the simplest inequality of this kind is

$$\|A^s Z A^t\|_\infty \le \|Z A^{s+t}\|_\infty. \tag{3}$$

Here, A is a positive operator (i.e. $A = X^*X$), Z is a normal operator, s et t are two nonnegative reals and $\|\cdot\|_\infty$ denotes the usual operator norm. This inequality raises two questions:

Can we replace the pair (A^s, A^t) by a more general pair (A, B), maybe at the cost of a numerical constant ?

The norm of an operator is also its first singular value. Is there a substitute of (3) for the other singular values ?

To generalize the pairs (A^s, A^t), we will say that two positive operators A and B form a monotone pair if there exist a positive operator C and two nondecreasing functions f, g, such that $A = f(C)$ and $B = g(C)$.

Proposition 4.1. *Let (A,B) be a monotone pair of positive operators. Let Z be a positive operator with largest and smallest nonzero eigenvalues a and b. Then*

$$\|AZB\|_\infty \leq \frac{a+b}{2\sqrt{ab}}\,\|ZAB\|_\infty.$$

If Z is a projection E, the previous inequality reduces to

$$\|AEB\|_\infty \leq \|EAB\|_\infty.$$

This will imply the following result:

Theorem 4.2. *Let (A,B) be a monotone pair of positive operators and let E be the projection onto a subspace $\mathcal{E}$. Then*

$$\mathrm{Sing}(AEB) \leq \mathrm{Sing}(EAB).$$

Consequently,

$$\mathrm{Eig}(A_{\mathcal{E}}B_{\mathcal{E}}) \leq \mathrm{Eig}((AB)_{\mathcal{E}}) \tag{4}$$

and

$$\mathrm{Eig}(A_{\mathcal{E}}B_{\mathcal{E}}A_{\mathcal{E}}) \leq \mathrm{Eig}((ABA)_{\mathcal{E}}).$$

Here, $\mathrm{Sing}(\cdot)$, resp. $\mathrm{Eig}(\cdot)$, stands for the sequence of singular values, resp. eigenvalues, arranged in decreasing order and counted with their multiplicities. An immediate consequence of (4) is the determinantal inequality

$$\det A_{\mathcal{E}}\cdot\det A_{\mathcal{E}} \leq \det(AB)_{\mathcal{E}}$$

in particular

$$\langle h, Ah\rangle\langle h, Bh\rangle \leq \langle h, ABh\rangle$$

for every norm one vector h. Actually, this inequality is the starting point of the previous results.

Finally, we will prove the following result for the Hilbert-Schmidt (or Frobenius) norm, $\|\cdot\|_2$.

Theorem 4.3. *Let Z be a norm operator and let A and B be two positive operators. Then:*

(1) *If (A,B) is monotone,*

$$\|AZB\|_2 \leq \|ZAB\|_2$$

(2) *If (A,B) is antimonotone,*

$$\|AZB\|_2 \geq \|ZAB\|_2. \tag{5}$$

Here, we say that (A, B) is antimonotone if there exist a positive operator C and two functions f, g, one nondecreasing, the other nonincreasing, such that $A = f(C)$ and $B = g(C)$. Note that (5) may fail in the infinite dimensional setting.

Chapter 5 ([3])

The results of this last chapter are specific to the infinite dimensional setting.

The notion of compression has an obvious extension: If B is an operator on a space $\mathcal{F}$ with $\dim \mathcal{F} \le \dim \mathcal{H}$, we still say that B is a compression of A if there is an isometry $V : \mathcal{F} \longrightarrow \mathcal{H}$ such that $B = V^*AV$. Thus, identifying B with VBV^* (equivalently, identifiyng $\mathcal{F}$ and $V(\mathcal{F})$), we can write

$$A = \begin{pmatrix} B & \star \\ \star & \star \end{pmatrix}.$$

One also says that A dilates B or that A is a dilation of B.

Let $\mathcal{H}$ be an infinite dimensional (separable) Hilbert space. We ask the following question:

What are the operators A on $\mathcal{H}$ with the property that any strict contraction X can be realized as a compression of A ?

We will answer this question and, actually, a more general one. Let us define the essential numerical range of A as

$$W_e(A) = \{\ \lambda \mid \quad \text{there is an orthonormal system } \{\mathrm{e_n}\}_{n=1}^{\infty} \text{ with } \lim \langle \mathrm{e_n}, \mathrm{Ae_n} \rangle = \lambda \}.$$

Let A be an operator on an infinite dimensional space $\mathcal{H}$ and consider an orthonormal decomposition $\mathcal{H} = \oplus_{n=1}^{\infty} \mathcal{H}_n$. Denote by E_n the projection onto $\mathcal{H}_n$. We have the pinching

$$\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_{\mathcal{H}_n} = \sum_n E_n A E_n$$

The notion of pinching has an obvious extension: if $\{A_n\}$ is a sequence of operators acting on separable Hilbert spaces with A_n unitarily equivalent to $A_{\mathcal{H}_n}$ for all n, we also naturally write

$$\mathcal{P}(A) = \oplus_{n=1}^{\infty} A_n.$$

Note that $W_e(A)$ contains the open unit disc $\mathcal{D}$ if and only if there is a basis (or an o.n.s) $\{e_n\}_{n=1}^{\infty}$ such that the convex hulls $\mathrm{co}\{\langle e_k, Ae_k\rangle \ : \ k > n\}$ contain $\mathcal{D}$ for all n.

We have the following answer to our question:

Theorem 5.1. *Let A be an operator with $W_e(A) \supset \mathcal{D}$ and let $\{A_n\}_{n=1}^{\infty}$ be a sequence of operators such that $\sup_n \|A_n\|_\infty < 1$. Then, there is a pinching*

$$\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_n.$$

So, we have in particular:

Corollary 5.2. *Let A be an operator with $W_e(A) \supset \mathcal{D}$. For any strict contraction X, there is an isometry V such that $X = V^*AV$.*

References

[1] J.-C. Bourin, Some inequalities for norms on matrices and operators, Linear Algebra Appl. 292 (1999) 139-154.
[2] J.-C. Bourin, Singular values of compressions, restrictions and dilations, Linear Algebra Appl. 360 (2003) 259-272.
[3] J.-C. Bourin, Compressions and pinchings, J. Operator Theory 50 (2003) 211-220.
[4] J.-C. Bourin, Total dilations, Linear Algebra Appl. 368 (2003) 159-169
[5] J.-C. Bourin, Total dilations II, Linear Algebra Appl. 374 (2003). 19-29
[6] J.-C. Bourin, Convexity or concavity inequalities for Hermitian operators, Math. Ineq. Appl., to appear.
[7] J.-C. Bourin, Symmetric norms and reverse inequalities to Davis and Hansen-Pedersen characterizations of operator convexity, submitted in Math. Ineq. Appl.

Chapter 1

Convexity or concavity inequalities for Hermitian operators

Introduction

Given an operator A on a separable Hilbert space $\mathcal{H}$ and a subspace $\mathcal{E} \subset \mathcal{H}$, we denote by $A_{\mathcal{E}}$ the compression of A onto $\mathcal{E}$, i.e. the restriction of EAE to $\mathcal{E}$, E being the projection onto $\mathcal{E}$. If $\mathcal{E}$ is a finite dimensional subspace, we show that, for any Hermitian operator A and any monotone convex function f defined on the spectrum of A, there exits a unitary operator U on $\mathcal{E}$ such that the operator inequality

$$f(A_{\mathcal{E}}) \leq Uf(A)_{\mathcal{E}}U^*.$$

holds. Here, $f(A)_{\mathcal{E}}$ must be read as $(f(A))_{\mathcal{E}}$. This result together with the elementary method of its proof motivate the whole chapter. In Section 1 we prove the above inequality and give a version when $\dim \mathcal{E} = \infty$. We also study the map $p \longrightarrow \{(A^p)_{\mathcal{E}}\}^{1/p}$, $A \geq 0$, $0 < p < \infty$.

Section 2 is concerned with eigenvalues inequalities (equivalently operator inequalities) which improve some trace inequalities of Brown-Kosaki and Hansen-Pedersen: Given a monotone convex function f defined on the real line with $f(0) \leq 0$, a Hermitian operator A and a contractive operator Z acting on a finite dimensional space, there exists a unitary operator U such that

$$f(Z^*AZ) \leq UZ^*f(A)ZU^*.$$

In Section 3, we prove that

$$\mathrm{Tr}\, f(Z^*AZ) \leq \mathrm{Tr}\, Z^*f(A)Z$$

for every positive operator A and expansive operator Z on a finite dimensional space, and every concave function f defined on an interval $[0, b]$, $b \geq \|Z^*AZ\|_\infty$, with $f(0) \geq 0$ ($\|\cdot\|_\infty$ denotes the usual operator norm).

The last section deals with unitarily invariant norms inequalities associated with orthogonal decompositions.

1. Compressions and convex functions

By a classical result of C. Davis [6] (see also [1, p. 117-9]), a function f on (a,b) is operator convex if and only if for every subspace $\mathcal{E}$ and every Hermitian operator A whose spectrum lies in (a,b) one has

$$f(A_{\mathcal{E}}) \leq f(A)_{\mathcal{E}}. \tag{1}$$

What can be said about convex, not operator convex functions ? Let g be operator convex on (a,b) and let ϕ be a nondecreasing, convex function on $g((a,b))$. Then, $f = \phi \circ g$ is convex and we say that f is *unitary* convex on (a,b). Since $t \longrightarrow -t$ is trivially operator convex, we note that *the class of unitary convex functions contains the class of monotone convex functions.* The following result holds:

Theorem 1.1. *Let f be a monotone convex, or more generally unitary convex, function on (a,b) and let A be a Hermitian operator whose spectrum lies in (a,b).*

(1) *If $\mathcal{E}$ is a finite dimensional subspace, then there exists a unitary operator U on $\mathcal{E}$ such that*

$$f(A_{\mathcal{E}}) \leq Uf(A)_{\mathcal{E}}U^*.$$

(2) *If $\mathcal{E}$ is an infinite dimensional subspace and $r > 0$ is arbitrarily small, then there exists a unitary operator U on $\mathcal{E}$ such that*

$$f(A_{\mathcal{E}}) \leq Uf(A)_{\mathcal{E}}U^* + rI.$$

Proof. We first consider the case when $\mathcal{E}$ has finite dimension d.

We begin by assuming that f is monotone. Let $\{\lambda_k(X)\}_{k=1}^d$ denote the eigenvalues of the Hermitian operator X on $\mathcal{E}$, arranged in decreasing order and counted with their multiplicities. Let k be an integer, $1 \leq k \leq d$. There exists a spectral subspace $\mathcal{F} \subset \mathcal{E}$ for $A_{\mathcal{E}}$ (hence for $f(A_{\mathcal{E}})$), $\dim \mathcal{F} = k$, such that

$$\begin{aligned}
\lambda_k[f(A_{\mathcal{E}})] &= \min_{h \in \mathcal{F};\ \|h\|=1} \langle h, f(A_{\mathcal{F}})h \rangle \\
&= \min\{f(\lambda_1(A_{\mathcal{F}}))\,;\, f(\lambda_k(A_{\mathcal{F}}))\} \\
&= \min_{h \in \mathcal{F};\ \|h\|=1} f(\langle h, A_{\mathcal{F}}h \rangle) \\
&= \min_{h \in \mathcal{F};\ \|h\|=1} f(\langle h, Ah \rangle)
\end{aligned}$$

where at the second and third steps we use the monotony of f. The convexity of f implies

$$f(\langle h, Ah \rangle) \leq \langle h, f(A)h \rangle$$

for all normalized vectors h. Therefore, by the minmax principle,

$$\begin{aligned}
\lambda_k[f(A_{\mathcal{E}})] &\leq \min_{h \in \mathcal{F};\ \|h\|=1} \langle h, f(A)h \rangle \\
&\leq \lambda_k[f(A)_{\mathcal{E}}].
\end{aligned}$$

This statement is equivalent to the existence of a unitary operator U on $\mathcal{E}$ satisfying to the conclusion of the theorem.

If f is unitary convex, $f = \phi \circ g$ with g operator convex and ϕ nondecreasing convex; inequality (1) applied to g combined with the fact that ϕ is nondecreasing yield a unitary operator V on $\mathcal{E}$ for which

$$\phi \circ g(A_{\mathcal{E}}) \leq V\phi[g(A)_{\mathcal{E}}]V^*.$$

Applying the first part of the proof to ϕ gives a unitary operator W on $\mathcal{E}$ such that

$$\phi[g(A)_{\mathcal{E}}] \leq W[\phi \circ g(A)]_{\mathcal{E}}W^*.$$

We then get the result by letting $U = VW$.

Now, we turn to the case when $\mathcal{E}$ has infinite dimension. Though the proof is similar to the previous one, we have to be more careful.

Given a Hermitian operator on an infinite dimensional space we may define, as in the finite dimensional setting, a sequence of numbers $\{\lambda_k(X)\}_{k=1}^{\infty}$ by

$$\lambda_k(X) = \sup_{\{\mathcal{F}\,:\,\dim \mathcal{F}=k\}} \inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, Xh\rangle$$

We note that $\{\lambda_k(X)\}_{k=1}^{\infty}$ is a nonincreasing sequence whose limit is the upper bound of the essential spectrum of X. We define another sequence of numbers, $\{\lambda_{-k}(X)\}_{k=1}^{\infty}$ by

$$\lambda_{-k}(X) = \sup_{\{\mathcal{F}\,:\,\operatorname{codim} \mathcal{F}=k-1\}} \inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, Xh\rangle.$$

Then, $\{\lambda_{-k}(X)\}_{k=1}^{\infty}$ is a nondecreasing sequence whose limit is the lower bound of the essential spectrum of X. The following fact (a) is obvious and fact (b) is proved in the addenda.

(a) If X and Y are two Hermitian operators such that $X \leq Y$, then $\lambda_k(X) \leq \lambda_k(Y)$ and $\lambda_{-k}(X) \leq \lambda_{-k}(Y)$ for all $k = 1, \ldots$.

(b) Let $r > 0$. If X and Y are two Hermitian operators such that $\lambda_k(X) \leq \lambda_k(Y)$ and $\lambda_{-k}(X) \leq \lambda_{-k}(Y)$ for all $k = 1, \ldots$, then there exists a unitary operator U such that $X \leq UYU^* + rI$.

From facts (a) and (b) we infer that, given two Hermitian operators X, Y with $X \leq Y$ and a nondecreasing continuous function ϕ, there exists a unitary operator U such that $\phi(X) \leq U\phi(Y)U^* + rI$. This observation allows us, by the same reasoning as in the first part of the proof, to restrict ourselves to the case when f is nondecreasing.

From fact (b) we also infer that it suffices to show that $\lambda_k(f(A_{\mathcal{E}})) \leq \lambda_k(f(A)_{\mathcal{E}})$ and $\lambda_{-k}(f(A_{\mathcal{E}})) \leq \lambda_{-k}(f(A)_{\mathcal{E}})$ for all integers k. We consider the case $\lambda_{-k}(\cdot)$, the other one being similar. We fix k and distinguish two cases.

1. $\lambda_{-k}(A_{\mathcal{E}})$ is an eigenvalue of $A_{\mathcal{E}}$. Then, for $1 \leq j \leq k$, $\lambda_{-j}(f(A_{\mathcal{E}}))$ are eigenvalues for $f(A_{\mathcal{E}})$. Consequently, there exists a subspace $\mathcal{F} \subset \mathcal{E}$, $\operatorname{codim}_{\mathcal{E}} \mathcal{F} = k-1$, such

that

$$\begin{aligned}
\lambda_{-k}(f(A_{\mathcal{E}})) &= \min_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, f(A_{\mathcal{E}})h\rangle \\
&= \min_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} f(\langle h, A_{\mathcal{E}}h\rangle) \\
&\le \inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, f(A)h\rangle \le \lambda_{-k}(f(A)_{\mathcal{E}})
\end{aligned}$$

where at the second and third lines we have used the nondecreasingness of f and its convexity, respectively.
2. $\lambda_{-k}(A_{\mathcal{E}})$ is not an eigenvalue of $A_{\mathcal{E}}$ (so, $\lambda_{-k}(A_{\mathcal{E}})$ is the lower bound of the essential spectrum of $A_{\mathcal{E}}$). Fix $\varepsilon > 0$ and choose $\delta > 0$ such that $|f(x) - f(y)| \le \varepsilon$ for all x, y are in the convex hull of the spectrum of A with $|x - y| \le \delta$. There exists a subspace $\mathcal{F} \subset \mathcal{E}$, $\operatorname{codim}_{\mathcal{E}} \mathcal{F} = k - 1$, such that

$$\lambda_{-k}(A_{\mathcal{E}}) \le \inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, A_{\mathcal{E}}h\rangle + \delta.$$

Since f is continuous nondecreasing we have $f(\lambda_{-k}(A_{\mathcal{E}})) = \lambda_{-k}(f(A_{\mathcal{E}}))$ so that, as f is nondecreasing,

$$\lambda_{-k}(f(A_{\mathcal{E}})) \le f\left(\inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} \langle h, A_{\mathcal{E}}h\rangle + \delta\right).$$

Consequently,

$$\lambda_{-k}(f(A_{\mathcal{E}})) \le \inf_{\{h\in\mathcal{F}\,:\,\|h\|=1\}} f(\langle h, A_{\mathcal{E}}h\rangle) + \varepsilon,$$

so, using the convexity of f and the definition of $\lambda_{-k}(\cdot)$, we get

$$\lambda_{-k}(f(A_{\mathcal{E}})) \le \lambda_{-k}(f(A)_{\mathcal{E}}) + \varepsilon.$$

By letting $\varepsilon \longrightarrow 0$, the proof is complete. $\square$

Later, we will see that Theorem 1.1 can *not* be extended to all convex functions f (Example 2.4).

Of course Theorem 1.1 holds with a reverse inequality for monotone concave functions f (or $f = \phi \circ g$, g operator convex and ϕ decreasing concave).

Given a compact positive operator A and a subspace $\mathcal{E}$, it is natural to study the behaviour of the map

$$p \longrightarrow \{(A^p)_{\mathcal{E}}\}^{1/p}$$

on $(0, \infty)$.

Theorem 1.2. *Let $A = \sum_k \lambda_k(A)\, f_k \otimes f_k$ be a positive, compact operator and let $\mathcal{E}$ be a subspace,* $\dim \mathcal{E} = d < \infty$. *Assume $\mathcal{E} \cap \operatorname{span}\{f_j : j > d\} = 0$. Then, for every integer $k \le d$, the map $p \longrightarrow \lambda_k(\{(A^p)_{\mathcal{E}}\}^{1/p})$ increases on $(0, \infty)$ and*

$$\lim_{p\to\infty} \lambda_k(\{(A^p)_{\mathcal{E}}\}^{1/p}) = \lambda_k(A).$$

Moreover the family $\{(A^p)_{\mathcal{E}}\}^{1/p}$ *converges in norm when* $p \to \infty$ *and the map* $p \longrightarrow \{(A^p)_{\mathcal{E}}\}^{1/p}$ *is increasing for the Loewner order on* $[1, \infty)$.

Proof. Let $p > 0$ and $r > 1$. By Theorem 1.1, there exists a unitary $U : \mathcal{E} \longrightarrow \mathcal{E}$ such that

$$\{(A^p)_{\mathcal{E}}\}^r \leq U(A^{pr})_{\mathcal{E}} U^*,$$

hence, for all k,

$$\lambda_k^r((A^p)_{\mathcal{E}}) \leq \lambda_k((A^{pr})_{\mathcal{E}}),$$

so,

$$\lambda_k(\{(A^p)_{\mathcal{E}}\}^{1/p}) \leq \lambda_k(\{(A^{pr})_{\mathcal{E}}\}^{1/pr}),$$

that is, the map $p \longrightarrow \lambda_k(\{(A^p)_{\mathcal{E}}\}^{1/p})$ increases on $(0, \infty)$. In order to study its convergence when $p \to \infty$, we first show that

$$\lim_{p\to\infty} \lambda_1((EA^pE)^{1/p}) = \lambda_1(A) \tag{2}$$

where E denotes the projection onto $\mathcal{E}$. We note that

$$\lim_{p\to\infty} \lambda_1((EA^pE)^{1/p}) \leq \lambda_1(A). \tag{3}$$

Recall that $A = \sum_k \lambda_k(A)\, f_k \otimes f_k$. Since by assumption $f_1 \notin \mathcal{E}^{\perp}$, there exists a normalized vector g in $\mathcal{E}$ such that $\langle g, f_1 \rangle \neq 0$. Setting $G = g \otimes g$, we have

$$\lambda_1((GA^pG)^{1/p}) = \langle g, A^p g \rangle^{1/p} = \Big(\sum_k \lambda_k^p(A) |\langle g, f_k \rangle|^2 \Big)^{1/p}.$$

The above expression is a weighted l^p-norm of the sequence $\{\lambda_k(A)\}$. When $p \to \infty$, this tends towards the l^∞-norm which is $\lambda_1(A)$. Since

$$\lambda_1((GA^pG)^{1/p}) \leq \lambda_1((EA^pE)^{1/p})$$

we then deduce with (3) that (2) holds.

In order to prove the general limit assertion, we consider antisymmetric tensor products. Let F be the projection onto $\mathcal{F} = \operatorname{span}\{f_j : j \leq \dim \mathcal{E}\}$. By assumption F maps $\mathcal{E}$ onto $\mathcal{F}$. Therefore $\wedge^k(F)$ maps $\wedge^k(\mathcal{E})$ onto $\wedge^k(\mathcal{F})$ and we may find a norm one tensor $\gamma \in \wedge^k(\mathcal{E})$ such that $\langle \gamma, f_1 \wedge \cdots \wedge f_k \rangle \neq 0$. Hence, with $\wedge^k E$ and $\wedge^k A$ in place of E and A, $1 \leq k \leq \dim \mathcal{E}$, we may apply (2) to obtain

$$\lim_{p\to\infty} \lambda_1(\wedge^k (EA^pE)^{1/p}) = \lambda_1(\wedge^k A)$$

meaning that

$$\lim_{p\to\infty} \prod_{1\leq j\leq k} \lambda_j((EA^pE)^{1/p}) = \prod_{1\leq j\leq k} \lambda_j(A).$$

From these relations we infer that, for every $k \leq \dim \mathcal{E}$, we have

$$\lim_{p\to\infty} \lambda_k((EA^pE)^{1/p}) = \lambda_k(A)$$

proving the main assertion of the theorem.

For $p, r \geq 1$ we have

$$(EA^{pr}E)^{1/r} \geq EA^pE.$$

by Hansen's inequality [6]. Since $t \longrightarrow t^{1/p}$ is operator monotone by the Loewner theorem [9, p. 2], we have

$$(EA^{pr}E)^{1/pr} \geq (EA^pE)^{1/p}.$$

Thus $p \longrightarrow (EA^pE)^{1/p}$ increases on $[1, \infty)$. Since this map is bounded, it converges in norm. $\square$

Question 1.4. Can we drop the rI term in Theorem 1.1 ?

2. Contractions and convex functions

In [7] and [8], the authors show that inequality (1) is equivalent to the following statement.

Theorem 2.1. (Hansen-Pedersen) *Let A and $\{A_i\}_{i=1}^m$ be Hermitian operators and let f be an operator convex function defined on an interval $[a, b]$ containing the spectra of A and A_i, $i = 1, \ldots m$.*

(1) *If Z is a contraction, $0 \in [a, b]$ and $f(0) \leq 0$,*

$$f(Z^*AZ) \leq Z^*f(A)Z.$$

(2) *If $\{Z_i\}_{i=1}^m$ is an isometric column,*

$$f(\sum_i Z_i^*A_iZ_i) \leq \sum_i Z_i^*f(A_i)Z_i.$$

Here, an isometric column $\{Z_i\}_{i=1}^m$ means that $\sum_{i=1}^m Z_i^*Z_i = I$.

In a similar way, Theorem 1.1 is equivalent to the next one. We state it in the finite dimensional setting, but an analogous version exists in the infinite dimensional setting by adding a rI term in the right hand side of the inequalities.

Theorem 2.2. *Let A and $\{A_i\}_{i=1}^m$ be Hermitian operators on a finite dimensional space and let f be a monotone, or more generally unitary, convex function defined on an interval $[a, b]$ containing the spectra of A and A_i, $i = 1, \ldots m$.*

(1) *If Z is a contraction, $0 \in [a, b]$ and $f(0) \leq 0$, then there exists a unitary operator U such that*

$$f(Z^*AZ) \leq UZ^*f(A)ZU^*.$$

(2) *If* $\{Z_i\}_{i=1}^m$ *is an isometric column, then there exists a unitary operator* U *such that*

$$f(\sum_i Z_i^* A_i Z_i) \le U\{\sum_i Z_i^* f(A_i) Z_i\}U^*.$$

Here, we give a first proof based on Theorem 1.1. A more direct proof is given at the end of the section.

Proof. Theorem 2.2 and Theorem 1.1 are equivalent. Indeed, to prove Theorem 1.1, it suffices to consider the case of monotone convex functions f. Then, by a limit argument, we may assume that f is defined on the whole real line. Since we may also assume that $f(0) = 0$, Theorem 1.1 follows from Theorem 2.2 by taking Z as the projection onto $\mathcal{E}$.

Theorem 1.1 entails Theorem 2.2(1): to see that, we introduce the partial isometry V and the operator $\tilde{A}$ on $\mathcal{H} \oplus \mathcal{H}$ defined by

$$V = \begin{pmatrix} Z & 0 \\ (I - |Z|^2)^{1/2} & 0 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$

Denoting by $\mathcal{H}$ the first summand of the direct sum $\mathcal{H} \oplus \mathcal{H}$, we observe that

$$f(Z^*AZ) = f(V^*\tilde{A}V) : \mathcal{H} = V^* f(\tilde{A}_{V(\mathcal{H})})V : \mathcal{H}.$$

Applying Theorem 1.1 with $\mathcal{E} = V(\mathcal{H})$, we get a unitary operator W on $V(\mathcal{H})$ such that

$$f(Z^*AZ) \le V^* W f(\tilde{A})_{V(\mathcal{H})} W^* V : \mathcal{H}.$$

Equivalently, there exists a unitary operator U on $\mathcal{H}$ such that

$$\begin{aligned} f(Z^*AZ) \le& UV^* f(\tilde{A})_{V(\mathcal{H})}(V : \mathcal{H})U^* \\ =& UV^* \begin{pmatrix} f(A) & 0 \\ 0 & f(0) \end{pmatrix} (V : \mathcal{H})U^* \\ =& U\{Z^* f(A) Z + (I - |Z|^2)^{1/2} f(0)(I - |Z|^2)^{1/2}\}U^*. \end{aligned}$$

Using $f(0) \le 0$ we obtain the first claim of Theorem 2.2.

Similarly, Theorem 1.1 implies Theorem 2.2(2) (we may assume $f(0) = 0$) by considering the partial isometry and the operator on $\oplus^m \mathcal{H}$,

$$\begin{pmatrix} Z_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ Z_m & 0 & \cdots & 0 \end{pmatrix}, \qquad \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix}.$$

□

We note that Theorem 2.2 strengthens some well-known trace inequalities:

Corollary 2.3. *Let A and $\{A_i\}_{i=1}^m$ be Hermitian operators on a finite dimensional space and let f be a convex function defined on an interval $[a,b]$ containing the spectra of A and A_i, $i=1,\dots m$.*

(1) (Brown-Kosaki [2]) *If Z is a contraction, $0 \in [a,b]$ and $f(0) \le 0$, then*

$$\operatorname{Tr} f(Z^*AZ) \le \operatorname{Tr} Z^*f(A)Z.$$

(2) (Hansen-Pedersen [8]) *If $\{Z_i\}_{i=1}^m$ is an isometric column, then*

$$\operatorname{Tr} f(\sum_i Z_i^*A_iZ_i) \le \operatorname{Tr} \{\sum_i Z_i^*f(A_i)Z_i\}.$$

Proof. By a limit argument, we may assume that f is defined on the whole real line and can be written as $f(x) = g(x) - \lambda x$ for some convex monotone function g and some scalar λ. We then apply Theorem 2.2 to g. $\square$

A very special case of Theorem 2.2(2) is: *Given two Hermitian operators A, B and a monotone convex or unitary convex function f on a suitable interval, there exists a unitary operator U such that*

$$f(\frac{A+B}{2}) \le U\frac{f(A)+f(B)}{2}U^*.$$

This shows that Theorem 2.2, and consequently Theorem 1.1, can not be valid for all convex functions:

Example 2.4. Theorems 1.1 and 2.2 are not valid for a simple convex function such as $t \longrightarrow |t|$. Indeed, it is well-known that the inequality

$$|A+B| \le U(|A|+|B|)U^* \tag{5}$$

is not always true, even for Hermitians A, B. We reproduce the counterexample [8, p. 1]: Take

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

Then, as the two eigenvalues of $|A+B|$ equal to $\sqrt{2}$ while $|A|+|B|$ has an eigenvalue equal to $2-\sqrt{2}$, inequality (5) can not hold.

In connection with Example 2.4, a famous result (e.g., [1, p. 74]) states the existence, for any operators A, B on a finite dimensional space, of unitary operators U, V such that

$$|A+B| \le U|A|U^* + V|B|V^*. \tag{6}$$

In the case of Hermitians A, B, the above inequality has the following generalization:

Proposition 2.5. *Let A, B be hermitian operators on a finite dimensional space and let f be an even convex function on the real line. Then, there exist unitary operators U, V such that*

$$f\left(\frac{A+B}{2}\right) \le \frac{Uf(A)U^* + Vf(B)V^*}{2}.$$

Proof. Since $f(X) = f(|X|)$, inequality (6) and the fact that f is increasing on $[0,\infty)$ give unitary operators U_0, V_0 such that

$$f\left(\frac{A+B}{2}\right) \le f\left(\frac{U_0|A|U_0^* + V_0|B|V_0^*}{2}\right).$$

Since f is monotone convex on $[0,\infty)$, Theorem 2.2 completes the proof. □

Question 2.6. Does Proposition 2.5 hold for all convex functions defined on the whole real line ?

We close this section by giving a direct and proof of Theorem 2.2, which is a simple adaptation of the proof of Theorem 1.1.

Proof. We restrict ourselves to the case when f is monotone. We will use the following observation which follows from the standard Jensen's inequality: for any vector u of norm less than or equal to one, since f is convex and $f(0) \le 0$,

$$f(\langle u, Au\rangle) \le \langle u, f(A)u\rangle.$$

We begin by proving assertion (1). We have, for each integer k less than or equal to the dimension of the space, a subspace $\mathcal{F}$ of dimension k such that

$$\begin{aligned}
\lambda_k[f(Z^*AZ)] &= \min_{h\in\mathcal{F};\ \|h\|=1} \langle h, f(Z^*AZ)h\rangle \\
&= \min_{h\in\mathcal{F};\ \|h\|=1} f(\langle h, Z^*AZh\rangle) \\
&= \min_{h\in\mathcal{F};\ \|h\|=1} f(\langle Zh, AZh\rangle).
\end{aligned}$$

where we have used the monotony of f. Then, using the above observation and the minmax principle,

$$\begin{aligned}
\lambda_k[f(Z^*AZ)] &\le \min_{h\in\mathcal{F};\ \|h\|=1} \langle Zh, f(A)Zh\rangle \\
&\le \lambda_k[Z^*f(A)Z].
\end{aligned}$$

We turn to assertion (2). For any integer k less than or equal to the dimension of the space, we have a subspace $\mathcal{F}$ of dimension k such that

$$\begin{aligned}
\lambda_k[f(\sum Z_i^* A_i Z_i)] &= \min_{h\in\mathcal{F};\ \|h\|=1} \langle h, f(\sum Z_i^* A_i Z_i)h\rangle \\
&= \min_{h\in\mathcal{F};\ \|h\|=1} f(\langle h, \sum Z_i^* A_i Z_i h\rangle) \\
&= \min_{h\in\mathcal{F};\ \|h\|=1} f(\sum \|Z_i h\|^2(\langle Z_i h, A_i Z_i h\rangle/\|Z_i h\|^2)) \\
&\le \min_{h\in\mathcal{F};\ \|h\|=1} \sum \|Z_i h\|^2 f(\langle Z_i h, A_i Z_i h\rangle/\|Z_i h\|^2) && (7)\\
&\le \min_{h\in\mathcal{F};\ \|h\|=1} \sum \langle Z_i h, f(A_i) Z_i h\rangle) && (8)\\
&\le \min_{h\in\mathcal{F};\ \|h\|=1} \langle h, \sum Z_i^* f(A_i) Z_i h\rangle) \\
&\le \lambda_k[\sum Z_i^* f(A_i) Z_i]
\end{aligned}$$

where we have used in (7) and (8) the convexity of f. $\square$

3. Inequalities involving expansive operators

In this section we are in the finite dimensional setting.

For two reals a, z, with $z > 1$, we have $f(za) \ge zf(a)$ for every convex function f with $f(0) \le 0$. In view of Theorem 2.2, one might expect the following result: If Z is an expansive operator (i.e. $Z^*Z \ge I$), A is a Hermitian operator and f is a convex function with $f(0) \le 0$, then there exists a unitary operator U such that

$$f(Z^*AZ) \ge UZ^*f(A)ZU^*. \tag{*}$$

But, as we shall see, this is not always true, even for $A \ge 0$ and f nonnegative with $f(0) = 0$. Let us first note the following remark:

Remark 3.1. Let $f : [0,\infty) \longrightarrow [0,\infty)$ be a continuous function with $f(0) = 0$. If

$$\operatorname{Tr} f(Z^*AZ) \le \operatorname{Tr} Z^*f(A)Z$$

for every positive operator A and every contraction Z, then f is convex.

To check this, it suffices to consider:

$$A = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \quad \text{and} \quad Z = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \end{pmatrix}$$

where x, y are arbitrary nonnegative scalars. Indeed, $\operatorname{Tr} f(Z^*AZ) = f((x+y)/2)$ and $\operatorname{Tr} Z^*f(A)Z = (f(x)+f(y))/2$.

We may now state

Proposition 3.2. *Let $f : [0,\infty) \longrightarrow [0,\infty)$ be a continuous one to one function with $f(0) = 0$ and $f(\infty) = \infty$. Then, the following conditions are equivalent:*

(1) *The function $g(t) = 1/f(1/t)$ is convex on $[0,\infty)$.*

(2) *For every positive operator A and every expansive operator Z, there exists a unitary operator U such that*

$$Z^*f(A)Z \leq Uf(Z^*AZ)U^*.$$

Proof. We may assume that A is invertible. If g is convex, (note that g is also nondecreasing) then Theorem 2.2 entails that

$$g(Z^{-1}A^{-1}Z^{-1*}) \leq U^*Z^{-1}g(A^{-1})Z^{-1*}U$$

for some unitary operator U. Taking the inverses, since $t \longrightarrow t^{-1}$ is operator decreasing on $(0,\infty)$, this is the same as saying

$$Z^*f(A)Z \leq Uf(Z^*AZ)U^*.$$

The converse direction follows, again by taking the inverses, from the above remark. $\square$

It is not difficult to find convex functions $f : [0,\infty) \longrightarrow [0,\infty)$, with $f(0) = 0$ which do not satisfy the conditions of Proposition 3.2. So, in general, (*) can not hold. Let us give an explicit simple example.

Example 3.3. Let $f(t) = t + (t-1)_+$ and

$$A = \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix}, \qquad Z = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Then $\lambda_2(f(ZAZ)) = 0.728.. < 0.767.. = \lambda_2(Zf(A)Z)$. So, (*) does not hold.

In spite of the previous example, we have the following positive result:

Lemma 3.4. *Let A be a positive operator, let Z be an expansive operator and β be a nonnegative scalar. Then, there exists a unitary operator U such that*

$$Z^*(A-\beta I)_+Z \leq U(Z^*AZ-\beta I)_+U^*.$$

Proof. We will use the following simple fact: If B is a positive operator with $\mathrm{Sp}B \subset \{0\} \cup (x,\infty)$, then we also have $\mathrm{Sp}Z^*BZ \subset \{0\} \cup (x,\infty)$. Indeed Z^*BZ and $B^{1/2}ZZ^*B^{1/2}$ (which is greater than B) have the same spectrum.

Let P be the spectral projection of A corresponding to the eigenvalues strictly greater than β and let $A_\beta = AP$. Since $t \longrightarrow t_+$ is nondecreasing, there exists a unitary operator V such that

$$(Z^*AZ-\beta I)_+ \geq V(Z^*A_\beta Z-\beta I)_+V^*$$

Since $Z^*(A-\beta I)_+Z = Z^*(A_\beta - \beta I)_+Z$ we may then assume that $A = A_\beta$. Now, the above simple fact implies

$$(Z^*A_\beta Z - \beta I)_+ = Z^*A_\beta Z - \beta Q$$

where $Q = \mathrm{supp} Z^*A_\beta Z$ is the support projection of $Z^*A_\beta Z$. Hence, it suffices to show the existence of a unitary operator W such that

$$Z^*A_\beta Z - \beta Q \geq WZ^*(A_\beta - \beta P)ZW^* = WZ^*A_\beta ZW^* - \beta WZ^*PZW^*.$$

But, here we can take $W = I$. Indeed, we have

$$\mathrm{supp} Z^*PZ = Q\ (*) \quad \text{and} \quad \mathrm{Sp} Z^*PZ \subset \{0\} \cup [1,\infty)\ (**)$$

where $(**)$ follows from the above simple fact and the identity $(*)$ from the observation below with $X = P$ and $Y = A_\beta$.

Observation. If X, Y are two positive operators with $\mathrm{supp} X = \mathrm{supp} Y$, then for every operator Z we also have $\mathrm{supp} Z^*XZ = \mathrm{supp} Z^*YZ$.

To check this, we establish the corresponding equality for the kernels,

$$\ker Z^*XZ = \{h \ :\ Zh \in \ker X^{1/2}\} = \{h \ :\ Zh \in \ker Y^{1/2}\} = \ker Z^*YZ.$$

□

Theorem 3.5. *Let A be a positive operator and Z be an expansive operator. Assume that f is a continuous function defined on $[0,b]$, $b \geq \|Z^*AZ\|_\infty$. Then,*

(1) *If f is concave and $f(0) \geq 0$,*

$$\mathrm{Tr}\, f(Z^*AZ) \leq \mathrm{Tr}\, Z^*f(A)Z.$$

(2) *If f is convex and $f(0) \leq 0$,*

$$\mathrm{Tr}\, f(Z^*AZ) \geq \mathrm{Tr}\, Z^*f(A)Z.$$

Example 3.6. Here, contrary to the Brown-Kosaki trace inequalities (Corollary 2.3(1)), the assumption $A \geq 0$ is essential. For instance, in the convex case, consider $f(t) = t_+$,

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

Then, we have $\mathrm{Tr}\, f(Z^*AZ) = 3 < 5 = \mathrm{Tr}\, Z^*f(A)Z$. Of course, the assumption $A \geq 0$ is also essential in Lemma 3.4.

We turn to the proof of Theorem 3.5.

Proof. Of course, assertions (1) and (2) are equivalent. Let us prove (2). Since Z is expansive we may assume that $f(0) = 0$. By a limit argument we may then assume that

$$f(t) = \lambda t + \sum_{i=1}^{m} \alpha_i (t - \beta_i)_+$$

for a real λ and some nonnegative reals $\{\alpha_i\}_{i=1}^m$ and $\{\beta_i\}_{i=1}^m$. The result then follows from the linearity of the trace and Lemma 3.4. $\square$

In order to extend Theorem 3.5(2) to all unitarily invariant norms, i.e. those norms $\|\cdot\|$ such that $\|UXV\| = \|X\|$ for all operators X and all unitaries U and V, we need a simple lemma. A family of positive operators $\{A_i\}_{i=1}^m$ is said to be *monotone* if there exists a positive operator Z and a family of nondecreasing nonnegative functions $\{f_i\}_{i=1}^m$ such that $f_i(Z) = A_i$, $i = 1, \ldots m$.

Lemma 3.7. *Let $\{A_i\}_{i=1}^m$ be a monotone family of positive operators and let $\{U_i\}_{i=1}^m$ be a family of unitary operators. Then, for every unitarily invariant norm $\|\cdot\|$, we have*

$$\|\sum_i U_i A_i U_i^*\| \le \|\sum_i A_i\|.$$

Proof. By the Ky Fan dominance principle, it suffices to consider the Ky Fan k-norms $\|\cdot\|_{(k)}$ [1, pp. 92-3]. There exists a rank k projection E such that

$$\|\sum_i U_i A_i U_i^*\|_{(k)} = \sum_i \operatorname{Tr} U_i A_i U_i^* E \le \sum_i \|A_i\|_{(k)} = \|\sum_i A_i\|_{(k)}$$

where the inequality comes from the maximal characterization of the Ky Fan norms and the last equality from the monotony of the family $\{A_i\}$. $\square$

Proposition 3.8. *Let A be a positive operator and Z be an expansive operator. Assume that f is a nonnegative convex function defined on $[0, b]$, $b \ge \|Z^*AZ\|_\infty$. Assume also that $f(0) = 0$. Then, for every unitarily invariant norm $\|\cdot\|$,*

$$\|f(Z^*AZ)\| \ge \|Z^*f(A)Z\|.$$

Proof. It suffices to consider the case when

$$f(t) = \lambda t + \sum_{i=1}^{m} \alpha_i (t - \beta_i)_+$$

for some nonnegative reals λ, $\{\alpha_i\}_{i=1}^m$ and $\{\beta_i\}_{i=1}^m$. By Lemma 3.4, we have

$$\begin{aligned} Z^*f(A)Z &= \lambda Z^*AZ + \sum_i Z^*\alpha_i(A - \beta_i I)_+ Z \\ &\le \lambda Z^*AZ + \sum_i U_i \alpha_i (Z^*AZ - \beta_i I)_+ U_i^* \end{aligned}$$

for some unitary operators $\{U_i\}_{i=1}^m$. Since λZ^*AZ and $\{\alpha_i(Z^*AZ-\beta_i I)_+\}_{i=1}^m$ form a monotone family, Lemma 3.7 completes the proof. $\square$

Theorem 3.9. *Let A be a positive operator, let Z be an expansive operator and let $f : [0,\infty) \longrightarrow [0,\infty)$ be a nondecreasing concave function. Then,*

$$\|f(Z^*AZ)\|_\infty \le \|Z^*f(A)Z\|_\infty.$$

Proof. Since Z is expansive we may assume $f(0) = 0$. By a continuity argument we may assume that f is onto. Let g be the reciprocal function. Note that g is convex and $g(0) = 0$. By Proposition 3.8,

$$\|g(Z^*AZ)\|_\infty \ge \|Z^*g(A)Z\|_\infty.$$

Hence

$$f(\|g(Z^*AZ)\|_\infty) \ge f(\|Z^*g(A)Z\|_\infty).$$

Equivalently,

$$\|Z^*AZ\|_\infty \ge \|f(Z^*g(A)Z)\|_\infty,$$

so, letting $B = g(A)$,

$$\|Z^*f(B)Z\|_\infty \ge \|f(Z^*BZ)\|_\infty,$$

proving the result because $A \longrightarrow g(A)$ is onto. $\square$

Our next result is a straightforward application of Theorem 2.2.

Corollary 3.10. *Let A be a positive operator and Z be an expansive operator. Assume that f is a nonnegative function defined on $[0,b]$, $b \ge \|Z^*AZ\|_\infty$. Then:*

(1) *If f is concave nondecreasing,*

$$\det f(Z^*AZ) \le \det Z^*f(A)Z.$$

(2) *If f is convex increasing and $f(0) = 0$,*

$$\det f(Z^*AZ) \ge \det Z^*f(A)Z.$$

Proof. For instance, consider the concave case. By Theorem 2.2, there exists a unitary operator U such that $Z^{*-1}f(Z^*AZ)Z^{-1} \le Uf(A)U^*$; hence the result follows. $\square$

We note the following fact about operator convex functions:

Proposition 3.11. *Let $f : [0,\infty) \longrightarrow [0,\infty)$ be a one to one continuous function with $f(0) = 0$ and $f(\infty) = \infty$. The following statements are equivalent:*

(i) *f(t) is operator convex.*

(ii) *1/f(1/t) is operator convex.*

Proof. Since the map $f(t) \longrightarrow \Psi(f)(t) = 1/f(1/t)$ is an involution on the set of all one to one continuous functions f on $[0,\infty)$ with $f(0)=0$ and $f(\infty)=\infty$, it suffices to check that (i) $\Rightarrow$ (ii). But, by the Hansen-Pedersen inequality [7], (i) is equivalent to

$$f(Z^*AZ) \le Z^* f(A) Z \tag{9}$$

for all $A \ge 0$ and all contractions Z. By a limit argument, it suffices to require (9) when both A and Z are invertible. Then, as $t \longrightarrow t^{-1}$ is operator decreasing, (9) can be written

$$f^{-1}(Z^*AZ) \ge Z^{-1} f^{-1}(A) Z^{*-1},$$

or

$$f^{-1}(A) \le Z f^{-1}(Z^*AZ) Z^*,$$

but this is the same as saying that (9) holds for $\Psi(f)$, therefore $\Psi(f)$ is operator convex. $\square$

We wish to sketch another proof of Proposition 3.11. By a result of Hansen and Pedersen [6], for a continuous function f on $[0,\infty)$, the following conditions are equivalent:

(i) $f(0) \le 0$ and f is operator convex.

(ii) $t \longrightarrow f(t)/t$ is operator monotone on $(0,\infty)$.

Using the operator monotony of $t \longrightarrow 1/t$ on $(0,\infty)$, we note that if $f(t)$ satisfies to (ii), then so does $1/f(1/t)$. This proves Proposition 3.11.

Remark 3.12. Theorem 3.9 remains true for infinte dimensional spaces. Indeed, Lemma 3.7 and Proposition 3.8 with the operator norm are valid in the infinite dimensional setting. Does Theorem 3.5 extend to all nonneggative concave functions on $[0,b]$ and/or all unitarily invariant norms ?

4. Unitarily invariant norms and orthogonal decompositions

This section deals with unitarily invariant norms on the space of operators on a finite dimensional space, or on operator ideals on an infinite dimensional, separable space (see Simon's book [10] for this notion – Simon uses the terminology of symmetric norms).

Let $p \ge 1$ and let $\|\cdot\|_\wedge$ be a unitarily invariant norm affiliated to another one $\|\cdot\|$ via

$$\|X\|_\wedge = \||X|^p\|^{1/p}$$

for all operators X. We then say that $\|\cdot\|_\wedge$ is a p-induced norm and that its dual norm $\|\cdot\|_\vee$ is a dual p-induced norm. We note that for $p=2$, p-induced norms are the quadratic norms or Q-norms $\|\cdot\|_Q$ [1, p. 95]:

$$\|X\|_\wedge = \|X^*X\|^{1/2}.$$

for all operators X. Clearly, if $\|\cdot\|_\wedge$ is p-induced, it is also r-induced, $1 < r < p$. The Schatten p-norms are p-induced norms affiliated to the trace norm. For $p \geq 2$, the Schatten p-norms are special cases of quadratic norms.

Let $\{E_i\}_{i=1}^k$ be a total sequence of mutually orthogonal projections on a space $\mathcal{H}$, that is

$$\mathcal{H} = \bigoplus_{i=1}^k E_i(\mathcal{H}).$$

Given an operator X on $\mathcal{H}$ we wish to compare X with the XE_i's and with the E_iXE_i's. In other words, we wish to compare X with its restrictions and with its compressions on the subspaces $E_i(\mathcal{H})$, $1 \leq i \leq k$. We have the following result:

Proposition 4.1. *Let $\|\cdot\|_\wedge$ be a p-induced norm, $1 \leq p \leq 2$, and let $\|\cdot\|_\vee$ be its dual norm. Then, for every operator X and every total sequence of projections $\{E_i\}_{i=1}^k$, we have*

$$\|X\|_\wedge \leq \Big(\sum_i \|XE_i\|_\wedge^p\Big)^{1/p}$$

and

$$\|X\|_\vee \geq \Big(\sum_i \|XE_i\|_\vee^q\Big)^{1/q}$$

with $1/q = 1 - 1/p$.

Proof. If $\|\cdot\|_\wedge$ is affiliated to $\|\cdot\|$, we have

$$\|X\|_\wedge^p = \||X|^p\| = \||X|^{p/2}(\sum_i E_i)|X|^{p/2}\| \leq \sum_i \||X|^{p/2}E_i|X|^{p/2}\| = \sum_i \|E_i|X|^pE_i\|.$$

We then note, by Theorem 1.1 (or by [6]) and the concavity of $t \longrightarrow t^{p/2}$, that

$$\|E_i|X|^pE_i\| \leq \|(E_i|X|^2E_i)^{p/2}\| = \|XE_i\|_\wedge^p.$$

To prove the case $\|\cdot\|_\vee$ we proceed by duality. For $i = 1, \ldots, k$ we consider the spaces of operators

$$C_i = \{\, X \;:\; X = XE_i \,\}.$$

Then, endowed with the norm $\|\cdot\|_\wedge$, C_i becomes a normed space whose dual is

$$R_i = \{\, Y \;:\; Y = E_iY \,\},$$

the duality being implemented by the trace:

$$\langle X, Y\rangle_i = \operatorname{Tr} XY = \operatorname{Tr} XE_iY.$$

Now, observe that the l^p-sum $C = \bigoplus_i C_i$ is canonically isomorphic to the space of all operators X on $\mathcal{H}$ equipped with the norm

$$\|X\|_{[\wedge]} = \big(\sum_i \|XE_i\|_{\wedge}^p\big)^{1/p}.$$

The dual of C is then the l^q-sum $R = \bigoplus_i R_i$ which is canonically isomorphic to the space of all operators Y on $\mathcal{H}$ equipped with the norm

$$\|Y\|_{[\vee]} = \big(\sum_i \|E_iY\|_{\vee}^q\big)^{1/q},$$

and the duality $\langle\cdot,\cdot\rangle$ between C and R is implemented by the trace since

$$\langle X, Y\rangle = \sum_i \langle XE_i, E_iY\rangle_i = \sum_i \operatorname{Tr} XE_iY = \operatorname{Tr} XY.$$

Consequently, we have for each operator Y on $\mathcal{H}$ an operator X on $\mathcal{H}$ such that

$$\|X\|_{[\wedge]} = 1 \quad \text{and} \quad \operatorname{Tr} XY = \|Y\|_{[\vee]}.$$

Hence,

$$\begin{aligned}\operatorname{Tr} XY &= \|X\|_{[\wedge]}\|Y\|_{[\vee]} \\ &= \big(\sum_i \|XE_i\|_{\wedge}^p\big)^{1/p}\big(\sum_i \|E_iY\|_{\vee}^q\big)^{1/q} \\ &\geq \|X\|_{\wedge}\big(\sum_i \|E_iY\|_{\vee}^q\big)^{1/q}. \qquad (*)\end{aligned}$$

Besides,

$$\operatorname{Tr} XY \leq \|X\|_{\wedge}\|Y\|_{\vee}. \qquad (**)$$

Combining (*) with (**),

$$\|Y\|_{\vee} \geq \big(\sum_i \|E_iY\|_{\vee}^q\big)^{1/q}$$

for every operator Y, which is equivalent to the second assertion of the Proposition. □

Let us consider some special cases and consequences of the previous result. We have

$$\|X\|_Q \leq \big(\sum_i \|XE_i\|_Q^2\big)^{1/2} \qquad (10)$$

for every operator X and all quadratic norms $\|\cdot\|_Q$. We also have

$$\|X\|_P \geq \big(\sum_i \|XE_i\|_P^2\big)^{1/2} \qquad (11)$$

for every operator X and all dual quadratic norms $\|\cdot\|_P$, a class of unitarily invariant norms containing the Schatten p-norms $\|\cdot\|_p$, $1 \leq p \leq 2$.

Applying (10) to Schatten q-norms, $2 \le q \le \infty$, and using Holder inequality we obtain

$$\|X\|_q \le k^{1/2-1/q}\big(\sum_i \|XE_i\|_q^q\big)^{1/q}$$

Similarly, (11) implies

$$\|X\|_p \ge k^{1/p-1/2}\big(\sum_i \|XE_i\|_p^p\big)^{1/p}$$

for Schatten p-norms, $1 \le p \le 2$. Proposition 4.1 also entails the following results essentially due to Gohberg and Markus [4]:

$$\big(\sum_i \|XE_i\|_q^q\big)^{1/q} \le \|X\|_q\,, \quad 2 \le q \le \infty,$$

and,

$$\big(\sum_i \|XE_i\|_p^p\big)^{1/p} \ge \|X\|_p\,, \quad 1 \le p \le 2.$$

Assuming that X acts on an n-dimensional space and given a matrix representation $(x_{i,j})$ of X we then derive the following wellknown inequalities [5] (see also [11, p. 50]):

$$\|X\|_p \le \big(\sum_{i,j} |x_{i,j}|^p\big)^{1/p}, \quad 1 \le p \le 2$$

and

$$\|X\|_q \ge \big(\sum_{i,j} |x_{i,j}|^q\big)^{1/q}, \quad 2 \le q \le \infty.$$

Given a total sequence of projections $\{E_i\}_{i=1}^k$ we say that the associated pinching

$$A \longrightarrow \mathcal{P}(A) = \sum_i E_iAE_i.$$

is a k-pinching.The pinching inequality states that pinchings are reducing for invariant norms. In the converse direction we note the following simple fact:

Proposition 4.2. *Let A be a positive operator and let $\mathcal{P}$ be a k-pinching. Then,*

$$A \le k\mathcal{P}(A)$$

Proof. We must show that, for all $h \in \mathcal{H}$,

$$\langle h, Ah\rangle \le k\sum_i \langle h, E_iA_ih\rangle$$

or, equivalently, that

$$\|A^{1/2}h\|^2 \le k\sum_i \|A^{1/2}E_ih\|^2. \tag{12}$$

But, we have

$$\|A^{1/2}h\| \le \sum_i \|A^{1/2}E_i h\|,$$

and, by convexity of $t \longrightarrow t^2$,

$$\left(\sum_i \frac{\|A^{1/2}E_i h\|}{k}\right)^2 \le \frac{\sum_i \|A^{1/2}E_i h\|^2}{k}$$

so that

$$\frac{1}{k^2}\|A^{1/2}h\|^2 \le \frac{1}{k}\sum_i \|A^{1/2}E_i h\|^2$$

hence (12) holds. $\square$

5. Addenda

5.1 Comparison of $f(A+B)$ and $f(A)+f(B)$

There exist several inequalities involving $f(A+B)$ and $f(A)+f(B)$ where A, B are Hermitians and f is a function with special properties. We wish to state and prove one of the most basic results in this direction which can be derived from a more general result due to Rotfel'd (see [1, p. 97]). The simple proof given here is inspired by that of Theorem 3.5.

Proposition 5.1. (Rotfel'd) *Let A, B be positive operators.*

(1) *If f is a convex nonnegative function on $[0,\infty)$ with $f(0) \le 0$, then*

$$\operatorname{Tr} f(A+B) \ge \operatorname{Tr} f(A) + \operatorname{Tr} f(B).$$

(2) *If f is a concave nonnegative function on $[0,\infty)$, then*

$$\operatorname{Tr} g(A+B) \le \operatorname{Tr} g(A) + \operatorname{Tr} g(B).$$

Proof. By limit arguments, we may assume that we are in the finite dimensional setting. Since, on any compact interval $[a,b]$, $a>0$, we may write $g(x) = \lambda x - f(x) + \mu$ for some scalar $\lambda, \mu \ge 0$ and some convex function f with $f(0)=0$, it suffices to consider the convex case. Clearly we may assume $f(0)=0$. Then, f can be uniformly approximated, on any compact interval, by a positive combination of functions $f_\alpha(x) = \max\{0, x-\alpha\} = (x-\alpha)_+$, $\alpha > 0$.

Therefore, still using the notation S_+ for the positive part of the Hermitian operator S, we need only to show that

$$\operatorname{Tr}(A+B-\alpha)_+ \ge \operatorname{Tr}(A-\alpha)_+ + \operatorname{Tr}(B-\alpha)_+.$$

To this end, consider an orthonormal basis $\{e_i\}_{i=1}^n$ of eigenvectors for $A+B$. We note that:

(a) If $\langle e_i, (A+B-\alpha)_+ e_i\rangle = 0$, then $\langle e_i, (A+B-\alpha)_+ e_i\rangle \leq \alpha$ so that we also have $\langle e_i, (A-\alpha)_+ e_i\rangle = \langle e_i, (B-\alpha)_+ e_i\rangle = 0$.

(b) If $\langle e_i, (A+B-\alpha)_+ e_i\rangle > 0$, then we may write

$$\langle e_i, (A+B-\alpha)_+ e_i\rangle = \langle e_i, Ae_i\rangle - \theta\alpha + \langle e_i, Be_i\rangle - (1-\theta)\alpha$$

for some $0 \leq \theta \leq 1$ chosen in such a way that $\langle e_i, Ae_i\rangle - \theta\alpha \geq 0$ and $\langle e_i, Be_i\rangle - (1-\theta)\alpha \geq 0$. Hence, we have

$$\langle e_i, Ae_i\rangle - \theta\alpha = \langle e_i, (A-\theta\alpha)_+ e_i\rangle \geq \langle e_i, (A-\alpha)_+ e_i\rangle$$

and

$$\langle e_i, Be_i\rangle - (1-\theta)\alpha = \langle e_i, (B-(1-\theta)\alpha)_+ e_i\rangle \geq \langle e_i, (B-\alpha)_+ e_i\rangle$$

by using the simple fact that for two commuting Hermitian operators S, T, $S \leq T \Rightarrow S_+ \leq T_+$.

From (a) and (b) we derive the desired trace inequality by summing over $i = 1, \ldots n$. $\square$

5.2 Extension to the von Neumann and C^*-algebras setting

We do not wish to discuss the possible extensions of our results to the setting of operator algebras. Nevertheless we mention that in [8], versions of trace inequalities of Brown-Kosaki and Hansen-Pedersen are established in the framework of a C^*-algebra endowed with a densely defined, lower semicontinuous trace. We also note that the paper by Nelson [9] and that one by Fack and Kosaki [4] form a good presentation of the theory of noncommutative integration in semifinite von Neumann algebras. In [4], Lemma 4.5 and Proposition 4.6, in the von Neumann algebra setting, state results which are very special cases of Theorem 2.1.

5.3 Proof of fact (b) occuring in the proof of Theorem 1.1.

Recall that we still have to check the following

Lemma 5.2. *Fix a real $r > 0$ and let X and Y be two Hermitian operators such that $\lambda_k(X) \leq \lambda_k(Y)$ and $\lambda_{-k}(X) \leq \lambda_{-k}(Y)$ for all $k = 1, \ldots$. Then there exists a unitary operator U such that*

$$X \leq UYU^* + rI.$$

Proof. Let $E(\lambda)$ be the strongly right continuous spectral measure of X. Let $F(\lambda)$ be the strongly left continuous spectral measure of X. Let a be the lower bound of the essential spectrum of X and b be its upper bound. Set

$$X_- = X_{E(a)} \quad \text{et} \quad X_+ = X_{(I-F(b))}$$

where the projections are identified to the corresponding subspaces. Thus, there is a direct sum

$$\mathcal{H} = \mathcal{H}_- \bigoplus \mathcal{H}_0 \bigoplus \mathcal{H}_+$$

such that

$$X = X_- \oplus X_0 \oplus X_+.$$

Similarly, there is another direct sum

$$\mathcal{H} = \mathcal{G}_- \bigoplus \mathcal{G}_0 \bigoplus \mathcal{G}_+$$

for which

$$Y = Y_- \oplus Y_0 \oplus Y_+.$$

Let us consider the case when

$$\dim \mathcal{H}_- = \dim \mathcal{H}_0 = \dim \mathcal{H}_+ = \dim \mathcal{G}_- = \dim \mathcal{G}_0 = \dim \mathcal{G}_+, = \infty,$$

the other cases being similar. By assumptions on the λ_+'s and λ_-'s of X and Y, there exist onto isometries

$$V : \mathcal{H}_0 \bigoplus \mathcal{H}_+ \longrightarrow \mathcal{G}_+$$

and

$$W : \mathcal{H}_- \longrightarrow \mathcal{G}_- \bigoplus \mathcal{G}_0$$

such that

$$V(X_0 \oplus X_+)V^{-1} \leq Y_+ + r$$

and

$$W(X_-)W^{-1} \leq (Y \oplus Y_0) + r;$$

Hence, we can take $U = W \oplus V$. Let us, for instance, check the existence of V. We may write

$$\begin{aligned} Y_+ &= \mathrm{diag}(\lambda_k(Y)) \\ &= \mathrm{diag}_{1\leq k\leq p}(\lambda_k(Y)) \oplus \mathrm{diag}_{k\in J}(\lambda_k(Y)) \oplus \mathrm{diag}_{k\in L}(\lambda_k(Y)) \end{aligned}$$

wherein $J \cup L = \{p+1, \dots\}$ and $|J| = |L| = \infty$. Choose p large enough to ensure that $\lambda_p(Y) \leq \lim_{k\to\infty} \lambda_k(Y) + r$. We then have one to one isometric operators V_1, V_2 such that

$$\big(\mathrm{diag}_{1\leq k\leq p}(\lambda_k(Y)) \oplus \mathrm{diag}_{k\in J}(\lambda_k(Y))\big) + rI \geq V_1 X_+ V_1^{-1}$$

and

$$\mathrm{diag}_{k\in L}(\lambda_k(Y)) \geq V_2 X_0 V_2^{-1}.$$

Hence, $V = V_1 \oplus V_2$ does the job. $\square$

References

[1] R. Bhatia, Matrix Analysis, Springer, Germany, 1996

[2] L. G. Brown and H. Kosaki, Jensen's inequality is semi-finite von Neumann algebras, J. Operator Theory 23 (1990) 3-19.

[3] C. Davis, A Shwarz inequality for convex operator functions, Proc. Amer. Math. Soc. 8 (1957) 42-44.

[4] T. Fack, H. Kosaki. Generalized s-numbers of τ-measurable operators, *Pacific J. Math.* 123, 269-300 (1986)
[5] I. C. Gohberg and A. S. Markus, Some relations between eigenvalues and matrix elements of linear operators, English trans., Amer. Math. Soc. trans. (2) 52 (1966) 201-216.
[6] F. Hansen, An operator inequality, Math. Ann. 258 (1980) 249-250.
[7] F. Hansen and G. K. Pedersen, Jensen's inequality for operator sand Lowner's Theorem, Math. Ann. 258 (1982) 229-241.
[8] F. Hansen and G. K. Pedersen, Jensen's operator inequality, Bull. London Math. Soc. 35 (2003) 553-564.
[9] E. Nelson, Notes on non-commutative integration, J. Funct. Anal., 15 (1974) 103-116. 1979
[10] B. Simon, Trace Ideals and Their Applications LMS lecture note, 35 Cambridge Univ. Press, Cambridge, 1979
[11] X. Zhan, Matrix Inequalities, LNM 1790, Springer, Berlin, 2002.

Chapter 2

Symmetric norms and reverse inequalities to Davis and Hansen-Pedersen characterizations of operator convexity

Introduction

Capital letters A, $B \dots Z$ mean n-by-n complex matrices, or operators on a finite dimensional Hilbert space $\mathcal{H}$; I stands for the identity. When A is positive semidefinite, resp. positive definite, we write $A \geq 0$, resp. $A > 0$. Let $\| \cdot \|$ be a general symmetric (or unitarily invariant) norm, i.e. $\|UAV\| = \|A\|$ for all A and all unitaries U, V. If A and B are such that the product AB is normal, then a classical inequality claims [1, p. 253]

$$\|AB\| \leq \|BA\| \tag{1}$$

Section 1 presents a generalization of (1) when $AB \geq 0$. Then, for $Z > 0$,

$$\|ZAB\| \leq \frac{a+b}{2\sqrt{ab}} \|BZA\| \tag{2}$$

where a, b are the extremal eigenvalues of Z. Several sharp inequalities are derived. For instance, if $0 \leq X \leq I$, then

$$XZX \leq \frac{(a+b)^2}{4ab} Z.$$

Another example concerns compressions $Z_{\mathcal{E}}$ of Z onto subspaces $\mathcal{E} \subset \mathcal{H}$,

$$(Z_{\mathcal{E}})^{-1} \geq \frac{4ab}{(a+b)^2} (Z^{-1})_{\mathcal{E}}. \tag{3}$$

This Kantorovich type inequality is due to Mond-Pecaric. In Section 2 we extend (3) to all operator convex functions $f : [0, \infty) \longrightarrow [0, \infty)$. Such inequalities are reverse inequalities to Davis' characterization of operator convexity via compressions. Equivalently, we show that, given any isometric column of operators

$\{A_i\}_{i=1}^m$, i.e. $\sum A_i^* A_i = I$, we have

$$f(\sum A_i^* Z_i A_i) \geq \frac{4ab}{(a+b)^2} \sum A_i^* f(Z_i) A_i.$$

This is a reverse inequality to the Hansen-Pedersen inequality.

1. Norms inequalities

Lemma 1.1. *Let $Z > 0$ with extremal eigenvalues a and b. Then, for every norm one vector h,*

$$\|Zh\| \leq \frac{a+b}{2\sqrt{ab}} \langle h, Zh \rangle.$$

Proof. Let $\mathcal{E}$ be any subspace of $\mathcal{H}$ and let a' and b' be the extremal eigenvalues of $Z_{\mathcal{E}}$. Then $a \geq a' \geq b' \geq b$ and, setting $t = \sqrt{a/b}$, $t' = \sqrt{a'/b'}$, we have $t \geq t' \geq 1$. Since $t \longrightarrow t + 1/t$ increases on $[1, \infty)$ and

$$\frac{a+b}{2\sqrt{ab}} = \frac{1}{2}\left(t + \frac{1}{t}\right), \qquad \frac{a'+b'}{2\sqrt{a'b'}} = \frac{1}{2}\left(t' + \frac{1}{t'}\right),$$

we infer

$$\frac{a+b}{2\sqrt{ab}} \geq \frac{a'+b'}{2\sqrt{a'b'}}.$$

Therefore, it suffices to prove the lemmma for $Z_{\mathcal{E}}$ with $\mathcal{E} = \mathrm{span}\{h, Zh\}$. Hence, we may assume $\dim \mathcal{H} = 2$, $Z = ae_1 \otimes e_1 + be_2 \otimes e_2$ and $h = xe_1 + (\sqrt{1-x^2})e_2$. Setting $x^2 = y$ we have

$$\frac{||Zh||}{\langle h, Zh \rangle} = \frac{\sqrt{a^2 y + b^2(1-y)}}{ay + b(1-y)}.$$

The righ hand side attains its maximum on $[0, 1]$ at $y = b/(a+b)$, and then

$$\frac{||Zh||}{\langle h, Zh \rangle} = \frac{a+b}{2\sqrt{ab}}$$

proving the lemma. $\square$

Theorem 1.2. *Let A, B such that $AB \geq 0$. Let $Z > 0$ with extremal eigenvalues a and b. Then, for every symmetric norm, the following sharp inequality holds*

$$\|ZAB\| \leq \frac{a+b}{2\sqrt{ab}} \|BZA\|.$$

Proof. For the sharpness see Remark 1.9 below.

It suffices to consider the Fan k-norms $\|\cdot\|_{(k)}$ [1, p. 93]. Fix k and let $\|\cdot\|_1$ denote the trace-norm. There exist two rank k projections E and F such that

$$\begin{aligned}\|ZAB\|_{(k)} &= \|ZABE\|_1 \\ &= \|Z(AB)^{1/2}F(AB)^{1/2}E\|_1 \\ &\leq \|Z(AB)^{1/2}F(AB)^{1/2}\|_1.\end{aligned}$$

Consider the canonical decomposition

$$(AB)^{1/2}F(AB)^{1/2} = \sum_{j=1}^{k} c_j\, h_j \otimes h_j$$

in which $\{h_j\}_{j=1}^k$ is an orthonormal system and $\{h_j \otimes h_j\}_{j=1}^k$ are the associated rank one projections. We have, using the trianle inequality and then the above lemma,

$$\begin{aligned}\|Z(AB)^{1/2}F(AB)^{1/2}\|_1 &\leq \sum_{j=1}^{k} c_j \|Zh_j \otimes h_j\|_1 \\ &= \sum_{j=1}^{k} c_j \|Zh_j\| \\ &\leq \frac{a+b}{2\sqrt{ab}} \sum_{j=1}^{k} c_j \langle h_j, Zh_j\rangle \\ &= \frac{a+b}{2\sqrt{ab}} \operatorname{Tr}(AB)^{1/2}F(AB)^{1/2}Z.\end{aligned}$$

Next, there exists a rank k projection G such that

$$\begin{aligned}\frac{a+b}{2\sqrt{ab}} \operatorname{Tr}(AB)^{1/2}F(AB)^{1/2}Z &= \frac{a+b}{2\sqrt{ab}} \operatorname{Tr}(AB)^{1/2}F(AB)^{1/2}ZG \\ &\leq \frac{a+b}{2\sqrt{ab}} \operatorname{Tr} GZ^{1/2}ABZ^{1/2}G \\ &\leq \frac{a+b}{2\sqrt{ab}} \|Z^{1/2}ABZ^{1/2}\|_{(k)} \\ &\leq \frac{a+b}{2\sqrt{ab}} \|BZA\|_{(k)}\end{aligned}$$

where at the last step we used the basic inequality (1). $\square$

One may ask wether our theorem can be improved to singular values inequalities. This is not possible as it is shown by the next example:

Take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \ B = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \ Z = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}.$$

Then the largest and smallest eigenvalues of Z are $a = 8$ and $b = 2$, so

$$\frac{a+b}{2\sqrt{ab}} = 1.25.$$

Besides, $\mu_2(ZAB) = 8$ and $\mu_2(AZB) = 4.604$, and since $4.604 \times 1.25 = 5.755 < 8$, Theorem 1.1 can not be extended to singular values inequalities.

We denote by Sing(X) the sequence of the singular values of X, arranged in decreasing order and counted with their multiplicities. Similarily, when X has only real eigenvalues, $\mathrm{Eig}(X)$ stands for the sequence of X's eigenvalues. Given two sequences of real numbers $\{a_j\}_{j=1}^n$ and $\{b_j\}_{j=1}^n$, we use the notation $\{a_j\}_{j=1}^n \prec_w \{b_j\}_{j=1}^n$ for weak-majorisation, that is $\sum_{j=1}^k a_j \le \sum_{j=1}^k b_j$, $k = 1, \ldots$.

A straightforward application of Theorem 1.2 is:

Corollary 1.3. *Let $A \ge 0$ and let $Z > 0$ with extremal eigenvalues a and b. Then,*

$$\mathrm{Sing}(AZ) \prec_w \frac{a+b}{2\sqrt{ab}}\, \mathrm{Eig}(AZ).$$

Proof. For each Fan norms, replace A and B by $A^{1/2}$ in Theorem 1.2. $\square$

Special cases of the above corollary are:

Corollary 1.4. *Let $A \ge 0$ and let $Z > 0$ with extremal eigenvalues a and b. Then,*

$$\|AZ\|_\infty \le \frac{a+b}{2\sqrt{ab}}\, \rho(AZ)$$

and

$$\|AZ\|_1 \le \frac{a+b}{2\sqrt{ab}}\, \mathrm{Tr}\, AZ.$$

Here, $\|\cdot\|_\infty$ stands for the standard operator norm and $\rho(\cdot)$ for the spectral radius.

From the preceding result, one may derive an interesting operator inequality:

Corollary 1.5. *Let $0 \le A \le I$ and let $Z > 0$ with extremal eigenvalues a and b. Then,*

$$AZA \le \frac{(a+b)^2}{4ab}\, Z.$$

Proof. The claim is equivalent to the operator nom inequalities

$$\|Z^{-1/2}AZAZ^{-1/2}\|_\infty \leq \frac{(a+b)^2}{4ab}$$

or

$$\|Z^{-1/2}AZ^{1/2}\|_\infty \leq \frac{a+b}{2\sqrt{ab}}.$$

But the previous corollary entails

$$\begin{aligned}\|Z^{-1/2}AZ^{1/2}\|_\infty &= \|Z^{-1/2}AZ^{-1/2}Z\|_\infty \\ &\leq \frac{a+b}{2\sqrt{ab}}\,\rho(Z^{-1/2}AZ^{-1/2}Z) \\ &= \frac{a+b}{2\sqrt{ab}}\,\|A\|_\infty \\ &\leq \frac{a+b}{2\sqrt{ab}},\end{aligned}$$

hence, the result holds. □

A special case of Corollary 1.5 gives a comparison bewtween Z and the compression EZE, for an arbitrary projection E.

Corollary 1.6. *Let $Z > 0$ with extremal eigenvalues a and b and let E be any projection. Then,*

$$EZE \leq \frac{(a+b)^2}{4ab}\,Z.$$

We may then derive a classical inequality:

Corollary 1.7. (Kantorovich) *Let $Z > 0$ with extremal eigenvalues a and b and let h be any norm one vector. Then,*

$$\langle h, Zh\rangle\langle h, Z^{-1}h\rangle \leq \frac{(a+b)^2}{4ab}.$$

Proof. Rephrase Corollary 1.6 as

$$\|Z^{-1/2}EZEZ^{-1/2}\|_\infty \leq \frac{(a+b)^2}{4ab}$$

and take $E = h \otimes h$. □

A classical inequality in Matrix theory, for positive definite matrices, claims that "The inverse of a principal submatrix is less than or equal to the corresponding submatrix of the inverse" [6, p. 474]. In terms of compressions, this means

$$(Z_{\mathcal{E}})^{-1} \leq (Z^{-1})_{\mathcal{E}} \tag{4}$$

for every subspace $\mathcal{E}$ and every $Z > 0$. Corollary 1.6 entails a reverse inequality, first proved by B. Mond and J.E. Pecaric [7]:

Corollary 1.8. (Mond-Pecaric) *Let $Z > 0$ with extremal eigenvalues a and b. Then, for every subspace $\mathcal{E}$,*

$$(Z_{\mathcal{E}})^{-1} \geq \frac{4ab}{(a+b)^2}\,(Z^{-1})_{\mathcal{E}}.$$

Note that Corollary 1.8 implies Corollary 1.7.

Proof. Let E be the projection onto $\mathcal{E}$. By Corollary 1.6, for every $r > 0$, there exists $x > 0$ such that

$$EZE + xE^{\perp} \leq \frac{(a+b)^2}{4ab}\,(Z + rI).$$

Since $t \longrightarrow -1/t$ is operator monotone we deduce

$$(EZE + xE^{\perp})^{-1} \geq \frac{4ab}{(a+b)^2}\,(Z + rI)^{-1}$$

so that

$$(Z_{\mathcal{E}})^{-1} \geq \frac{4ab}{(a+b)^2}\,\{(Z + rI)^{-1}\}_{\mathcal{E}}$$

and the result follows by letting $r \longrightarrow 0$. $\square$

Remark 1.9. All the previous inequalities are sharp. Indeed, let h be a norm one vector for which equality occurs in Lemma 1.1. Then, replacing A, B, E by $h \otimes h$ and $\mathcal{E}$ by span$\{h\}$ in the above statements, yields equality cases.

Remark 1.10. As for a standard proof of (1) [1, p. 253], it is tempting to first prove Theorem 1.2 for the operator norm and then to use an antisymmetric tensor product argument to derive the general case. Such an approach seems impossible. Indeed if a_k and b_k are the extremal eigenvalues of $\wedge^k(Z)$, then the relation

$$\frac{(a_k + b_k)^2}{4a_k b_k} \leq \left(\frac{(a+b)^2}{4ab}\right)^k$$

is not true in general.

The next result states a companion inequality to Corollary 1.8.

Proposition 1.11. *Let $Z > 0$ with extremal eigenvalues a and b and let $1 \leq p \leq 2$. Then, for every subspace $\mathcal{E}$,*

$$(Z_{\mathcal{E}})^p \geq \frac{4ab}{(a+b)^2}\,(Z^p)_{\mathcal{E}}.$$

Proof. Let E be the projection onto $\mathcal{E}$. For any norm one vector $h \in \mathcal{E}$, Lemma 1.1 implies

$$\begin{aligned} \langle h, (Z^p)_{\mathcal{E}} h\rangle &= \langle h, EZ^pEh\rangle \\ &= \|Z^{p/2}h\|^2 \\ &\le \frac{(a+b)^2}{4ab}\,\langle h, Z^{p/2}h\rangle^2. \end{aligned}$$

Then, using the concavity of $t \longrightarrow t^{p/2}$ and next the convexity of $t \longrightarrow t^p$, we deduce

$$\begin{aligned} \langle h, (Z^p)_{\mathcal{E}} h\rangle &\le \frac{(a+b)^2}{4ab}\,\langle h, Zh\rangle^p \\ &= \frac{(a+b)^2}{4ab}\,\langle h, EZEh\rangle^p \\ &\le \frac{(a+b)^2}{4ab}\,\langle h, (Z_{\mathcal{E}})^p h\rangle. \end{aligned}$$

and the proof is complete. $\square$

2. Operator convexity

Davis' characterization of operator convexity [2] claims: f is operator convex on $[a, b]$ if and only if for every subspace $\mathcal{E}$ and every Hermitian Z with spectrum in $[a, b]$,

$$f(Z_{\mathcal{E}}) \le (f(Z))_{\mathcal{E}} \tag{D}$$

Since $t \longrightarrow t^p$, $1 \le p \le 2$ and $t \longrightarrow 1/t$ are operator convex on $(0, \infty)$, both Proposition 1.11 and Corollary 1.8 are reverse inequalities to Davis' characterization of operator convexity.

Proposition 1.11 is a special case of the next theorem.

Theorem 2.1. *Let $f : [0, \infty) \longrightarrow [0, \infty)$ be operator convex and let $Z > 0$ with extremal eigenvalues a and b. Then, for every subspace $\mathcal{E}$,*

$$f(Z_{\mathcal{E}}) \ge \frac{4ab}{(a+b)^2}\,(f(Z))_{\mathcal{E}}.$$

Proof. We have the integral representation [xx]

$$f(t) = \alpha + \beta t + \gamma t^2 + \int_0^\infty \frac{\lambda t^2}{\lambda + t}\, d\mu(\lambda),$$

where α, β, γ are nonnegative scalars and μ is a positive finite measure. Therefore, it suffices to prove the result for

$$\alpha + \beta t + \gamma t^2$$

and

$$f_\lambda(t) = \frac{\lambda t^2}{\lambda + t}.$$

The quadratic case is a staightforward application of Proposition 1.11. To prove the f_λ case, note that f_λ is convex meanwhile $f_\lambda^{1/2}$ is convave and then proceed as in the proof of Proposition 1.11. $\square$

Davis' characterization (D) of operator convexity is equivalent to the following result of Hansen-Pedersen [5].

Recall that a family $\{A_i\}_{i=1}^m$ form an isometric column when $\sum A_i^* A_i = I$.

Theorem 2.2. (Hansen-Pedersen) *Let $\{Z_i\}_{i=1}^m$ be Hermitians with spectrum lying in $[a,b]$ and let f be operator convex $[a,b]$. Then, for every isometric column $\{A_i\}_{i=1}^m$,*

$$f(\sum A_i^* Z_i A_i) \leq \sum A_i^* f(Z_i) A_i. \tag{J}$$

(J) is the operator version of Jensen's inequality: operator convex combinations and operator convex functions replace the ordinary ones. As a sthraightforward consequence, we have the following contractive version of (J):

Corollary 2.3. (Hansen-Pedersen) *Let $\{Z_i\}_{i=1}^m$ be Hermitians with spectrum lying in $[a,b]$ and let f be operator convex $[a,b]$ with $0 \in [a,b]$ and $f(0) \leq 0$. Then, for every contraction A,*

$$f(A^* Z A) \leq A^* f(Z) A. \tag{C}$$

Exactly as Theorem 2.1 is a reverse inequality to (D), the following results is a reverse inequality to (J).

Theorem 2.4. *Let $f : [0,\infty) \longrightarrow [0,\infty)$ be operator convex and let $\{Z_i\}_{i=1}^m$ be positive with spectrum lying in $[a,b]$, $a > 0$. Then, for every isometric column $\{A_i\}_{i=1}^m$,*

$$f(\sum A_i^* Z_i A_i) \geq \frac{4ab}{(a+b)^2} \sum A_i^* f(Z_i) A_i.$$

Let us consider a very special case: For every $A, B > 0$ with spectrum lying on $[r, 2r]$, $r > 0$, and for every operator convex $f : [0, \infty) \longrightarrow [0, \infty)$, we have

$$\frac{8}{9} \cdot \frac{f(A) + f(B)}{2} \leq f\left(\frac{A+B}{2}\right) \leq \frac{f(A) + f(B)}{2}.$$

The left inequality gives a negative answer to an approximation problem: Let f be an operator convex function on $[a, b]$, 0¡a¡b, and let $\varepsilon > 0$. Then, in general, there is no operator convex function g on $[0, \infty)$ such that

$$\max_{x \in [a,b]} |f(x) - g(x)| < \varepsilon.$$

From Theorem 2.4 we obtain a reverse inequality to (C):

Corollary 2.5. *Let $f : [0, \infty) \longrightarrow [0, \infty)$ be operator convex and let $Z > 0$ with extremal eigenvalues a and b. Then, for every contraction A,*

$$f(A^*ZA) \geq \frac{4ab}{(a+b)^2} A^* f(Z) A.$$

We turn to the proof of Theorem 2.4 and Corollary 2.5.

Proof. Consider the following operators acting on $\oplus^m \mathcal{H}$,

$$V = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ A_m & 0 & \cdots & 0 \end{pmatrix}, \qquad \tilde{Z} = \begin{pmatrix} Z_1 & & \\ & \ddots & \\ & & Z_m \end{pmatrix}$$

and note that V is a partial isometry. Denoting by $\mathcal{H}$ the first summand of the direct sum $\oplus^m \mathcal{H}$ and by $X : \mathcal{H}$ the restriction of X to $\mathcal{H}$, we observe that

$$f(\sum A_i^* Z_i A_i) = f(V^* \tilde{Z} V) : \mathcal{H} = V^* f(\tilde{Z}_{V(\mathcal{H})}) V : \mathcal{H}.$$

Applying Theorem 2.1 with $\mathcal{E} = V(\mathcal{H})$, we get

$$\begin{aligned} f(\sum A_i^* Z_i A_i) &\geq \frac{4ab}{(a+b)^2} V^* f(\tilde{Z})_{V(\mathcal{H})} V : \mathcal{H} \\ &= \frac{4ab}{(a+b)^2} \sum A_i^* f(Z_i) A_i. \end{aligned}$$

and the proof of Theorem 2.4 is complete. To obtain its corollary, take an operator B such that $A^*A + B^*B = I$. Then, note that, using $f(0) \geq 0$,

$$\begin{aligned} f(A^*ZA) = f(A^*ZA + B^*0B) &\geq \frac{4ab}{(a+b)^2} \{A^* f(Z) A + B^* f(0) B\} \\ &\geq \frac{4ab}{(a+b)^2} A^* f(Z) A \end{aligned}$$

by application of Theorem 2.4. $\square$

Remark 2.6. Corollary 1.8 and Proposition 1.11 for $p = 2$ have been obtained by Mond-Pecaric in the more general form of Theorem 2.4. Note that Proposition 1.11 with $p = 2$ immediately implies Lemma 1.1; hence, we have no pretention of originality in establishing this basic lemma.

Remark 2.7. Hansen-Pedersen first prove the contractive version (C) in [4] and then, some twenty years later [5], prove the more general form (Jo). When proving (C) they noted a technical difficulty to derive (Jo) when $0 \notin [a, b]$]. In fact, this difficulty can be easily overcomed: Note that if (Jo) is valid for every operator convex functions on an interval $[a, b]$, then (Jo) is also valid on every interval of the type $[a + r, b + r]$.

Remark 2.8. (D), (Jo), (C) are equivalent statements. Similarly, Theorems 2.1, 2.4 and Corollary 2.5 are equivalent.

Clearly, the previous results can be suitably restated for operators acting on infinite dimensional spaces.

Inspired by the seminal paper [3], we note that Corollary 2.5 can be stated in a still more general framework. Let $\mathrm{B}(\mathcal{H})$ denote the algebra of all (bounded) linear operators on a separable Hilbert space $\mathcal{H}$.

Corollary 2.9. *Let $\Phi : \mathcal{Z} \longrightarrow \mathrm{B}(\mathcal{H})$ be a positive, linear contraction on a C^*-algebra $\mathcal{Z}$. Let $Z \in \mathcal{Z}$, $Z > 0$ with $\mathrm{Sp}(Z) \subset [a, b]$, $a > 0$. Then, for every operator convex function $f : [0, \infty) \longrightarrow [0, \infty)$,*

$$f \circ \Phi(Z) \geq \frac{4ab}{(a + b)^2} \, \Phi \circ f(Z).$$

Proof. Restricting Φ to the commutative C^*-subalgebra generated by Z, one may suppose Φ completely positive. By Stinepring's dilation Theorem [8], there exist a larger Hilbert space $\mathcal{F} \supset \mathcal{H}$, a linear contraction $A : \mathcal{H} \longrightarrow \mathcal{F}$ and a $*$-homomorphism $\pi : \mathcal{Z} \longrightarrow \mathrm{B}(\mathcal{F})$ such that $\Phi(\cdot) = A^*(\pi(\cdot))_{\mathcal{F}} A$. Therefore

$$\begin{aligned} f \circ \Phi(Z) &= f(A^* \pi(Z) A) \\ &\geq \frac{4ab}{(a + b)^2} \, A^* f(\pi(Z)) A \\ &= \frac{4ab}{(a + b)^2} \, A^* \pi(f(Z)) A \\ &= \frac{4ab}{(a + b)^2} \, \Phi \circ f(Z) \end{aligned}$$

where at the second step we apply Corollary 2.5 which can be extended to this situation by inspection of the proof of Theorem 2.4 and Corollary 2.5. $\square$

References

[1] R. Bhatia, Matrix Analysis, Springer, Germany, 1996.
[2] C. Davis, A Shwarz inequality for convex operator functions, Proc. Amer. Math. Soc. 8 (1957) 42-44.
[3] F. Hansen, An operator inequality, Math. Ann. 246 (1980) 249-259.
[4] F. Hansen and G. K. Pedersen, Jensen's inequality for operator sand Lowner's Theorem, Math. Ann. 258 (1982) 229-241.
[5] F. Hansen and G. K. Pedersen, Jensen's operator inequality, Bull. London Math. Soc. 35 (2003) 553-564.
[6] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, Cambridge, 1985.
[7] B. Mond, J.E. Pecaric, A matrix version of the Ky fan generalization of the Kantorovich inequality, Linear and Multilinear Algebra 36 (1994) 217-221.
[8] W. F. Stinepring, Positive functions on C^*-algebras, Proc. Amer. Math. Soc. 6 (1955) 211-216.

Chapter 3

Commuting dilations and Total dilations

Introduction

The letter $\mathcal{H}$ denotes a separable Hilbert space. $\mathcal{H}$ can be real or complex, finite or infinite dimensional. An operator is a bounded linear operator. An operator Z on $\mathcal{H} \oplus \mathcal{H}$ or $\oplus^k \mathcal{H}$ is a dilation of the operator A on $\mathcal{H}$ if

$$Z = \begin{pmatrix} A & \star \\ \star & \star \end{pmatrix}, \text{ or } Z = \begin{pmatrix} A & \star & \dots \\ \star & \star & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}.$$

How to dilate an operator or a family of operators into operators with special properties is the purpose of Dilation theory. In the above dilations there is a lack of symmetry between the summands in $\mathcal{H} \oplus \mathcal{H}$ or $\oplus^k \mathcal{H}$. We then introduce the following natural notion: An operator Z on $\oplus^k \mathcal{H}$ is said to be a *total dilation* of the operator A on $\mathcal{H}$ if the operator diagonal of Z consists of a repetition of A,

$$Z = \begin{pmatrix} A & \star & \dots \\ \star & A & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}.$$

We express this fact by writing

$$\operatorname{diag}(Z) = \oplus^k A.$$

Let $\{A_j\}_{j=0}^n$ be a family of operators on $\mathcal{H}$ and let $\{Z_j\}_{j=0}^n$ be a family of operators on $\oplus^k \mathcal{H}$. We say that $\{Z_j\}_{j=0}^n$ *totally* dilates $\{A_j\}_{j=0}^n$ if we can write

$$Z_0 = \begin{pmatrix} A_0 & \star & \dots \\ \star & A_0 & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}, \quad \dots \quad Z_n = \begin{pmatrix} A_n & \star & \dots \\ \star & A_n & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}.$$

In the first Section we give several simple examples of total dilation and we show that any operator A on a space of even finite dimension can be written as

$$A = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

for some suitable decomposition of the space.

Section 2 is concerned with the relationship between operator inequalities and (total) dilations. In particular we show that, for positive invertible operators A and B, the condition $A \leq B^{-1}$ is equivalent to the existence of a positive invertible operator Z such that

$$Z = \begin{pmatrix} A & \star \\ \star & A \end{pmatrix} \qquad \textit{and} \qquad Z^{-1} = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

The third section is devoted to other commuting dilations and to open problems.

1. Total dilations: some examples

We give some examples of total dilations:

Example 1.1. A $2n \times 2n$ antisymmetric real matrix A totally dilates the n-dimensional zero operator: with respect to a suitable decomposition

$$A = \begin{pmatrix} 0 & -B^T \\ B & 0 \end{pmatrix}$$

for some symmetric real n-by-n matrix B.

Example 1.2. Any operator A on $\mathcal{H}$ can be totally dilated into a normal operator N on $\mathcal{H} \oplus \mathcal{H}$ by setting

$$N = \begin{pmatrix} A & A^* \\ A^* & A \end{pmatrix}.$$

Example 1.3. Denote by $\tau(A)$ the normalized trace $(1/n)\mathrm{Tr}\, A$ of an operator A on an n-dimensional space. Then the scalar $\tau(A)$ can be totally dilated into A. For an operator acting on a real space and for a hermitian operator the proof is easy. When A is a general operator on a complex space, this result, called Parker's theorem, follows from the Hausdorff-Toeplitz Theorem (see [4, p. 20]).

Example 1.4. Any contraction A on a finite dimensional space $\mathcal{H}$ can be totally dilated into a unitary operator U on $\oplus^k \mathcal{H}$ for any integer k. Indeed by considering the polar decomposition $A = V|A|$, it suffices to construct a total unitary dilation W of $|A|$ and then to take $U = (\oplus^k V) \cdot W$. The construction of a total unitary dilation on $\oplus^k \mathcal{H}$ for a positive contraction X on $\mathcal{H}$ is easy: Let $\{x_j\}_{j=1}^n$ be the eigenvalues of X repeated according to their multiplicities and let $\{U_j\}_{j=1}^n$ be

$k \times k$ unitary matrices such that $\tau(U_j) = x_j$. Example 1.3 and an obvious matrix manipulation show that $\oplus_{j=1}^n U_j$ totally dilates X.

Example 1.5. Let (A, B) be a pair of (positive) operators on $\mathcal{H}$. Then the dilations

$$\begin{pmatrix} A & A \\ A & A \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} B & -B \\ -B & B \end{pmatrix}$$

(are positive and) commute (with a zero product).

Example 1.6. Let (A, B) be a pair of hermitian operators on $\mathcal{H}$ and suppose that A is a strict contraction. For any $\varepsilon > 0$ the Hermitian operators

$$X = \begin{pmatrix} A & \varepsilon B \\ \varepsilon B & A \end{pmatrix} \qquad \text{and} \qquad Y = \begin{pmatrix} B & A/\varepsilon \\ A/\varepsilon & B \end{pmatrix}$$

commute and, when ε is small enough, X is a strict contraction.

In contrast to the previous example, it is not possible in general to dilate a pair of Hermitian strict contractions into a commuting pair of Hermitian strict contractions. I thank Chi-Kwong Li for showing me the following simple example:

Example 1.7. (1) Consider the Hermitian contractions:

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

Then any normal dilation N of $A+iB$ must satisfy $||N||_\infty \geq ||A+iB||_\infty = 2 > \sqrt{2}$. Consequently A and B can not be dilated into commuting Hermitians with norms arbitrarily close to 1.

(2) Now, let P, Q be positive, strict contractions. Then, $P = (A+I)/2$ and $Q = (B+I)/2$ for some Hermitian, strict contractions A, B. Therefore, dilating P, Q into a commuting pair of positive, strict contractions is a problem equivalent to that of dilating A, B into a commuting pair of Hermitian, strict contractions. By the preceding example, it may be impossible.

For an operator A, its numerical range and its numerical angular range are $W(A) = \{\langle h, Ah\rangle \mid \ \|h\| = 1\}$ and $W'(A) = \{\langle h, Ah\rangle \mid \ h \in \mathcal{H}\}$. If Z dilates A then $W(Z) \supset W(A)$ and $W'(Z) \supset W'(A)$. From Examples 1.5 and 1.6 we derive two normal dilation results:

Proposition 1.8. *Let A be an operator on $\mathcal{H}$. Then A can be totally dilated into a normal operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that $W'(Z) = W'(A)$.*

Proof. By the rotation property of the numerical range, we may assume that

$$W'(A) = \{z = x + iy \mid \ 0 \leq x, \ -x \leq ay \leq x\}$$

for some $a > 0$. Therefore, the Cartesian decomposition $A = X + iY$ satisfies to $X \geq 0$ and $-X \leq aY \leq 2X$. Since $X + aY$ and $X - aY$ are both positive, we may totally dilate them into a commuting pair of positive operators, as in Example 1.4. Let S and T denote these dilations, respectively. We observe that $P = (S+T)/2$ and $Q = (S-T)/2a$ are a commuting pair of Hermitian operators, with $P \geq 0$, such that

$$-P \leq aQ \leq P \tag{1}$$

Furthermore, by construction $Z = P + iQ$ totally dilates $A = X + iY$. Hence Z is a normal operator which totally dilates A and (1) then ensures that $W'(Z) = W'(A)$. □

Proposition 1.9. *Let A be an operator on $\mathcal{H}$ and let S be an open strip with $S \supset W(A)$. Then A can be totally dilated into a normal operator Z on $\mathcal{H} \oplus \mathcal{H}$ with $S \supset W(Z)$.*

Proof. By rotation and translation, we may assume that

$$S = \{z = x + iy \mid \quad -1 < x < 1\}.$$

Considering the Cartesian decomposition $A = X + iY$ and applying Example 1.5 to X and Y in place of A and B, respectively, we obtain the proposition. □

Example 1.10. Let $\{A_k\}_{k=1}^n$ be a family of operators on $\mathcal{H}$ and let $\{B_k\}_{k=1}^n$ be the family of operators acting on $\oplus^n \mathcal{H}$ defined by

$$B_k = \begin{pmatrix} A_k & A_{k-1} & \dots \\ A_{k+1} & A_k & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Then $\{B_k\}_{k=1}^n$ is a commuting family which totally dilates $\{A_k\}_{k=1}^n$. (we set $A_0 = A_n$, $A_{-1} = A_{n-1}$, $\dots$.)

In the previous example, the dilations do not preserve properties such as positivity, self-adjointness or normality. Using larger dilations we may preserve these properties:

Proposition 1.11. *Let $\{A_j\}_{j=0}^n$ be operators on a space $\mathcal{H}$. Then there exist operators $\{B_j\}_{j=0}^n$ on $\oplus^k \mathcal{H}$, where $k = 2^n$, such that*

(1) *For $i \neq j$, $B_i B_j = 0$.*

(2) *$\{B_j\}_{j=0}^n$ totally dilates $\{A_j\}_{j=0}^n$.*

(3) *If the A_j's are positive (resp. Hermitian, normal) then the B_j's are of the same type.*

Proof. Given a pair A_0, A_1 of operators, construct

$$S = \begin{pmatrix} A_0 & A_0 \\ A_0 & A_0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} A_1 & -A_1 \\ -A_1 & A_1 \end{pmatrix}.$$

Then $ST = TS = 0$. We then proceed by induction. We have just proved the case of $n = 1$. Assume that the result holds for $n - 1$. Thus we have a family $\mathcal{C} = \{C_j\}_{j=0}^{n-1}$ which totally dilates $\{A_j\}_{j=0}^{n-1}$. Moreover $\mathcal{C}$ acts on a space $\mathcal{G}$, $\dim \mathcal{G} = 2^{n-1} \dim \mathcal{H}$. We dilate A_n to an operator C_n on $\mathcal{G}$ by setting $C_n = A_n \oplus \cdots \oplus A_n$, 2^{n-1} terms . We then consider the operators on $\mathcal{F} = \mathcal{G} \oplus \mathcal{G}$ defined by

$$B_j = \begin{pmatrix} C_j & C_j \\ C_j & C_j \end{pmatrix} \quad \text{for} \quad 0 \le j < n \quad \text{and} \quad B_n = \begin{pmatrix} C_n & -C_n \\ -C_n & C_n \end{pmatrix}.$$

The family $\{B_j\}_{j=0}^{n}$ has the required properties. $\square$

We turn to the main result of this section. If $\mathcal{H}$ is a space with an even finite dimension, we then say that the orthonormal decomposition $\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^\perp$ is a *halving* decomposition whenever $\dim \mathcal{F} = (1/2) \dim \mathcal{H}$.

Theorem 1.12. *Let A be an operator on a space $\mathcal{H}$ with an even finite dimension. Then there exists a halving decomposition $\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^\perp$ for which we have a total dilation*

$$A = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

Proof. Choose a halving decomposition of $\mathcal{H}$ for which we have a matrix representation of ReA of the following form

$$\mathrm{Re}A = \begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix}.$$

Consequently in respect to this decomposition we must have

$$A = \begin{pmatrix} Y & X \\ -X^* & Z \end{pmatrix}.$$

Let $X = U|X|$ and $Y_0 = U^* Y U$. We have

$$\begin{pmatrix} U^* & 0 \\ 0 & I \end{pmatrix} A \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} U^* & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} Y & U|X| \\ -|X|U^* & Z \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} Y_0 & |X| \\ -|X| & Z \end{pmatrix}.$$

Now observe that

$$\frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix} \begin{pmatrix} Y_0 & |X| \\ -|X| & Z \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} = \begin{pmatrix} (Y_0 + Z)/2 & \star \\ \star & (Y_0 + Z)/2 \end{pmatrix}.$$

Thus, using two unitary congruences we have exhibited an operator totally dilated into A. $\square$

We note that, in the very special case of $\dim \mathcal{H} = 2$, theorem 1.12 gives the key step of a classical proof of the Hausdorff-Toeplitz Theorem [6, p. 18].

Remark 1.13. The proof of Theorem 1.12 is easy for a normal operator: consider a representation $A = \begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix}$ and use the unitary conjugation by $\frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ -I & I \end{pmatrix}$. Applying this to X^*X, for an operator X on an even dimensional space, we note that there exists a halving projection E such that XE and $XE^\perp$ have the same singular values (indeed EX^*XE and $E^\perp X^*XE^\perp$ are unitarily equivalent).

Problem 1.14. Does the theorem hold for infinite dimensional spaces ? Let $\mathrm{Tot}(A)$ be the set of operators B which can be totally dilated into A. This set is invariant under unitary congruences. Is it a closed set ? a connected set ? a Riemann measurable set ? Can we extend the theorem to spaces of dimensions kn instead of $2n$?

2. Total Dilations: commuting dilations

The symbol $X_{\mathcal{H}}$ means the compression onto the first summand of an operator X acting on a space of the form $\oplus^k \mathcal{H}$ and an expression such as $f(X)_{\mathcal{H}}$ must be understood as $(f(X))_{\mathcal{H}}$.

Theorem 2.1. *Let A, B be strictly positive operators on $\mathcal{H}$. Then, the condition $A \geq B^{-1}$ is equivalent to the existence of a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that*

$$Z = \begin{pmatrix} A & \star \\ \star & A \end{pmatrix} \qquad \textit{and} \qquad Z^{-1} = \begin{pmatrix} B & \star \\ \star & B \end{pmatrix}.$$

(The stars hold for unspecified entries.)

Proof. Clearly the existence of Z implies $A \geq B^{-1}$ by operator convexity of $t \longrightarrow t^{-1}$. To prove the converse implication we set $C = [I - A^{-1/2}B^{-1}A^{-1/2}]^{1/2}$ and

$$Z = \begin{pmatrix} A & A^{1/2}CA^{1/2} \\ A^{1/2}CA^{1/2} & A \end{pmatrix}.$$

From $A \geq B^{-1}$ we deduce that $I \geq A^{-1/2}B^{-1}A^{-1/2}$; hence C is a contraction. Since $A^{-1/2}B^{-1}A^{-1/2}$ is strictly positive, C is even a strict contraction, i.e. $\|C\| < 1$. Therefore $A > A^{1/2}CA^{1/2}$ so that Z is a strictly positive operator and we may apply the inversion formula for a partitioned matrix (see [5, p. 18]) to obtain Z^{-1} as

$$\begin{pmatrix} [A - A^{1/2}C^2A^{1/2}]^{-1} & A^{-1/2}CA^{1/2}[A^{1/2}C^2A^{1/2} - A]^{-1} \\ [A^{1/2}C^2A^{1/2} - A]^{-1}A^{1/2}CA^{-1/2} & [A - A^{1/2}C^2A^{1/2}]^{-1} \end{pmatrix}$$

that is

$$Z^{-1} = \begin{pmatrix} B & -A^{-1/2}CA^{1/2}B \\ -BA^{1/2}CA^{-1/2} & B \end{pmatrix}$$

and the proof is complete. $\square$

Concerning dilations of the form (Z, Z^{-p}) we have the following, not very surprising fact:

Proposition 2.2. *Let A, B be positive operators on $\mathcal{H}$. The statement $A \geq I$ and $B \geq I$ is equivalent to each of the following:*

(1) *For each $p > 0$, there exists a strictly positive operator Z on $\mathcal{F} \supset \mathcal{H}$ such that*

$$A = Z_{\mathcal{H}} \qquad \text{and} \qquad B = (Z^{-p})_{\mathcal{H}}.$$

(2) *For each $p > 0$, there exists a strictly positive operator Z on $\oplus^4 \mathcal{H}$ such that*

$$\operatorname{diag}(Z) = \oplus^4 A \qquad \text{and} \qquad \operatorname{diag}(Z^{-p}) = \oplus^4 B.$$

Proof. The proof requires some familiarity with basic properties of the numerical range (or field of values) of a normal operator and its connection with elementary dilation and compression results (see [6, chapter 1] for this background).

The implication (b)⇒(a) is obvious. Let us show that (a) implies $A \geq I$ and $B \geq I$. Fix $p > 0$. By assumption there exists a strictly positive operator Z on a larger space $\mathcal{F} \supset \mathcal{H}$ such that

$$A = Z_{\mathcal{H}} \qquad \text{and} \qquad B = (Z^{-p})_{\mathcal{H}}.$$

Therefore $A + iB = N_{\mathcal{H}}$ where N is the normal operator $Z + iZ^{-p}$. Consequently we must have

$$W(A + iB) \subset W(N)$$

where $W(\cdot)$ denotes the numerical range. For a normal operator M its numerical range equals to the convex hull of its spectrum: $W(M) = \operatorname{co}\operatorname{Sp}(M)$ (well, in case of $\dim\mathcal{H} = \infty$, this equality holds for the closure of the numerical range); hence, we must have

$$W(A + iB) \subset \operatorname{co}\operatorname{Sp}(N).$$

Now, we note that

$$\operatorname{Sp}(N) = \operatorname{Sp}(Z + iZ^{-p}) \subset \{z \in \mathbf{C} \ : \ z = t + it^{-p}, \ t > 0\}$$

so that we necessarily have

$$\begin{aligned} W(A + iB) \subset & \bigcap_{p>0} \operatorname{co}\{z \in \mathbf{C} \ : \ z = t + it^{-p}, \ t > 0\} \\ =& \{z \in \mathbf{C} \ : \ z = x + iy, \ x \geq 1, \ y \geq 1\} \end{aligned}$$

and this ensures that $A \geq I$ and $B \geq I$.

Now, let us prove that $A \geq I$ and $B \geq I$ imply (b). Fix $p > 0$. Both

$$S = \begin{pmatrix} A & A - I \\ A - I & A \end{pmatrix} \qquad \text{and} \qquad T = \begin{pmatrix} B & -B + I \\ -B + I & B \end{pmatrix}$$

are strictly positive operators on $\oplus^2\mathcal{H}$; moreover $S \geq I$ and $T \geq I$. Since $ST = TS$, $S + iT$ is a normal operator. From $S \geq I$ and $T \geq I$ we infer that

$$W(S + iT) \subset \{z \in \mathbf{C} \ : \ z = x + iy, \ x \geq 1, \ y \geq 1\} = \Gamma.$$

Some elementary geometric considerations then show that any point in Γ is the middle of two points lying in the curve

$$\Gamma_p = \{z \in \mathbf{C} \ : \ z = t + it^{-p}, \ t > 0\}.$$

In particular, any point z of the spectrum of $S + iT$ is the middle of two points α and β lying in Γ_p. Since the one-dimensional operator z can be dilated into the normal operator $G = \begin{pmatrix} z & (\alpha-\beta)/2 \\ (\alpha-\beta)/2 & z \end{pmatrix}$ with $\mathrm{Sp}(G) = \{\alpha, \beta\}$, a standard argument shows that $S + iT$ can be dilated into a normal operator, say N, acting on $(\oplus^2\mathcal{H}) \oplus (\oplus^2\mathcal{H}) = \oplus^4\mathcal{H}$ with $\mathrm{Sp}(N) \subset \Gamma_p$. This means that $N = Z + iZ^{-p}$ for some strictly positive operator Z on $\oplus^4\mathcal{H}$ and we deduce that

$$N_{\mathcal{H}} = Z_{\mathcal{H}} + i(Z^{-p})_{\mathcal{H}} = A + iB$$

so that $A = Z_{\mathcal{H}}$ and $B = (Z^{-p})_{\mathcal{H}}$ as wanted. $\square$

Proposition 2.3. *Let A, B be strictly positive operators on $\mathcal{H}$. Then, the condition $A^2 \leq B$ is equivalent to the existence of a strictly positive operator Z on $\oplus^k\mathcal{H}$, where k is any integer for which $B \leq kA^2$, such that*

$$\mathrm{diag}(Z) = \oplus^k A \qquad \textit{and} \qquad \mathrm{diag}(Z^2) = \oplus^k B.$$

Proof. Obviously the existence of Z implies $A^2 \leq B$ by the simple fact that $(ETE)^2 \leq ET^2E$ for any positive operator T and projection E acting on the same space. Conversely, we have

$$0 \leq B - A^2 \leq (k-1)A^2$$

and since $t \longrightarrow t^{1/2}$ is operator monotone [1, p. 115],

$$0 \leq \frac{(B - A^2)^{1/2}}{\sqrt{k-1}} \leq A.$$

From this we derive that the operator Z on $\oplus^k\mathcal{H}$ defined by

$$Z = \begin{pmatrix} A & \frac{(B-A^2)^{1/2}}{\sqrt{k-1}} & \cdots \\ \frac{(B-A^2)^{1/2}}{\sqrt{k-1}} & A & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}$$

is a positive operator. We then observe that

$$Z^2 = \begin{pmatrix} B & \star & \dots \\ \star & B & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}$$

and this completes the proof. $\square$

Proposition 2.4. *Let A and B be two positive operators on a space $\mathcal{H}$ and suppose that $I \geq B \geq (1/k)I$ for some integer k. Then there exist two positive operators S and T on $\oplus^k \mathcal{H}$ such that:*

(1) *S and T form a monotone pair of positive operators.*

(2) $\mathrm{diag}(S) = \oplus^k A$ and $\mathrm{diag}(T) = \oplus^k B$.

(3) $I \geq T$.

Proof. To simplify the notation we assume that $\mathcal{H}$ has a finite dimension. We dilate A into S with block-matrix representation

$$S = \begin{pmatrix} A & A & \dots \\ A & A & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

and we dilate B into

$$T = \begin{pmatrix} B & \frac{I-B}{k-1} & \dots \\ \frac{I-B}{k-1} & B & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}$$

We observe that $ST = T$ hence S commutes with T. Let $\{a_j\}_{j=1}^n$ be an orthonormal basis of $\mathcal{H}$ such that

$$A = \sum_{1 \leq j \leq n} \mu_j(A) a_j \otimes a_j.$$

Notice that $\{s_j = (1/\sqrt{k})(a_j \oplus \dots \oplus a_j)\}_{j=1}^n$ is a family of normalised eigenvectors of S whose corresponding eigenvalues are the numbers $\{k\mu_j(A)\}_{j=1}^n$. Moreover, setting $\mathcal{E} = \mathrm{span}\{s_j, 1 \leq j \leq n\}$ and $\mathcal{E}_j = \mathrm{span}\{(x_1 a_j \oplus \dots \oplus x_k a_j) \mid x_1 + \dots x_k = 0\}$ we note that

$$\bigoplus_{1 \leq j \leq n} \mathcal{E}_j = \mathcal{E}^{\perp} \subset \ker S.$$

Consequently,

$$S = k \sum_{1 \leq j \leq n} \mu_j(A) s_j \otimes s_j.$$

Now we observe that $\mathcal{E}$ reduces T and that $T_{\mathcal{E}}$ is the identity operator on $\mathcal{E}$. To prove the theorem it therefore remains to check that $0 \le T \le I$. For this purpose let $\{b_j\}_{j=1}^n$ be an orthonormal basis of $\mathcal{H}$ such that

$$B = \sum_{1 \le j \le n} \mu_j(B) b_j \otimes b_j$$

and consider the orthogonal decomposition $\mathcal{F} = \mathcal{G}_1 \bigoplus \cdots \bigoplus \mathcal{G}_n$ in which

$$\mathcal{G}_j = \operatorname{span}\{(b_j, 0, \ldots, 0); (0, b_j, \ldots, 0); \ldots; (0, \ldots, 0, b_j)\}.$$

Relatively to this decomposition $T = T_1 \bigoplus \cdots \bigoplus T_n$ with

$$T_j = \begin{pmatrix} \mu_j(B) & \frac{1-\mu_j(B)}{k-1} & \cdots \\ \frac{1-\mu_j(B)}{k-1} & \mu_j(B) & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}$$

Let $I_{(k)}$ be the k-by-k identity matrix and let $P_{(k)}$ be the k-by-k matrix whose entries all equal 1. We have

$$T_j = \frac{1 - \mu_j(B)}{k-1} P_{(k)} + \frac{k\mu_j(B) - 1}{k-1} I_{(k)}.$$

Since $1/k \le \mu_j(B) \le 1$ we infer that $T_j \ge 0$. Since $\|P_{(k)}\| = k$ we have $\|T_j\| = 1$. Thus $0 \le T_j \le I$ and consequently $0 \le T \le I$. $\square$

3. Other commuting dilations

Besides the total dilations obtained in the preceding section, it is natural to search commuting dilations on $\oplus^k \mathcal{H}$, and if possible, on $\mathcal{H} \oplus \mathcal{H}$. Our next result presents a particularly simple monotone dilation for Hermitian operators.

Proposition 3.1. *Let A, B be two Hermitian operators acting on $\mathcal{H}$. Then, there exists a Hermitian operator Z acting on $\mathcal{H} \oplus \mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = (Z^3)_{\mathcal{H}}$.*

Proof. Assume first that A is invertible and set

$$Z = \begin{pmatrix} A & A \\ A & A^{-1}BA^{-1} - 3A \end{pmatrix}.$$

Then (Z, Z^3) is a monotone pair of Hermitian operators dilating (A, B). If A is no longer invertible we take

$$Z = \begin{pmatrix} A & A \\ A & A_0^{-1}BA_0^{-1} - 3A \end{pmatrix}$$

in which $A_0 = A + P$, P being the projection onto $\ker A$. $\square$

Remark 3.2 The above proof is not valid when $\dim \mathcal{H} = \infty$ since A_0^{-1} may be unbounded. However the curve $\Gamma_3 = \{z \in \mathbf{C} \ : \ z = t + it^3, \ t \in (-\infty, \infty)\}$

satisfies the following property: any bounded region of $\mathbf{C}$ is contained in a triangle whose vertices belong to Γ_3. From Mirman's theorem we deduce that $A+iB$ can be dilated into a normal operator N acting on $\oplus^3\mathcal{H}$ with $\mathrm{Sp}(N) \subset \Gamma_3$. Consequently $(\mathrm{Re}N, \mathrm{Im}N) = (\mathrm{Re}N, (\mathrm{Re}N)^3)$ dilates (A, B). Since $\oplus^3\mathcal{H}$ and $\oplus^2\mathcal{H}$ are basically the same when $\dim\mathcal{H} = \infty$, Proposition 2.1 remains valid in the infinite dimensional case. For convenience to the reader we recall Mirman's theorem [7, 8]: Let A be an operator on $\mathcal{H}$ and suppose that $W(A)$ is contained in a triangle with vertices (α, β, γ). Then A can be dilated into a normal operator N on $\oplus^3\mathcal{H}$ with $\mathrm{Sp}(N) = \{\alpha, \beta, \gamma\}$. We also refer the reader to [4] for a transparent proof of Mirman's theorem.

For finite families of Hermitians, we have:

Proposition 3.3. *Let $\{A_j\}_{j=0}^n$ be Hermitian operators on a space $\mathcal{H}$. Then we can dilate them into a monotone family of Hermitian operators on a larger space $\mathcal{F}$ with $\dim\mathcal{F} = 2(n+1)\dim\mathcal{H} - 1$.*

Of course, if $\dim\mathcal{H} = \infty$ then we may take $\mathcal{F} = \mathcal{H}\bigoplus\mathcal{H}$. For sake of simplicity we suppose that $\mathcal{H}$ has a finite dimension and, in this setting, we first state an elementary lemma. Let us say that an operator B *essentially* acts on a subspace $\mathcal{E}$ if both the range and the corange of B are contained in $\mathcal{E}$ (equivalently, $\mathrm{ran}B \subset \mathcal{E}$ and $(\ker B)^\perp \subset \mathcal{E}$).

Lemma 3.4. *Fix an integer n and a space $\mathcal{H}$. Then there exist a larger space $\mathcal{F}$, $\dim\mathcal{F} = (n+1)\dim\mathcal{H}$, and an orthogonal decomposition $\mathcal{F} = \mathcal{E}_0 \oplus \cdots \oplus \mathcal{E}_n$, in which $\dim\mathcal{E}_j = \dim\mathcal{H}$ for each j, such that: for every family of operators $\{A_j\}_{j=0}^n$ on $\mathcal{H}$ there is a family $\{B_j\}_{j=0}^n$ of operators on $\mathcal{F}$ with B_j essentially acting on $\mathcal{E}_j$ and $A_j = (B_j)_\mathcal{H}$, $0 \le j \le n$. Moreover when the A_j's are Hermitian or positive, the B_j's can be taken of the same type.*

Let us sketch the elementary proof of this lemma. First, choose subspaces $\{\mathcal{E}_j\}_{j=0}^n$ of $\mathcal{F} = \oplus^{n+1}\mathcal{H}$ in such a way that for each j (a) $\dim\mathcal{E}_j = \dim\mathcal{H}$, (b) The projection E_j from $\mathcal{F}$ onto $\mathcal{E}_j$ verifies: $(E_j)_\mathcal{H}$ is a strictly positive operator on $\mathcal{H}$. Now, fix an integer j and observe that any vector $h \in \mathcal{H}$ can be lifted to a unique vector $h_j \in \mathcal{E}_j$ such that $Hh_j = h$, where H is the projection onto $\mathcal{H}$. Consequently any rank one operator of the form $R = h\otimes h$, $h \in \mathcal{H}$, can be lifted into a positive rank one operator T essentially acting on $\mathcal{E}_j$ such that $T_\mathcal{H} = R$. This ensures that given a general (resp. hermitian, positive) operator A on $\mathcal{H}$ there exists a general (resp. Hermitian, positive) operator B essentially acting on $\mathcal{E}_j$ such that $B_\mathcal{H} = A$.

We turn to the proof of proposition 3.3.

Proof. By Lemma 3.4 we may dilate $\{A_j\}_{j=0}^n$ into a commuting family of Hermitians $\{S_j\}_{j=0}^n$ on a larger space $\mathcal{G}$ with $\dim\mathcal{G} = (n+1)\dim\mathcal{H} = d$. Thus, there is a basis $\{g_k\}_{k=0}^d$ in $\mathcal{G}$ and real numbers $\{s_{j,k}\}$ such that

$$S_j = \sum_{k=0}^{d} s_{j,k} g_k \otimes g_k \qquad (0 \le j \le n).$$

We take for $\mathcal{F}$ a space of the form

$$\mathcal{F} = \mathcal{E}_0 \oplus \mathcal{E}_1 \oplus \cdots \oplus \mathcal{E}_d$$

in which $\dim \mathcal{E}_0 = 1$ and $g_0 \in \mathcal{E}_0$; and for $k > 0$, $\dim \mathcal{E}_k = 2$ and $g_k \in \mathcal{E}_k$. Hence, we have $\dim \mathcal{F} = 2(n+1) \dim \mathcal{H} - 1$.

For $k > 0$, let $\{e_{1,k}; e_{2,k}\}$ be a basis of $\mathcal{E}_k$ and suppose that $g_k = (e_{1,k} + e_{2,k})/\sqrt{2}$ (*). We set, for $0 \le j \le n$,

$$B_j = s_{j,0} g_0 \otimes g_0 + \sum_{k=1}^{d} (r_{j,k} e_{1,k} \otimes e_{1,k} + t_{j,k} e_{2,k} \otimes e_{2,k})$$

where the reals $r_{j,k}$ and $t_{j,k}$ are chosen in such a way that:

(1) $s_{j,k} = (r_{j,k} + t_{j,k})/2, \qquad j = 0, \dots n.$

(2) $r_{j,d} < \cdots < r_{j,1} < s_{j,0} < t_{j,1} < \cdots < t_{j,d}, \qquad j = 0, \dots n.$

From (1) and (*) we deduce that $S_j = (B_j)_{\mathcal{G}}$ so that $A_j = (B_j)_{\mathcal{H}}$. From (2) we infer that $\{B_j\}_{j=0}^n$ is a monotone family. $\square$

Now, we focus on dilations of the type (Z, Z^p), $Z \ge 0$, $p > 0$. In connection with Proposition 3.1, we have:

Proposition 3.5. *Let A, B be strictly positive operators on a finite dimensional space $\mathcal{H}$ and let $p = 2, 3$. The condition $A^p < B$ ensures the existence of a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = Z^p_{\mathcal{H}}$.*

Proof. We already know that the existence of Z entails $A^2 \le B$. In the converse direction it suffices to take

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

in which $\lambda > 0$ is chosen large enough to ensure $Z > 0$ (this is possible since A is invertible).

Now, suppose that $A^3 < B$ and choose $1 > \varepsilon > 0$ small enough to ensure that $B \ge (1 + 3\varepsilon^2) A^3$. We then take

$$Z = \begin{pmatrix} A & \varepsilon A \\ \varepsilon A & \varepsilon^{-2}[A^{-1} B A^{-1} - (1 + 2\varepsilon^2) A] \end{pmatrix}.$$

We observe that, with our choice for ε, $Z > 0$. A direct computation then shows $B = (Z^3)_{\mathcal{H}}$. $\square$

For sake of completeness we state the next proposition which is an easy application of Mirman's Theorem.

Proposition 3.6. *Let A, B be strictly positive operators on $\mathcal{H}$ and assume that $A \le I \le B$. Then, for each $p > 1$, there exists a strictly positive operator Z on $\oplus^3\mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = Z^p_{\mathcal{H}}$.*

Proof. $A + iB$ is an operator whose numerical range lies in the region

$$\Gamma = \{z \in \mathbf{C} \ : \ z = x + iy, \ \varepsilon \le x \le 1, \ 1 \le y\}$$

in which $\varepsilon = ||A^{-1}||^{-1}$. Now, fix $p > 1$ and let

$$\Gamma_p = \{z \in \mathbf{C} \ : \ z = t + it^p, \ t > 0\}.$$

We observe that $W(A + iB)$ is contained in a triangle whose vertices are three points in Γ_p. Mirman's theorem entails that $A + iB$ can be dilated into a normal operator N acting on $\oplus^3\mathcal{H}$ with $\mathrm{Sp}(N) \subset \Gamma_p$. We then deduce that $N = Z + iZ^p$ for some strictly positive operator on $\oplus^3\mathcal{H}$. Therefore $A = Z_{\mathcal{H}}$ and $B = (Z^p)_{\mathcal{H}}$. $\square$

From Theorem 2.1 and Proposition 3.5 we derive

Corollary 3.7. *Let A, B be strictly positive operators on $\mathcal{H}$. The following statements are equivalent:*

(1) $A \le B$.

(2) *There exists a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that*

$$A = (Z_{\mathcal{H}})^{-1} \qquad \textit{and} \qquad B = (Z^{-1})_{\mathcal{H}}.$$

(3) *There exists a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that*

$$A = (Z_{\mathcal{H}})^2 \qquad \textit{and} \qquad B = (Z^2)_{\mathcal{H}}.$$

Proof. Implications (2)⇒(1) and (3)⇒(1) are known.

Let us check the implication (1)⇒(2). Let $X = A^{-1}$ and $Y = B$. Since $t \longrightarrow -t^{-1}$ is operator monotone, $X \ge Y^{-1}$. Hence, Theorem 2.1 entails the existence of a strictly positive operator Z on $\mathcal{F} \supset \mathcal{H}$ such that $X = Z_{\mathcal{H}}$ and $Y = Z^{-1}_{\mathcal{H}}$, that is $A = (Z_{\mathcal{H}})^{-1}$ and $B = (Z^{-1})_{\mathcal{H}}$.

To check the implication (1)⇒(3) we set $X = A^{1/2}$ and $Y = B$. Consequently $X^2 \le Y$ and Proposition 3.5 ensures the existence of a strictly positive operator Z on $\mathcal{F} \supset \mathcal{H}$ such that $X = Z_{\mathcal{H}}$ and $Y = (Z^2)_{\mathcal{H}}$. Thus, we have $A = (Z_{\mathcal{H}})^2$ and $B = (Z^2)_{\mathcal{H}}$. $\square$

In view of the above corollary it seems natural to pose:

Conjecture 3.8. *Let A, B be strictly positive operators on $\mathcal{H}$ and let $f : (0,\infty) \longrightarrow (0,\infty)$ be onto, nonlinear and operator convex. Then, the following statements are equivalent:*

(1) $A \le B$.

(2) *There exists a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that*

$$A = f(Z_{\mathcal{H}}) \qquad \text{and} \qquad B = f(Z)_{\mathcal{H}}.$$

A convex function $f : (0,\infty) \to (-\infty,\infty)$ is said to be strongly convex if its epigraph is the convex hull of its graph.

Proposition 3.9. *Let A, B be two commuting strictly positive operators on $\mathcal{H}$ and let $f : (0,\infty) \to (-\infty,\infty)$ be strongly convex. Then, the condition $f(A) \le B$ ensures the existence of a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = f(Z)_{\mathcal{H}}$.*

Proof. Let $T = A + iB$ and let $z \in W(T)$. So, there exists a norm one vector h such that

$$z = x + iy = \langle h, Ah\rangle + i\langle h, Bh\rangle.$$

By convexity of f, we have

$$f(x) = f(\langle h, Ah\rangle) \le \langle h, f(A)h\rangle \le \langle h, Bh\rangle = y.$$

But the relation $f(x) \le y$ entails that $W(T)$ lies in the region

$$R = \{(x + iy) \ : \quad x > 0,\ y \ge f(x)\}.$$

By assumptions on f, any point of R is then the convex combination of two points of the curve

$$\Gamma = \{x + if(x),\ x > 0\}.$$

In particular, any point of the spectrum of T is a convex combination of two points in Γ. Since T is normal, a standard dilation argument (the same as that one of the proof of Proposition 2.2) shows that T can be dilated into a normal operator N on $\mathcal{H} \oplus \mathcal{H}$ with spectrum in Γ. Therefore $N = Z + if(Z)$ for some strictly positive operator Z and $A = Z_{\mathcal{H}}$, $B = f(Z)_{\mathcal{H}}$ as desired. $\square$

Conjecture 3.10. *Let A, B be strictly positive operators on a finite dimensional space $\mathcal{H}$ and let $f : (0,\infty) \to (-\infty,\infty)$ be strongly convex. Then, the condition $f(A) < B$ ensures the existence of a strictly positive operator Z on $\mathcal{H} \oplus \mathcal{H}$ such that $A = Z_{\mathcal{H}}$ and $B = f(Z)_{\mathcal{H}}$.*

4. Addenda

In [2], we established the following extension of Proposition 2.4

Proposition 4.1. *Let $\{A_j\}_{j=0}^n$ be positive operators on $\mathcal{H}$. Assume that for $j > 0$ we have integers $k_j > 0$ such that $I \geq A_j \geq (1/k_j)I$. Then there exist positive operators $\{B_j\}_{j=0}^n$ on $\oplus^k\mathcal{H}$, where $k = \prod_{j=1}^n k_j$, such that:*

(1) *$\{B_j\}_{j=0}^n$ is a monotone family of positive operators.*

(2) $\mathrm{diag}(B_j) = \oplus^k A_j$, $0 \leq j \leq n$.

Furthermore, we may require that $\|B_1\|_\infty \leq 1$.

Corollary 4.2. *Let $\{A_j\}_{j=0}^n$ be hermitian operators on $\mathcal{H}$ with $\|A_0\|_\infty \leq 1$. Then we can totally dilate them into a monotone family of hermitian operators $\{B_j\}_{j=0}^n$ on $\oplus^k\mathcal{H}$, $k = 2^n$, in such a way that $\|B_0\|_\infty \leq 1$.*

References

[1] R. Bhatia, Matrix Analysis, Springer, Germany, 1996
[2] J.-C. Bourin, Singular values of compressions, restrictions and dilations, Linear Algebra Appl. 360 (2003) 259-272.
[3] J.-C. Bourin, Total dilations, Linear Algebra Appl. 368 (2003) 159-169.
[4] M.-D. Choi, C.-K. Li, Numerical ranges and dilations, Linear Multilinear Algebra 47 (2000) 35-48.
[5] R.A. Horn, C.R. Johnson, Matrix analysis, Cambrige University Press, 1985.
[6] R.A. Horn, C.R. Johnson, Topics in matrix analysis, Cambrige University Press, 1991.
[7] B.A. Mirman, Numerical range and norm of a linear operator, Trudy Seminara po Funkcional' nomu Analyzu 10 (1968) 51-55.
[8] Y. Nakamura, Numerical range and norm, Math. Japonica 27 (1982) 149-150.

Chapter 4

Inequalities for some commuting pairs of positive operators

Introduction

For all positive operator A and normal operator Z on a separable Hilbert space, the interpolation inequality

$$\|A^s Z A^t\|_\infty \leq \|Z A^{s+t}\|_\infty, \quad s, t > 0$$

holds (we denote by $\|\cdot\|_\infty$ the usual operator norm). Such a result belongs to the folklore in Matrix/operator theory (there might be a precise reference). The aim of this first chapter is to establish several inequalities for pairs (A, B) generalizing (A^s, A^t).

Let A, B be two commuting Hermitian operators (on a separable Hilbert space). If there exist a positive operator C and two nondecreasing functions f, g such that

$$A = f(C) \quad \text{and} \quad B = g(C),$$

we then say that A and B form a *monotone* pair of positive operators. If instead f is still nondecreasing but g is now nonincreasing, we then say that (A, B) is an *antimonotone* pair of positive operators. Some classical inequalities can be rephrased in terms of monotone pairs: For instance, in the finite dimensional setting, von Neumann's trace inequality claims that

$$|\mathrm{Tr}\, UAVB| \leq \mathrm{Tr}\, AB$$

for all monotone pairs (A, B) of positive operators and all unitary operators U, V (see [1, p. 94-95] for stronger results)

Section 1 is devoted to a basic lemma about some pairs of integrable functions on a probability space and some applications to operators. This lemma is also the key to the results of Section 2 in which we show several inequalities involving compressions of monotone pairs. Recall that, given an operator X and a projection E onto a subspace $\mathcal{E}$, the compression of X onto $\mathcal{E}$, denoted by $X_{\mathcal{E}}$, is defined as the restriction of EXE to $\mathcal{E}$. An example of the obtained results is

$$\det A_{\mathcal{E}} \cdot \det B_{\mathcal{E}} \leq \det(AB)_{\mathcal{E}}$$

for every monotone pairs (A, B) of positive operators and every finite dimensional subspace $\mathcal{E}$.

In section 3 we prove the following inequalities for the Hilbert-Schmidt norm $\|\cdot\|_2$: for every normal (or even hyponormal) operator Z and every monotone pair (A, B) of positive operators, we have

$$\|AZB\|_2 \leq \|ZAB\|_2,$$

meanwhile if (A, B) is antimonotone and Z is normal, Hilbert-Schmidt, then the reverse inequality holds.

1. A basic lemma

Let γ be a real valued function on a set Ω and let ϕ, ψ be two real valued functions on the real line. Set

$$f = \phi \circ \gamma \quad \text{and} \quad g = \psi \circ \gamma.$$

If ϕ and ψ are both nondecreasing, we then say that f and g form a *monotone* pair. If ϕ is nondecreasing and ψ is nonincreasing, we then say that f and g form an *antimonotone* pair. Monotone (resp. antimonotone) pairs of functions (f, g) on Ω satisfies the property: For all x and y in Ω,

$$[f(x) - f(y)] \cdot [g(x) - g(y)] \geq 0 \,(\text{resp. } \leq)\, 0.$$

We have the following basic fact:

Lemma 1.1. *Let f and g be two positive measurable functions on a probability space. Then:*

(1) *If (f, g) is monotone,*

$$\int_\Omega f \,\mathrm{d}P \cdot \int_\Omega g \,\mathrm{d}P \leq \int_\Omega fg \,\mathrm{d}P$$

(2) *If (f, g) is antimonotone,*

$$\int_\Omega f \,\mathrm{d}P \cdot \int_\Omega g \,\mathrm{d}P \geq \int_\Omega fg \,\mathrm{d}P$$

Proof. We prove the monotone case. For any x and y in the probability space Ω, we have

$$[f(x) - f(y)] \cdot [g(x) - g(y)] \geq 0,$$

consequently

$$f(x)g(y) + f(y)g(x) \leq f(x)g(x) + f(y)g(y). \tag{1}$$

Now, we compute

$$
\begin{aligned}
2\int_\Omega f\,\mathrm{d}P\cdot\int_\Omega g\,\mathrm{d}P &= \int_{\Omega\times\Omega}[f(x)g(y)+f(y)g(x)]\,\mathrm{dP}(x)\mathrm{dP}(y)\\
&\le \int_{\Omega\times\Omega}[f(x)g(x)+f(y)g(y)]\,\mathrm{dP}(x)\mathrm{dP}(y)\\
&= 2\int_\Omega fg\,\mathrm{d}P.
\end{aligned}
$$

When (f,g) is antimonotone, inequality (1) is reversed and the proof is similar. □

Let A be a Hermitian operator with spectrum Ω and spectral decomposition

$$A=\int_\Omega \lambda\,\mathrm{d}E(\lambda).$$

For every bounded Borel function f on Ω and every norm one vector h, we have

$$\langle h, f(A)h\rangle = \int_\Omega f(\lambda)\,\mathrm{d}P(\lambda)$$

where P is a probability measure on Ω, $\mathrm{d}P(\lambda)=\mathrm{d}\langle h,E(\lambda)h\rangle$. Consequently, Lemma 1.1 admits the following operator formulation:

Lemma 1.2. *Let A, B be a pair of positive operators and let h be a norm one vector.*

(1) *If (A,B) forms a monotone pair,*

$$\langle h,Ah\rangle\langle h,Bh\rangle \le \langle h,ABh\rangle \quad \textit{and} \quad \|Ah\|\cdot\|Bh\|\le\|ABh\|.$$

(2) *If (A,B) forms an antimonotone pair,*

$$\langle h,Ah\rangle\langle h,Bh\rangle \ge \langle h,ABh\rangle \quad \textit{and} \quad \|Ah\|\cdot\|Bh\|\ge\|ABh\|.$$

The scalar product inequalities imply the norm inequalities by replacing A and B by A^2 and B^2.

As an application of the above lemma, we have

Proposition 1.3. *Let X,Y be two positive operators with Y invertible and $X\le Y$. If M is a positive trace class operator which commutes with X and α, β are two reals such that $\alpha\ge 0$, $\beta\ge -1$ and $\alpha+\beta\ge 0$, then we have*

$$\operatorname{Tr} MX^\alpha Y^\beta \le \operatorname{Tr} MY^{\alpha+\beta}.$$

Proof. By repeating the process, we may assume that $\alpha\le 1$. By a limit argument, we may assume that X is invertible. There exists an orthonormal system $\{e_n\}$

and two sequences of reals $\{x_n\}$ and $\{m_n\}$ such that $M = \sum_n m_n e_n \otimes e_n$ and $X(e_n) = x_n e_n$. Thus

$$\operatorname{Tr} MY^{\alpha+\beta} = \sum_n m_n x_n^\alpha \langle e_n, X^{-\alpha} e_n\rangle \langle e_n, Y^{\alpha+\beta} e_n\rangle.$$

By Loewner's theorem, $t \longrightarrow t^\alpha$, $0 < \alpha \leq 1$ is operator monotone [1, p. 115]. Since $t \to t^{-1}$ is operator decreasing [1, p. 114], then so is $t \to t^{-\alpha}$. Consequently,

$$\operatorname{Tr} MY^{\alpha+\beta} \geq \sum_n m_n x_n^\alpha \langle e_n, Y^{-\alpha} e_n\rangle \langle e_n, Y^{\alpha+\beta} e_n\rangle.$$

Since $t \to t^{-\alpha}$ decreases and $t \to t^{\alpha+\beta}$ increases, Lemma 1.2 implies

$$\begin{aligned} \operatorname{Tr} MY^{\alpha+\beta} &\geq \sum_n m_n x_n^\alpha \langle e_n, Y^{\beta} e_n\rangle. \\ &= \operatorname{Tr} MX^\alpha Y^\beta \end{aligned}$$

and we get the result. $\square$

Remark 1.4. A special case of Proposition 1.3 is when $\beta \geq 0$. Then, in order to prove the proposition, it suffices to use the most basic case of Loewner's Theorem, namely that $t \longrightarrow t^{1/2}$ is operator monotone (and so are $t \longrightarrow t^{1/2^n}$, $n = 1, 2, \ldots$). We also note that the proposition with $\beta \geq 0$ still holds for a not necessarily invertible Y. Proposition 1.3 gives an immediate proof of the McCarthy inequality (cf [5] p.20, theorem 1.22):

$$\operatorname{Tr}(X+Y)^p \geq \operatorname{Tr} X^p + \operatorname{Tr} Y^p \quad (0 \leq X,\, Y;\ p \geq 1)$$

Indeed,

$$\operatorname{Tr}(X+Y)^p = \operatorname{Tr} X(X+Y)^{p-1} + \operatorname{Tr} Y(X+Y)^{p-1} \geq \operatorname{Tr} X^p + \operatorname{Tr} Y^p.$$

Similarly we also get:

$$\operatorname{Tr}(X+Y)^p \geq \operatorname{Tr} X^p + \operatorname{Tr} Y^p + \operatorname{Tr}(XY^{p-1} + YX^{p-1}) \quad (0 \leq X,\, Y;\ p \geq 2).$$

Example 1.5. Proposition 1.3 is not valid for all reals β and all nonnegative reals α: For instance taking $\alpha = 3$, $\beta = -7$ and

$$M = X = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \quad ; \quad Y = \begin{pmatrix} 9 & 0 \\ 0 & 2 \end{pmatrix}$$

one has $\operatorname{Tr} X^4 Y^{-7} > \operatorname{Tr} XY^{-4}$.

Our purpose is now to apply Lemma 1.2 to a problem about compressions. If X is a Hermitian operator acting on an n-dimensional space, we denote by $\operatorname{Eig}(X)$ the sequence $\{\lambda_k(X)\}_{k=1}^n$ of the eigenvalues of X arranged in decreasing order and counted with their multiplicities. Let $\mathcal{H}$ be a finite dimensional space. Fix a subspace $\mathcal{E}$ of $\mathcal{H}$ and a Hermitian operator A on $\mathcal{H}$. We look for the collection of subspaces $\mathcal{F}$ with the same dimension as $\mathcal{E}$ for which $\operatorname{Eig}(A_{\mathcal{E}}) \leq \operatorname{Eig}(A_{\mathcal{F}})$. We may answer a closely related problem.

Proposition 1.6. *Let A, X be operators on a finite dimensional space. Suppose that A is Hermitian and X is invertible. The following conditions are equivalent.*

(1) $AX = XA$ *and* $(A, |X|)$ *is a monotone pair.*

(2) *For every subspace* $\mathcal{E}$,

$$\mathrm{Eig}(A_{X(\mathcal{E})}) \geq \mathrm{Eig}(A_{\mathcal{E}}).$$

We will give a statement equivalent to Proposition 1.6. Given a subspace $\mathcal{E}$ and an operator A, denote by $A\!:\!\mathcal{E}$ the restriction of A to $\mathcal{E}$ and by $\mathrm{Sing}(A\!:\!\mathcal{E})$ the sequence of the singular values $\{\mu_k(A\!:\!\mathcal{E})\}$ arranged in decreasing order and counted with their multiplicities.

Proposition 1.7. *Let A, X be operators on a finite dimensional space. Suppose that X is invertible. The following conditions are equivalent.*

(1) $|A|X = X|A|$ *and* $(|A|, |X|)$ *is a monotone pair.*

(2) *For every subspace* $\mathcal{E}$,

$$\mathrm{Sing}(A\!:\!X(\mathcal{E})) \geq \mathrm{Sing}(A\!:\!\mathcal{E}).$$

Proof. We may assume that A is positive, $A = |A|$.

Suppose that X satisfies condition (1). Then, if h is a norm one vector, Lemma 1.2 implies $\|AXh\| \geq \|Ah\|.\|Xh\|$. Equivalently, $\|Au\| \geq \|Ah\|$ where $u = Xh/\|Xh\|$. Denoting by $\mathcal{L}$ the line spanned by h, this can be read as

$$\|A\!:\!X(\mathcal{L})\| \geq \|A\!:\!\mathcal{L}\|. \qquad (*)$$

Fix a subspace $\mathcal{E}$. We recall the version of the minimax principle for singular values: for $j \leq \dim \mathcal{E}$,

$$\mu_j(A\!:\!\mathcal{E}) = \min \|A\!:\!\mathcal{F}\|$$

where the minimum runs over all the subspaces $\mathcal{F} \subset \mathcal{E}$ with $\mathrm{codim}_{\mathcal{E}}\mathcal{F} = j - 1$. Using this basic principle and the invertibility of X we then deduce that there exists $\mathcal{F} \subset \mathcal{E}$, $\mathrm{codim}_{\mathcal{E}}\mathcal{F} = j - 1$, such that

$$\begin{aligned} \mu_j(A\!:\!X(\mathcal{E})) &= \|A\!:\!X(\mathcal{F})\| \\ &\geq \|A\!:\!X(\mathcal{L})\| \quad \text{for all lines } \mathcal{L} \subset \mathcal{F} \\ &\geq \|A\!:\!\mathcal{L}\| \quad \text{for all lines } \mathcal{L} \subset \mathcal{F} \text{ by } (*). \end{aligned}$$

Therefore

$$\mu_j(A\!:\!X(\mathcal{E})) \geq \|A\!:\!\mathcal{F}\| \geq \mu_j(A\!:\!\mathcal{E}).$$

Conversely, suppose that $\mathrm{Sing}(A : X(\mathcal{E})) \geq \mathrm{Sing}(A : \mathcal{E})$ for each $\mathcal{E}$. Denote by $s_1(A) > \cdots > s_l(A)$ the singular values of A arranged in decreasing order but not counted with their multiplicities. Let $\mathcal{S}_1, \ldots, \mathcal{S}_l$ be the corresponding spectral subspaces of A. We note the following fact, easily proved by arguing by the contrary:

If $\dim \mathcal{F} = \dim(\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k)$ and $\mathcal{F} \neq \mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k$ then, for at least one integer $j \leq \dim \mathcal{F}$, we have $\mu_j(A\!:\!\mathcal{F}) < \mu_j(A\!:\!\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k)$.

It follows that our operator X must satisfy

$$X(\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k) = \mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k$$

for each integer k, $k \leq l$. Let $\mathcal{L}$ be a line in $\mathcal{S}_k$. The preceding identity shows that there is a line $\mathcal{T}$ in $\mathcal{S}_1 \oplus \cdots \oplus \mathcal{S}_k$ such that $X(\mathcal{T}) = \mathcal{L}$. Since $\|A\!:\!X(\mathcal{T})\| \geq \|A\!:\!\mathcal{T}\|$ we must have $\mathcal{T} \subset \mathcal{S}_k$. Thus $X(\mathcal{S}_k) = \mathcal{S}_k$. This is exactly the same as saying that X commutes with A. Then $|X|$ also commutes with A. We claim that $(A, |X|)$ is a monotone pair. By the contrary there would exist an integer k and two norm one eigenvectors for $|X|$, u and v, such that : $u \in \mathcal{S}_k$, $v \in \mathcal{S}_{k+1}$ and $a = \|Xu\| < \|Xv\| = b$. Denoting by $\mathcal{L}$ the line spanned by $u + v$ we compute

$$\|A\!:\!\mathcal{L}\|^2 = \frac{s_k^2(A) + s_{k+1}^2(A)}{2} \quad \text{and} \quad \|A\!:\!X(\mathcal{L})\|^2 = \frac{a^2 s_k^2(A) + b^2 s_{k+1}^2(A)}{a^2 + b^2}.$$

Since $a < b$ and $s_k(A) > s_{k+1}(A)$ we conclude that $\|A\!:\!X(\mathcal{L})\| < \|A\!:\!\mathcal{L}\|$ and we reach a contradiction. Hence X must verify condition (1). $\Box$

Proof of Proposition 1.6. Replacing A by $A + rI$ with r large enough we may assume $A \geq 0$. Now, note that conditions (1) and (2) of the proposition are equivalent to

(i) $AX = XA$ and $(A, |X|)$ is a monotone pair

and

(ii) For every subspace $\mathcal{E}$,

$$\mathrm{Sing}^2(A^{1/2}\!:\!X(\mathcal{E})) \geq \mathrm{Sing}^2(A^{1/2}\!:\!\mathcal{E})$$

respectively. Since (i) is also equivalent to "$A^{1/2}X = XA^{1/2}$ and $(A^{1/2}, |X|)$ is a monotone pair", Proposition 1.7 completes the proof. $\Box$

2. Compressions of monotone pairs

Recall the following fact (Chapter 2, Corollary 1.4):

Lemma 2.1. *Let Z be a positive operator on a finite dimensional space. Let a and b be the extremal nonzero eigenvalues of Z. Then, for every projection E whose range is contained in the range of Z,*

$$EZE \leq \frac{(a+b)^2}{4ab} Z$$

Proposition 2.2. *Let A, B, Z be positive operators on a finite dimensional space. Let a and b be the extremal nonzero eigenvalues of Z. If (A,B) is monotone, then,*

$$\|AZB\|_\infty \le \frac{a+b}{2\sqrt{ab}}\,\|ZAB\|_\infty,$$

in particular, for every projection E,

$$\|AEB\|_\infty \le \|EAB\|_\infty. \tag{3}$$

Proof. We first establish (3). To this end, let f be a norm one vector such that $\|AEBf\| = \|AEB\|_\infty$ and let $h = EBf/\|EBf\|$. We then have, using Lemma 1.2,

$$\begin{aligned}\|AEB\|_\infty &\le \|A\,h\otimes h\,B\|_\infty \\ &= \|Ah\|\|Bh\| \\ &\le \|ABh\| \le \|EAB\|_\infty.\end{aligned}$$

Now, we consider the case of a general positive operator Z. Let E be the projection onto the support $\mathcal{E}$ of Z. There exists a norm one vector h in $\mathcal{E}$ such that

$$\begin{aligned}\|AZB\|_\infty &\le \|AZ^{1/2}\,h\otimes h\,Z^{1/2}B\|_\infty \\ &= \|A(Z^{1/2}h\otimes Z^{1/2}h)B\|_\infty.\end{aligned}$$

Therefore, using (3),

$$\|AZB\|_\infty \le \|(Z^{1/2}h\otimes Z^{1/2}h)AB\|_\infty. \tag{4}$$

Now, observe that

$$(Z^{1/2}h\otimes Z^{1/2}h)^2 \le \frac{(a+b)^2}{4ab}\,Z^2, \tag{5}$$

indeed, this is equivalent to

$$Z^{-1/2}(Z^{1/2}h\otimes Z^{1/2}h)^2 Z^{-1/2} \le \frac{(a+b)^2}{4ab}\,Z$$

which can also be written as

$$h\otimes h\cdot Z\cdot h\otimes h \le \frac{(a+b)^2}{4ab}\,Z$$

and which holds by Lemma 2.1. The obvious identity $\|X\|_\infty^2 = \|X^*X\|_\infty$ for all operators X combined with (4) and (5) then yield the proposition. □

Theorem 2.3. *Assume that A and B form a monotone pair of positive operators and let E be the projection onto a finite dimensional subspace $\mathcal{E}$. Then, we have*

$$\mathrm{Sing}(AEB) \le \mathrm{Sing}(EAB).$$

Consequently,

$$\mathrm{Eig}(A_{\mathcal{E}}B_{\mathcal{E}}) \le \mathrm{Eig}((AB)_{\mathcal{E}})$$

and

$$\mathrm{Eig}(A_{\mathcal{E}}B_{\mathcal{E}}A_{\mathcal{E}}) \le \mathrm{Eig}((ABA)_{\mathcal{E}}).$$

If we take $A = B$ in the third inequality we obtain

$$\mathrm{Eig}((A_{\mathcal{E}})^3) \le \mathrm{Eig}((A^3)_{\mathcal{E}}),$$

a special case of results of Chapter 1.

Proof. Let $\mathcal{E}$ be the range of E. By a continuity argument we may assume that B is invertible. If $k > \mathrm{rank}E$ then, obviously, $\mu_k(AEB) = \mu_k(ABE) = 0$.

Let $1 \le k \le \mathrm{rank}E$. By the minimax principle for singular values we have

$$\mu_k(AEB) = \min_F \|AEBF\|_\infty,$$

where the minimum runs over all the projections F with $\mathrm{corank}F = k-1$. Therefore,

$$\mu_k(AEB) \le \|AEBG\|_\infty \tag{6}$$

where G is the projection onto the subspace $\mathcal{G}$ of codimension $k-1$ defined by

$$\mathcal{G} = \mathrm{span}\{B^{-1}(\mathcal{E}^\perp),\, B^{-1}(\mathcal{E}_0)\},$$

$\mathcal{E}_0$ being a subspace of $\mathcal{E}$ such that $\mathrm{codim}\,(\mathcal{E}^\perp \bigoplus \mathcal{E}_0) = k-1$ and

$$\mu_k(ABE) = \|ABE(E^\perp + E_0)\|_\infty = \|ABE_0\|_\infty. \tag{7}$$

By definition of $\mathcal{G}$ we have $\|AEBG\|_\infty = \|AE_0BG\|_\infty$ so that

$$\begin{aligned}\|AEBG\|_\infty \le& \|AE_0B\|_\infty \\ \le& \|ABE_0\|_\infty \quad \text{(by (3))}.\end{aligned}$$

From (6) and (7) we then obtain the first assertion of the theorem.

For each integer k, $1 \le k \le \dim\mathcal{E}$, we have

$$\begin{aligned}\lambda_k(A_{\mathcal{E}}B_{\mathcal{E}}) = \lambda_k(EAEBE) =& \lambda_k(A^{1/2}EBEA^{1/2}) \\ =& \mu_k^2(A^{1/2}EB^{1/2}) \\ \le& \mu_k^2(EA^{1/2}B^{1/2}) = \lambda_k((AB)_{\mathcal{E}}),\end{aligned}$$

thus the second assertion of the theorem holds. The proof of the third one is similar:

$$\begin{aligned}\lambda_k(A_{\mathcal{E}}B_{\mathcal{E}}A_{\mathcal{E}}) &= \mu_k^2(B^{1/2}EAE) \\ &\le \mu_k^2(B^{1/2}EA) \\ &\le \mu_k^2(B^{1/2}AE) = \lambda_k((ABA)_{\mathcal{E}})\end{aligned}$$

□

As for an invertible positive operator A and $s, t > 0$ we have $\|A^sZA^{-t}\|_\infty \ge \|ZA^{s-t}\|_\infty$ for every normal operator Z, one may ask whether the previous inequalties are reversed for antimonotone pairs of positive operators. This is not true, as shown by the next examples.

Example 2.4. (1) Consider the following antimonotone pair (A, B) and projection E.

$$A = \begin{pmatrix} 1+\varepsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad E = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix},$$

then, $\|AEB\|_\infty < \|EAB\|_\infty$ if ε is small enough(for $\varepsilon = 0$, it is $0.772 < 0.816$).

(2) If (A, B) is an antimonotone pair the reverse inequality of the third claim of Theorem 2.4 is not valid. Take

$$A = \begin{pmatrix} 9 & 9 & 8 \\ 9 & 10 & 9 \\ 8 & 9 & 9 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then, setting $\mathcal{E} = \operatorname{ran} E$, $\mu_2((A^{-1})_{\mathcal{E}}(A^2)_{\mathcal{E}}(A^{-1})_{\mathcal{E}}) = 0.667.. < 1 = \mu_2(I_{\mathcal{E}})$.

Proposition 2.5. *Let A, B, Z be positive operators on a finite dimensional space. Let a and b be the largest and the smallest nonzero eigenvalues of Z. If (A, B) is monotone, then,*

$$\operatorname{Sing}(AZB) \le \sqrt{a/b}\,\operatorname{Sing}(ZAB).$$

In view of Proposition 2.2, may one replace $\sqrt{a/b}$ by $(a+b)/(2\sqrt{ab})$? We can not prove nor disprove it.

Proof. Suppose that the operators act on $\mathcal{H}$, $\dim \mathcal{H} = n$.

By homogenity, it suffices to consider the case $\|Z\|_\infty = 1$ and we must then show that

$$\operatorname{Sing}(AZB) \le \|Z^{-1}\|_\infty^{1/2} \operatorname{Sing}(ZAB) \tag{8}$$

(Z^{-1} still stands for the inverse or the generalized inverse). Define the operators on $\mathcal{H} \bigoplus \mathcal{H}$,

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{Z} = \begin{pmatrix} Z & Z^{1/2}(I-Z)^{1/2} \\ Z^{1/2}(I-Z)^{1/2} & I - Z \end{pmatrix}.$$

Then, $(\tilde{A}, \tilde{B})$ is a monotone pair and $\tilde{Z}$ is a projection. Theorem 2.3 entails that, for $1 \le k \le n$,

$$\begin{aligned} \mu_k^2(AZB) &= \mu_k^2(\tilde{A}\tilde{Z}\tilde{B}) \\ &\le \mu_k^2(\tilde{Z}\tilde{A}\tilde{B}) \\ &= \mu_k(\tilde{A}\tilde{B}\tilde{Z}\tilde{A}\tilde{B}) = \mu_k(ABZAB). \end{aligned}$$

But, $Z \le \|Z^{-1}\|_\infty Z^2$ (recall that $\|Z\|_\infty = 1$), consequently,

$$\mu_k^2(AZB) \le \|Z^{-1}\|_\infty \mu_k(ABZ^2AB)$$

meaning that (8) holds. $\square$

Proposition 2.6. *Let A, B be two commuting positive operators on a finite dimensional space and let $\mathcal{E}$ be a subspace.*

(1) *If (A, B) is monotone,*

$$\det A_{\mathcal{E}} \cdot \det B_{\mathcal{E}} \leq \det(AB)_{\mathcal{E}}.$$

(2) *If (A, B) is antimonotone and $\mathcal{E}$ is a hyperplane,*

$$\det A_{\mathcal{E}} \cdot \det B_{\mathcal{E}} \geq \det(AB)_{\mathcal{E}}.$$

Proof. (1) immediatly follows from Theorem 2.3. To prove (2) we first observe that for an operator X acting on $\mathcal{H}$ and a k-dimensional subspace $\mathcal{F}$,

$$|\det X_{\mathcal{F}}| = \| \wedge^k (X_{\mathcal{F}})\| = \|(\wedge^k X)_{\wedge^k \mathcal{F}}\|$$

where $\wedge^k$ stands for the kth antisymmetric tensor power. Consequently, letting $\mathcal{H}$ be the space on which A and B act and setting $\dim \mathcal{H} = n$, we have

$$\det A_{\mathcal{E}} \cdot \det B_{\mathcal{E}} = \|(\wedge^{n-1} A)_{\wedge^{n-1}\mathcal{E}}\| \cdot \|(\wedge^{n-1} B)_{\wedge^{n-1}\mathcal{E}}\|. \tag{9}$$

Since (A, B) is an antimonotone pair, there exist a family $\{E_j\}_{j=1}^n$ of mutually orthogonal rank one projections, a decreasing sequence $\{a_j\}_{j=1}^n$ and an increasing sequence $\{b_j\}_{j=1}^n$ such that

$$A = \sum_{j=1}^n a_j E_j \quad \text{and} \quad B = \sum_{j=1}^n b_j E_j.$$

We infer that

$$\wedge^{n-1} A = \sum_{j=1}^n (\Pi_{k\neq j} a_k)(\wedge_{k\neq j} E_k) \quad \text{and} \quad \wedge^{n-1} B = \sum_{j=1}^n (\Pi_{k\neq j} b_k)(\wedge_{k\neq j} E_k)$$

form an antimonotone pair of positive operators acting on $\wedge^{n-1}\mathcal{H}$. We note that $\wedge^{n-1}\mathcal{E}$ is a line in $\wedge^{n-1}\mathcal{H}$. Let h be a normalised vector spanning $\wedge^{n-1}\mathcal{E}$. We deduce from Lemma 1.2 that

$$\begin{aligned}\|(\wedge^{n-1} A)_{\wedge^{n-1}\mathcal{E}}\| \cdot \|(\wedge^{n-1} B)_{\wedge^{n-1}\mathcal{E}}\| &= \langle h, (\wedge^{n-1} A)h\rangle\langle h, (\wedge^{n-1} B)h\rangle \\ &\geq \langle h, (\wedge^{n-1} A \wedge^{n-1} B)h\rangle \\ &= \det(AB)_{\mathcal{E}}.\end{aligned}$$

Comparing with (9) we obtain the result. $\square$

Does Proposition 2.6(2) hold for every subspace ? Proposition 2.6(1) does not extend to other functionals such as the norm $\|\cdot\|_\infty$ (except if $\mathcal{E}$ is 1-dimensional), as shown by the next example.

Example 2.7. Consider the monotone pair (A, B) and the range $\mathcal{E}$ of the projection E,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad E = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

Then, one has $\|A_{\mathcal{E}}\|_\infty.\|B_{\mathcal{E}}\|_\infty = 6.668.. > 6.309.. = \|(AB)_{\mathcal{E}}\|_\infty$. The same inequality holds for the Schatten p-norms $\|\cdot\|_p$ when p runs over an interval (a, ∞) where a is a real lying into $(1, 2)$.

Theorem 2.3 suggests to compare $\mathrm{Sing}((AB)_{\mathcal{E}})$ and $\mathrm{Sing}(A_{\mathcal{E}}B_{\mathcal{E}})$ when A and B form a monotone pair of positive operators. We do not know whether the inequality $\mathrm{Sing}((AB)_{\mathcal{E}}) \geq \mathrm{Sing}(A_{\mathcal{E}}B_{\mathcal{E}})$ is valid, but our next proposition goes in this direction.

We denote by Γ the class of positive functions f defined on $[0, \infty)$ with $f(0) = 0$ and such that f is convex or concave. In particular the power functions $x \longrightarrow x^t$, $t > 0$, lie in Γ.

Proposition 2.8. *Let A, B be two commuting positive operators on a finite dimensional space and let $\mathcal{E}$ be a subspace. Assume that $B = f(A)$ for some $f \in \Gamma$. Then, there exist two unitary operators U and V on $\mathcal{E}$ such that*

$$|A_{\mathcal{E}}B_{\mathcal{E}}| \leq \frac{1}{2}\left(U(AB)_{\mathcal{E}}U^* + V(AB)_{\mathcal{E}}V^*\right).$$

We shall use a remarkable result of R. Bhatia and F. Kittaneh (see [1 p. 262]). This basic theorem, called the arithmetic-geometric mean operator inequality, states that for any operators A and B we have

$$\mathrm{Sing}(AB) \leq \frac{1}{2}\mathrm{Sing}(A^*A + BB^*).$$

Proof. By a limit argument we may assume that A is invertible and $f : [0, \infty) \longrightarrow [0, \infty)$ is bijective. By replacing if necessary B by A and A by $f^{-1}(B)$ we may then assume that f is convex. Therefore $x \longrightarrow f(x)/x$ is increasing on $[0, \infty)$. We write

$$EAEBE = EAE(f(A)A^{-1})^{1/2} \cdot (Af(A))^{1/2}E$$

and we apply the Bhatia-Kittaneh theorem to get a unitary W such that

$$|EAEBE| \leq \frac{1}{2}W\{(f(A)A^{-1})^{1/2}EAEAE(f(A)A^{-1})^{1/2} \\ +(Af(A))^{1/2}E(Af(A))^{1/2}\}W^*.$$

Since ZEZ^* and EZ^*ZE are unitarily congruent for any operator Z, there are unitaries U_0 and V_0 such that

$$|EAEBE| \leq \frac{1}{2}U_0(EAEf(A)A^{-1}EAE)U_0^* + \frac{1}{2}V_0(EAf(A)E)V_0^*.$$

Since $(A, f(A)A^{-1})$ is a monotone pair, Theorem 2.4 implies that, for a unitary U_1,

$$|EAEBE| \leq \frac{1}{2}U_1(EAf(A)E)U_1^* + \frac{1}{2}V_0(EAf(A)E)V_0^*.$$

We note that EU_1E and EV_0E can be viewed as contractions acting on $\mathcal{E}$; hence, we have

$$|A_{\mathcal{E}}B_{\mathcal{E}}| \leq \frac{1}{2}\left(U(AB)_{\mathcal{E}}U^* + V(AB)_{\mathcal{E}}V^*\right)$$

for some unitaries U and V acting on $\mathcal{E}$. $\square$

It is not surprising that specific inequalities hold for monotone or antimonotone pairs of the type (A^x, A^y).

Proposition 2.9. *Let A be a positive, invertible operator on a finite dimensional space and let $\mathcal{E}$ be a subspace. Then, we have*

$$\prod_{j=1}^{k}\mu_j((A^s)_{\mathcal{E}}(A^t)_{\mathcal{E}}) \leq \prod_{j=1}^{k}\mu_j((A^{s+t})_{\mathcal{E}}) \quad\text{and}\quad \prod_{j=1}^{k}\mu_j((A^s)_{\mathcal{E}}(A^{-t})_{\mathcal{E}}) \geq \prod_{j=1}^{k}\mu_j((A^{s-t})_{\mathcal{E}})$$

for every $1 \leq k \leq \dim \mathcal{E}$ and all $s, t > 0$. Therefore, for all unitarily invariant norms $\|\cdot\|$,

$$\|(A^s)_{\mathcal{E}}(A^t)_{\mathcal{E}}\| \leq \|(A^{s+t})_{\mathcal{E}}\| \quad\text{and}\quad \|(A^s)_{\mathcal{E}}(A^{-t})_{\mathcal{E}}\| \geq \|(A^{s-t})_{\mathcal{E}}\|.$$

Proof. The norm inequalities follow from the singular values inequalities and by standard antisymetric tensor arguments [5, Chapter 1] it suffices to prove the operator norm inequalities:

$$\|(A^s)_{\mathcal{E}}(A^t)_{\mathcal{E}}\|_\infty \leq \|(A^{s+t})_{\mathcal{E}}\|_\infty \tag{10}$$

and

$$\|(A^s)_{\mathcal{E}}(A^{-t})_{\mathcal{E}}\|_\infty \geq \|(A^{s-t})_{\mathcal{E}}\|_\infty. \tag{11}$$

(10) is a consequence of Proposition 2.10. We prove (11): Let E denote the projection onto $\mathcal{E}$. We have

$$\begin{aligned}\|EA^{s-t}E\| = \|EA^{(s-t)/2}\|^2 &\leq \|A^{s/2}EA^{-t/2}\|^2 \\ &= \|A^{s/2}EA^{-t}EA^{s/2}\| \leq \|EA^{-t}EA^sE\|\end{aligned}$$

where in the first inequality we use Proposition 3.5 below and in the second one the simple fact that $\|XY\| \leq \|YX\|$ for operators X, Y whose product XY is normal. $\square$

3. An inequality for the Hilbert-Schmidt norm

We first introduce the *hyponormality index* of an operator. It measures the lack of normality of an operator on a finite dimensional space $\mathcal{H}$. If $\mathcal{H}$ has an infinite dimension, then this number measures the lack of hyponormality. We recall that an operator X is hyponormal when $XX^* \leq X^*X$.

The hyponormality index $\nu(X)$ of an operator X is defined by

$$\nu^2(X) = \min\{a \in \mathrm{R}_+ \mid XX^* \leq aX^*X\}$$

Equivalently,

$$\nu(X) = \sup \frac{\|X^*h\|}{\|Xh\|} \quad (\text{and } \nu(0) = 1)$$

where the supremum runs over all the vectors h such that $\|Xh\| \neq 0$.

Thus, for an invertible operator X,

$$\nu(X) = \|X^*X^{-1}\|_\infty.$$

If X is no longer invertible,

$$\nu(X) = \lim_{\varepsilon \to 0} \|X^*(|X| + \varepsilon)^{-1}\|_\infty.$$

If $\nu(X)$ is finite, we have $XX^* \leq \nu^2(X)X^*X$, so $\|X^*\|_\infty \leq \nu(X)\|X\|_\infty$. This shows that $\nu(X) \in [1, \infty]$. Moreover $\nu(X) = 1$ if and only if X is hyponormal. In particular, if X is compact, then $\nu(X) = 1$ implies the normality of X. Indeed, it is easy to check that a compact hyponormal operator is normal; more generally Putnam inequality [4] ensures that a hyponormal operator whose spectrum has zero area is normal.

We now state the main result of this section.

Theorem 3.1. *Let Z be an operator whose hyponormality index $\nu(Z)$ is finite and let A and B be two commuting positive operators. Then:*

(1) *If A and B form a monotone pair,*

$$\|AZB\|_2 \leq \nu(Z)\|ZAB\|_2$$

where the constant $\nu(Z)$ is optimal. In particular, if Z is normal, or hyponormal, the inequality holds with $\nu(Z) = 1$.

(2) *If A and B form an antimonotone pair, if Z is normal and if either Z is in the Hilbert-Schmidt class or Z is self-adjoint or A or B is compact,*

$$\|AZB\|_2 \geq \|ZAB\|_2.$$

Example 3.2. The theorem is no longer true for the Schatten p-norms, $2 < p \leq \infty$. Consider

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

then, for any $p > 2$, $\|AZB\|_p > \|ZAB\|_p$. It would be desirable either to find counterexamples for $1 \leq p < 2$ or to extend the theorem to the Schatten p-norms with $p < 2$.

Proof. 1. Proof of the monotone case.

There exist $C \geq 0$ and two nondecreasing functions $f, g : \mathbf{R}_+ \to \mathbf{R}_+$ such that $A = f(C)$ and $B = g(C)$. One of the following inclusion relations must hold

$$\{t \ : \ f(t) = 0\} \subset \{t \ : \ g(t) = 0\}, \quad \{t \ : \ f(t) = 0\} \supset \{t \ : \ g(t) = 0\}$$

Suppose that the left inclusion holds (the other case leads to a similar proof). Then, for every $\varepsilon > 0$, we may find a positive $\tilde{A}$ and a nondecreasing $\Psi : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$(1-\varepsilon)\tilde{A} \leq A \leq (1+\varepsilon)\tilde{A} \quad \text{and} \quad (1-\varepsilon)\Psi(\tilde{A}) \leq B \leq (1+\varepsilon)\Psi(\tilde{A})$$

Moreover, we may require that A commutes with $\tilde{A}$ and B with $\Psi(\tilde{A})$. Therefore, it suffices to prove the theorem for monotone pairs of the type $(A, \Psi(A))$. We then proceed as follows.

1.1. First, we suppose that A has a finite rank, and we follow two steps.

- If r is a fixed positive real, the function f defined on $(-r, r)$ by

$$f(s) = \|A^{r+s} Z A^{r-s}\|_2^2$$

is convex. This can be seen by computing the second derivative of

$$s \to \operatorname{Tr} A^{r-s} Z^* A^{2(r+s)} Z A^{r-s},$$

or, more quickly by remarking that if $(e_1, \dots, e_n)$ is an orthonormal system associated to A's non-zero eingenvalues and if $z_{i\,j} = \langle e_i, Ze_j \rangle$, then

$$f(s) = \sum_{i,j} a_i^{2(r+s)} |z_{i\,j}|^2 a_j^{2(r-s)}$$

which is obviously convex.

Besides, f can be extended by continuity to r and $-r$. If we call E the projection onto the range of A, we have

$$f(-r) = \|EZA^{2r}\|_2^2 \leq \|ZA^{2r}\|_2^2$$

and

$$f(r) = \|A^{2r} Z E\|_2^2 \leq \|A^{2r} Z\|_2^2 \leq \nu^2(Z) \|ZA^{2r}\|_2^2.$$

The convexity of f entails $f(s) \leq \sup\{f(-r)\,;\, f(r)\}$, hence:

$$f(s) \leq \nu^2(Z) \|ZA^{2r}\|_2^2$$

which can also be written as:

$$\|A^s Z A^t\|_2 \leq \nu(Z)\|ZA^{s+t}\|_2 \qquad (0 \leq s,t).$$

• We turn to the main step of the proof. Let us show that for any nondecreasing function$\Psi : \mathbf{R}_+ \to \mathbf{R}_+$, we have

$$\|AZ\Psi(A)\|_2 \leq \nu(Z)\|ZA\Psi(A)\|_2,$$

or

$$\mathrm{Tr}\, Z^* A^2 Z \Psi^2(A) \leq \nu^2(Z)\mathrm{Tr}\, |Z|^2 A^2 \Psi^2(A)\,.$$

Setting $C = A^2$ and $\varphi = \Psi^2 \circ \sqrt{\ }$, we have to prove that for any positive operator of finite rank C and any positive nondecreasing $\varphi : \mathbf{R}_+ \to \mathbf{R}_+$,

$$\mathrm{Tr}\, Z^* C Z \varphi(C) \leq \nu^2(Z)\mathrm{Tr}\, |Z|^2 C\varphi(C)\ . \tag{$*$}$$

Let Φ be the set of all positive, nondecreasing functions φ which verify $(*)$ and

$$\mathrm{Tr}\, Z^* \varphi(C) Z C \leq \nu^2(Z)\mathrm{Tr}\, |Z|^2 C\varphi(C)\ . \tag{$**$}$$

Let us show that Φ coincides with the set of all positive, nondecreasing functions on $\mathbf{R}_+$. Φ is stable under:

(a) a linear combination with positive coefficients
(b) "dilation": $\varphi \in \Phi \Rightarrow \varphi_\lambda(x) = \varphi(\lambda x) \in \Phi$
(c) a pointwise limit
(d) if $\varphi \in \Phi$ is continuous and strictly increasing, with $\varphi(0) = 0$ and $\varphi(\infty) = \infty$; then the reciprocal function φ^{-1} is also an element of Φ.

By (a) (b) (c) we just have to prove that $\chi_{[1,\infty)} \in \Phi$. By the first step, the functions $x \to x^s$, $(s \geq 0)$, belong to Φ. So,

$$\varphi_n(x) = \frac{1}{n}x^n + x^{1/n}$$

is an element of Φ. Since $\lim_{n\to\infty} \varphi_n(0) = 0$, $\lim_{n\to\infty} \varphi_n(x) = 1$ for $x \in [0,1)$ and $\lim_{n\to\infty} \varphi_n(x) = \infty$ for $x \in (1,\infty)$, the reciprocal functions φ_n^{-1} pointwise converge to $\chi_{[1,\infty)}$ and the theorem is proved with (c) and (d).

1.2. Now, A no longer has a finite rank. If A can be diagonalized, there exists an increasing sequence $\{A_n\}$ of finite rank operators which pairwise commute and strongly converge to A. We have:

$$\|A_n Z\Psi(A_n)\|_2 \uparrow \|ZA\Psi(A)\|_2 \quad \text{and} \quad \|ZA_n\Psi(A_n)\|_2 \uparrow \|ZA\Psi(A)\|_2$$

which proves the theorem when A can be diagonalized. The general case can be deduced from it, because for any $\varepsilon > 0$ there exists A_ε which can be diagonalized and which commutes with A, such that

$$(1-\varepsilon)A \leq A_\varepsilon \leq (1+\varepsilon)A \quad \text{and} \quad (1-\varepsilon)\Psi(A) \leq \Psi(A_\varepsilon) \leq (1+\varepsilon)\Psi(A).$$

We still have to check that $\nu(Z)$ is the best constant. Let $\varepsilon > 0$ and let h be a norm-one vector for which

$$\frac{\|Z^*h\|}{\|Zh\|} \geq \nu(Z) - \varepsilon.$$

We take $A = h \otimes h$ and $\Psi(A) = I$, where I is the identity. Letting ε tend to 0 proves the claim.

2. Proof of the antimonotone case.

Let us observe that if A is a positive operator, Θ is a positive nondecreasing function defined on the spectrum of A and Z is a normal operator, the Hilbert-Schmidt norm inequality that we have proved in the monotone case implies the following trace norm inequality:

$$\|AZ\Theta(A)Z^*A\|_1 \leq \|ZA^2\Theta(A)Z^*\|_1. \tag{12}$$

2.1. Suppose first that Z is a normal Hilbert-Schmidt operator. For an arbitrarily small $\varepsilon > 0$, we may find a positive $\tilde{A}$ and a nonincreasing $\Psi : \mathbf{R}_+ \to \mathbf{R}_+$ such that

$$\|\tilde{A} - A\|_\infty \leq \varepsilon \quad \text{and} \quad \|\Psi(\tilde{A}) - B\|_\infty \leq \varepsilon.$$

Hence, it suffices to prove the theorem for antimonotone pairs of the type $(A, \Psi(A))$. Ψ is bounded and we can write Ψ^2 as the difference between a constant $k = \Psi^2(0)$ and an increasing function Θ: $\Psi^2 = k - \Theta$. So,

$$\begin{aligned}\|AZ\Psi(A)\|_2^2 =& \|AZ\Psi^2(A)Z^*A\|_1 \\ =& \|AZ(k-\Theta)Z^*A\|_1 = k\|AZZ^*A\|_1 - \|AZ\Theta(A)Z^*A\|_1,\end{aligned}$$

where the assumption that Z is Hilbert-Schmidt is *essential* for the last expression to be meaningful. Using the normality of Z and (12), we can conclude:

$$\|AZ\Psi(A)\|_2^2 \geq k\|AZ^*ZA\|_1 - \|ZA^2\Theta(A)Z^*\|_1 = \|ZA\Psi(A)\|_2^2.$$

2.2. We now assume that Z is self-adjoint. We may suppose that both A and B can be diagonalized, so that there exits an increasing sequence $\{E_n\}$ of finite rank projections, commuting with A and B, such that

$$\|AZB\|_2 = \lim \|AE_nZE_nB\|_2$$

By step 2.1 $\|AE_nZE_nB\|_2 \geq \|E_nZE_nAB\|_2$ and we deduce the result by letting n tend to the infinite.

2.3. Finally we assume that A is compact. There exists a sequence $\{A_n\}$ of positive Hilbert-Schmidt operators increasing to A such that (A_n, B) are antimonotone pairs. Therefore, $\|A_nZB\|_2 \uparrow \|AZB\|_2$ and $ZA_nB \to ZAB$ in Strong Operator Topology. Thanks to the SOT lower semi-continuity of the Hilbert-Schmidt norm,

$$\|ZAB\|_2 \leq \liminf \|ZA_nB\|_2.$$

So, it suffices to show the inequality when A is Hilbert-Schmidt. We may reproduce the argument of step 2.1, now using as an *essential* assumption the fact that A is Hilbert-Schmidt. $\square$

We say that a normal operator S is *semi-unitary* if its restriction to $\mathrm{ran}(S)$ is a unitary operator.

Corollary 3.2. *Let (A, B) be a monotone pair of positive operators and let S be a semi-unitary operator. Then,*

$$\|ASB\|_\infty \le \sqrt{2}\|SAB\|_\infty,$$

moreover, $\sqrt{2}$ is the best constant possible.

Proof. By a limit argument, we may assume that there is a norm one vector h such that

$$\|ASB\|_\infty = \|ASBh\|$$

Let E be the projection onto the range of S and set $f = Bh$. Since $\|Ef\| = \|Sf\|$ we obtain a semi-unitary operator R of rank 2 (or 1) such that $\mathrm{ran}(R) = \mathrm{Span}\{Ef, SEf\}$ and $REf = Sf$. Since $R^*R \le E$ we have $Sf = Rf$ and

$$\begin{aligned}
\|ASB\|_\infty = \|AEf\| &= \|ARf\| \\
&\le \|ARB\|_\infty \\
&\le \|ARB\|_2 \\
&\le \|RAB\|_2 \quad \text{(by Theorem 3.1)} \\
&\le \sqrt{2}\|RAB\|_\infty \quad (\mathrm{rank}(R) \le 2) \\
&\le \sqrt{2}\|EAB\|_\infty \quad (R^*R \le E^*E).
\end{aligned}$$

To see that the constant $\sqrt{2}$ can not be improved, we consider

$$A_n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & n \end{pmatrix} \quad B_n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S_n = \frac{1}{\sqrt{n^2+1}} \begin{pmatrix} 0 & n & 0 \\ n & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Then,

$$\lim_{n\to\infty} \frac{\|A_nS_nB_n\|_\infty}{\|S_nA_nB_n\|_\infty} = \sqrt{2},$$

thus the constant $\sqrt{2}$ can not be improved. $\square$

It would be interesting to find substitutes for Corollary 3.2 when S is a general normal (or hermitian, or positive) operator. In this direction we note the following straightforward consequence of Theorem 3.1: If (A, B) is a monotone pair of positive operators and Z is a normal operator on an n-dimensional space,

$$\|AZB\|_\infty \le \sqrt{n}\|ZAB\|_\infty.$$

If A and B form an antimonotone pair, then the reverse inequality holds.

For monotone or antimonotone pairs of the type (A^x, A^y), there are specific results. Although these results are not very original, we include them for the sake of completeness.

We denote by $\mathcal{I}$ an ideal of compact operators endowed with a unitarily invariant norm $\|\cdot\|$ making it a Banach space (one also says that $\mathcal{I}$ is a symmetrically normed ideal).

Lemma 3.4. *For an operator $Z \in \mathcal{I}$ and two invertible operators A and B, the map $t \to \||A|^t Z|B|^t\|$ is log-convex, equivalently:*

$$\|Z\| \leq \|A^* Z B^*\|^{1/2} \|A^{-1} Z B^{-1}\|^{1/2}$$

Proof. By unitary invariance of the norm, the analytic map $f(z) = |A|^z Z|B|^z$ satisfies $\|f(x+iy)\| = \|f(x)\|$ for all reals x and y. Hence the lemma is a straightforward application of the Banach space valued version of the Three lines theorem. $\square$

Proposition 3.5. *Let Z be an operator in $\mathcal{I}$ and let A be a positive operator. Then, for all $s, t \geq 0$, we have*

$$\|A^s Z A^t\| \leq \nu(Z)^{s/(s+t)} \|Z A^{s+t}\|.$$

Moreover, if A is invertible and $0 \leq s < t$,

$$\|A^s Z A^{-t}\| \geq \nu(Z)^{s/(s-t)} \|Z A^{s-t}\|.$$

Equivalently, if A is invertible and $0 \leq t < s$,

$$\|A^s Z A^{-t}\| \geq \nu(Z^*)^{s/(t-s)} \|Z A^{s-t}\|.$$

In the first inequality of the proposition, $\|ZA^{s+t}\| = 0 \Rightarrow ZA^{s+t} = 0 \Rightarrow A^s Z A^t = 0$ $\Rightarrow \|A^s Z A^t\| = 0$. Hence we may adopt the convention that $\nu(Z) \cdot 0 = 0$ when $\nu(Z) = \infty$.

Proof. By lemma 3.4, $f(r) = \|A^{s-r} Z A^{t+r}\|$ is log-convex on $]-t, s[$. From the lower semi-continuity of $\|\cdot\|$ in WOT we easily deduce that f can be extended by continuity to $-t$ and s with $f(-t) = \|A^{s+t} Z E\|$ and $f(s) = \|E Z A^{s+t}\|$ where $E = A^0$ is the support projection of A. Hence,

$$f(0) \leq f(-t)^{\frac{s}{s+t}} f(s)^{\frac{t}{s+t}}$$

where

$$f(0) = \|A^s Z A^t\|, \qquad f(-t) \leq \|A^{s+t} Z\|, \qquad f(s) \leq \|Z A^{s+t}\|$$

As $\|A^{s+t} Z\| \leq \nu(Z) \|Z A^{s+t}\|$, we get the first inequality.

To prove the second inequality, we consider the function $f(r) = \|A^{s-r} Z A^{r-t}\|$. Since $g(r) = \log f(r)$ is convex, the graphic representation of g shows us that the point $(0, g(0))$ is above the line passing by $(s, g(s))$ and $(t, g(t))$. Hence,

$$g(0) \geq g(s) + \frac{g(t) - g(s)}{t - s} \cdot (0 - s),$$

thus

$$\log f(0) \geq \log f(s) + \frac{-s}{t-s} \log \frac{f(t)}{f(s)},$$

or

$$\log f(0) \geq \log \left(f(s)^{\frac{t}{t-s}} f(t)^{\frac{-s}{t-s}} \right),$$

so

$$f(0) \geq f(s)^{\frac{t}{t-s}} f(t)^{\frac{-s}{t-s}}.$$

Then, using $f(0) = \|A^s Z A^{-t}\|$ and $f(t) = \|A^{s-t}Z\| \leq \nu(Z)\|ZA^{s-t}\| = \nu(Z)f(s)$, we deduce the result. The proof of the third inequality is similar. Actually it is not difficult to see that the second and third inequalities are equivalent. $\square$

Lemma 3.4 has been derived from a general principle of complex Analysis. It is also possible (and preferable) to deduce it from matrix theoretical technics. We then obtain a more precise result involving weak log-majorisation. Our next lemma follows such an approach and extends Lemma 3.4.

Lemma 3.6. *Let A, B, Z be operators on a finite dimensional space with A and B invertibles. Let p, q, r be positive reals with $1/p + 1/q = 1$. Then, for every unitarily invariant norm $\| \cdot \|$,*

$$\| \, |Z|^r \| \leq \| \, |A^* Z B^*|^{\frac{rp}{2}} \|^{1/p} \cdot \| \, |A^{-1} Z B^{-1}|^{\frac{rq}{2}} \|^{1/q},$$

in particular

$$\| \, |Z|^r \| \leq \| \, |A^* Z B^*|^r \|^{1/2} \cdot \| \, |A^{-1} Z B^{-1}|^r \|^{1/2}.$$

Proof . Note that, $\rho(\cdot)$ denoting the spectral radius,

$$\begin{aligned} \|Z\|_\infty^2 = \rho(Z^* Z) = \rho(BZ^* Z B^{-1}) &\leq \|BZ^* Z B^{-1}\|_\infty \\ &= \|BZ^* A A^{-1} Z B^{-1}\|_\infty \\ &\leq \|A^* Z B^*\|_\infty \|A^{-1} Z B^{-1}\|_\infty \end{aligned}$$

Thus the inequality is proved in case of the operator norm. An antisymetric tensor product argument then shows that, if $\{z_n\}$, $\{b_n\}$, $\{c_n\}$ denote the sequences of the respective singular values of Z, $A^* Z B^*$, $A^{-1} Z B^{-1}$ arranged in decreasing order and repeated according to their multiplicity, we have

$$\prod_{n=1}^{N} z_n \leq \prod_{n=1}^{N} b_n^{1/2} c_n^{1/2} \quad \text{for each integer } N.$$

This implies that $\{z_n^r\}$ is weakly majorized by $\{b_n^{r/2} c_n^{r/2}\}$, so we have

$$\begin{aligned} \| \, |Z|^r \| &= \Phi(z_1^r, z_2^r, \dots) \\ &\le \Phi(b_1^{r/2} c_1^{r/2}, b_2^{r/2} c_2^{r/2}, \dots) \\ &\le \Phi^{1/p}(b_1^{\frac{rp}{2}}, b_2^{\frac{rp}{2}}, \dots) \Phi^{1/q}(c_1^{\frac{rq}{2}}, c_2^{\frac{rq}{2}}, \dots) \\ &\le \| \, |A^* Z B^*|^{\frac{rp}{2}} \|^{1/p} \cdot \| \, |B^{-1} A C^{-1}|^{\frac{rq}{2}} \|^{1/q} \end{aligned}$$

by the Holder inequality for the symmetric gauge function Φ corresponding to the norm $\| \cdot \|$ (cf [1], p 87). □

Proposition 3.7. *Let A, B, Z be operators on a finite dimensional space and let p, q, r be positive reals with $1/p + 1/q = 1$. Then, for every unitarily invariant norm $\| \cdot \|$,*

$$\| \, |AZB|^r \| \le \| \, |A^* A Z|^{\frac{rp}{2}} \|^{1/p} \cdot \| \, |Z B B^*|^{\frac{rq}{2}} \|^{1/q}$$

in particular (Bhatia-Davis [2]),

$$\| \, |AZB|^r \| \le \| \, |A^* A Z|^r \|^{1/2} \cdot \| \, |Z B B^*|^r \|^{1/2}.$$

Proof. We may assume that both A and B are invertibles. Then we apply Lemma 3.6 with AZB, A and B^{*-1} in place of Z, A and B respectively. □

4. Addenda

4.1. If (A, B) is a monotone pair of positive operators, then we can find a positive operator C and two nondecreasing *continuous* functions f, g such that $A = f(C)$, $B = g(C)$. A similar statement holds for antimonotone pairs.

4.2. Several well known facts can be derived from Theorem 3.1 when Z is a unitary:

(1) The von Neumann trace inequality mentionned in the introduction: For two operators X, Y on an n-dimensional space,

$$|\mathrm{Tr}\, XY| \le \sum_k \mu_k(X) \mu_k(Y)$$

Recall that rearrangement inequalites of Hardy-Littlewood-Polya type can be derived from von Neumann's trace inequality.

(2) Given two Hermitian operators A, B on an n-dimensional space,

$$\sum_k |\lambda_k(A) - \lambda_k(B)|^2 \le \|A - B\|_2^2 \le \sum_k |\lambda_{n+1-k}(A) - \lambda_k(B)|^2.$$

Actually, these inequalities remain valid for all unitarily invariant norms [1, p. 71].

(3) Given an arbitrary basis $\{x_i\}_{i=1}^n$ of an n-dimensional space, the problem of how to find an orthonormal basis $\{u_i\}_{i=1}^n$ minimizing $\sum_i \|x_i - u_i\|^2$ was first

considered by the chemist Lowdin [1, p. 87]. This problem is equivalent to that of finding the best unitary approximant of an invertible operator X, in the Hilbert-Schmidt norm. If $X = U|X|$, one has

$$\min_{V^*V=I} \|X - V\|_2 = \|X - U\|_2.$$

This result remains true for all unitarily invariant norm [1, p. 276].

4.3. Theorem 3.1 entails the following inequalities for real and imaginary parts:

Corollary 4.1. *Let A, B be commuting, positive operators and let Z be a normal Hilbert-Schmidt operator.*

(a) *If (A, B) is monotone,*

$$\|\mathrm{Re}(AZB)\|_2 \leq \|\mathrm{Re}(ZAB)\|_2.$$

(b) *If (A, B) is antimonotone,*

$$\|\mathrm{Re}(AZB)\|_2 \geq \|\mathrm{Re}(ZAB)\|_2.$$

The same results hold for the imaginary parts.

Proof. As $\mathrm{Im}(AZB) = -\mathrm{Re}(A(iZ)B)$ and $\mathrm{Im}(ZAB) = -\mathrm{Re}((iZ)AB)$, it suffices to prove the corollary for the real parts. We do case (a), case (b) being similar. We have

$$\begin{aligned} 4\|\mathrm{Re}(AZB)\|_2^2 &= 2\mathrm{Tr}\, AZB^2Z^*A + \mathrm{Tr}\,(AZB)^2 + \mathrm{Tr}\,(BZ^*A)^2 \\ &= 2\|AZB\|_2^2 + \mathrm{Tr}\,(AZB)^2 + \mathrm{Tr}\,(BZ^*A)^2 \\ &\leq 2\|ZAB\|_2^2 + \mathrm{Tr}\,(AZB)^2 + \mathrm{Tr}\,(BZ^*A)^2 \\ &= 4\|\mathrm{Re}(ZAB)\|_2^2 \end{aligned}$$

where the last equality follows from commutatibility of A and B. $\square$

References

[1] R. Bhatia, Matrix Analysis, Springer, Germany, 1996

[2] R. Bhatia, C. Davis, A Cauchy-Schwarz inequality for operators with applications, Linear Algebra Appl. 223 (1995) 119-129

[3] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, Cambridge, 1985

[4] C.R. Putnam, An inequality for the area hyponormal spectra, Math. Z. 116 (1970) 323-330

[5] B. Simon, Trace Ideals and Their Applications LMS lecture note, 35 Cambridge Univ. Press, Cambridge, 1979

Chapter 5

Dilations for strict contractions

Introduction

By an operator, we mean an element of the algebra $\mathrm{L}(\mathcal{H})$ of all bounded linear operators acting on the usual (i.e. complex, separable, infinite dimensional) Hilbert space $\mathcal{H}$. We will denote by the same letter a projection and the corresponding subspace. Thus, if F is a projection and A is an operator, we denote by A_F the compression of A by F, that is the restriction of FAF to the subspace F. Given a total sequence of nonzero mutually orthogonal projections $\{E_n\}$, we consider the pinching

$$\mathcal{P}(A) = \sum_{n=1}^{\infty} E_n A E_n = \bigoplus_{n=1}^{\infty} A_{E_n}.$$

If $\{A_n\}$ is a sequence of operators acting on separable Hilbert spaces with A_n unitarily equivalent to A_{E_n} for all n, we also naturally write $\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_n$. The main result of this chapter can then be stated as:

Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of operators acting on separable Hilbert spaces. Assume that $\sup_n \|A_n\|_\infty < 1$*. Then, we have a pinching*

$$\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_n$$

for any operator A whose essential numerical range contains the unit disc.

This result is proved in the second section of the chapter. We have included a first section concerning some well-known properties of the essential numerical range. The third section is concerned with some related results.

1. Properties of the essential numerical range

We denote by $\langle \cdot, \cdot \rangle$ the inner product (linear in the second variable), by co$S$ the convex hull of a subset S of the complex plane $\mathbf{C}$. The numerical range of an operator A is

$$W(A) = \{\langle h, Ah \rangle \mid \|h\| = 1\}.$$

We denote by $\overline{W}(A)$ the closure of $W(A)$. The celebrated Hausdorff-Toeplitz theorem (cf [6] chapter 1) states that $W(A)$ is convex. A corollary is Parker's theorem ([6], p.20): Given an n by n matrix A, there is a matrix B unitarily equivalent to A and with all its diagonal elements equal to $\mathrm{Tr}A/n$.

Let us give three equivalent definitions of the essential numerical range of A, denoted by $W_e(A)$.

(1) $W_e(A) = \cap \overline{W}(A+K)$, the intersection running over the compact operators K

(2) Let $\{E_n\}$ be any sequence of finite rank projections converging strongly to the identity and denote by B_n the compression of A to the subspace $E_n^{\perp}$. Then $W_e(A) = \cap_{n\geq 1}\overline{W}(B_n)$

(3) $W_e(A) = \{\lambda \mid \text{there is an orthonormal system } \{\mathrm{e_n}\}_{\mathrm{n=1}}^{\infty} \text{ with } \lim\langle \mathrm{e_n}, \mathrm{Ae_n}\rangle = \lambda\}$.

It follows that $W_e(A)$ is a compact convex set containing the essential spectrum of A, $Sp_e(A)$. The equivalence between these definitions has been known since the early seventies if not sooner (see for instance [1]). The very first definition of $W_e(A) =$ is (1); however (3) is also a natural notion and easily entails convexity and compactness of the essential numerical range. We mention the following result of Chui-Smith-Smith-Ward [4] :

Proposition 1.1. *Every operator A admits some compact perturbation $A+K$ for which $W_e(A) = \overline{W}(A+K)$.*

Another characterization of the essential numerical range of A is

$$W_e(A) = \{\lambda \mid \text{there is a basis } \{\mathrm{e_n}\}_{\mathrm{n=1}}^{\infty} \text{ with } \lim\langle \mathrm{e_n}, \mathrm{Ae_n}\rangle = \lambda\}.$$

Let us check the equivalence between our definition (3) with orthonormal system and the above identity which seems to be due to Q. F. Stout [11]. Let $\{x_n\}_{n=1}^{\infty}$ be an orthonormal system such that $\lim_{n\to\infty}\langle x_n, Ax_n\rangle = \lambda$. If $\mathrm{span}\{x_n\}_{n=1}^{\infty}$ is of finite codimension p we immediately get a basis $e_1, \ldots, e_p; e_{p+1} = x_1, \ldots, e_{p+n} = x_n, \ldots$ such that $\lim_{n\to\infty}\langle e_n, Ae_n\rangle = \lambda$. If $\mathrm{span}\{x_n\}_{n=1}^{\infty}$ is of infinite codimension, we may complete this system with $\{y_n\}_{n=1}^{\infty}$ in order to obtain a basis. Let P_j be the subspace spanned by y_j and $\{x_n | 2^{j-1} \leq n < 2^j\}$. By Parker's theorem, there is a basis of P_j, say $\{e_l^j\}_{l\in\Lambda_j}$, with

$$\langle e_l^j, Ae_l^j\rangle = \frac{1}{\mathrm{dim}P_j}\mathrm{Tr}AP_j.$$

Since

$$\frac{1}{\mathrm{dim}P_j}\mathrm{Tr}AP_j \to \lambda \quad \text{as} \quad j \to \infty,$$

we may index $\{e_l^j\}_{j\in\mathrm{N};l\in\Lambda_j}$ in order to obtain a basis $\{f_n\}_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty}\langle f_n, Af_n\rangle = \lambda.$$

The essential numerical range appears closely related to the diagonal set of A which we define by

$$\Delta(A) = \{\lambda \mid \text{there is a basis } \{\mathrm{e_n}\}_{\mathrm{n=1}}^{\infty} \text{ with } \langle \mathrm{e_n}, \mathrm{Ae_n}\rangle = \lambda\}.$$

The next result is a straightforward consequence of a lemma of Peng Fan [5]. A real operator means an operator acting on a real Hilbert space and int$X$ denotes the interior of $X \subset \mathbf{C}$.

Proposition 1.2 *Let A be an operator. Then* $\operatorname{int} W_e(A) \subset \Delta(A) \subset W_e(A)$. *Consequently, an open set $\mathcal{U}$ is contained in $\Delta(A)$ if and only if there is a basis $\{e_n\}_{n=1}^{\infty}$ such that $\mathcal{U} \subset \mathrm{co}\{\langle e_k, Ae_k\rangle | k \geq n\}$ for all n. Finally, the diagonal set of a real operator is symetric about the real axis. (For A self-adjoint, the result holds with* int *denoting the interior of subsets of* $\mathbf{R}$.*)*

Curiously enough, it seems difficult to answer the following questions: Is the diagonal set always a (possibly vacuous) convex set ? Is there an operator of the form self-adjoint + compact with a disconnected diagonal set ?

An elementary but very important property of $W(\cdot)$ is the so named projection property: Re$W(A) = W(\mathrm{Re}A)$ (see [6] p. 9), where Re stands for real part. $W_e(\cdot)$ also has this property. This result and the Hausdorff-Toeplitz Theorem are the keys to prove the following fact:

Proposition 1.3. *Let A be an operator.*

(1) *If $W_e(A) \subset W(A)$ then $W(A)$ is closed.*

(2) *There exist normal, finite rank operators R of arbitrarily small norm such that $W(A+R)$ is closed.*

Proof. Assertion (1) is due to J. S. Lancaster [8]. We prove the second assertion and implicitly prove Lancaster's result. We may find an orthonormal system $\{f_n\}$ such that the closure of the sequence $\{\langle f_n, Af_n\rangle\}$ contains the boundary $\partial W_e(A)$. Fix $\varepsilon > 0$. It is possible to find an integer p and scalars z_j, $1 < j < p$, with $|z_j| < \varepsilon$, such that :

$$\mathrm{co}\{\langle f_j, Af_j\rangle + z_j \; |1 < j < p\} \supset \partial W_e(A).$$

Thus, the finite rank operator $R = \sum_{1<j<p} z_j f_j \otimes f_j$ has the property that W$(A+R)$ contains $W_e(A)$.

We need this operator R. Indeed, setting $X = A + R$, we also have W$(X) \supset W_e(X)$. We then claim that W(X) is closed (this claim implies assertion (1)). By the contrary, there would exist $z \in \partial\overline{W}(X) \setminus W_e(X)$. Furthermore, since $\overline{W}(X)$ is the convex hull of its extreme points, we could assume that such a z is an extreme of $\overline{W}(X)$. By suitable rotation and translation, we could assume that $z = 0$ and that the imaginary axis is a line of support of $\overline{W}(X)$. The projection property for $W(\cdot)$ would imply that $W(\mathrm{Re}X) = (x, 0)$ for a certain negative number x, so that

$0 \in W_e(\mathrm{Re}X)$. Thus we would deduce from the projection property for $W_e(\cdot)$ that $0 \in W_e(X)$: a contradiction. $\square$

The perturbation R in Proposition 1.3 can be taken real if A is real. We mention that the set of operators with nonclosed numerical ranges is not dense in $\mathrm{L}(\mathcal{H})$. Proposition 3 improves the following result of I.D. Berg and B. Sims [3]: operators which attain their numerical radius are norm dense in $\mathrm{L}(\mathcal{H})$. A motivation for Berg and Sims was the following fact: Given an arbitrary operator A, a small rank one perturbation of A yields an operator which attains its norm. Indeed, the polar decomposition allows us to assume that A is positive, an easy case when reasoning as in the proof of Proposition 1.3.

Let us say that a convex set in $\mathbf{C}$ is relatively open if either it is a single point, an open segment or an usual open set. Using similar methods as in the previous proof, or applying Propositions 1.2 and 1.3, we obtain

Proposition 1.4. *For an operator A the following assertions are equivalent*

(1) $W(A)$ *is relatively open.*

(2) $\Delta(A) = W(A)$.

From the previous results we may derive some information about $W(\cdot)$, $W_e(\cdot)$ and $\Delta(\cdot)$ for various classes of operators:

(a) Let S be either the unilateral or bilateral Shift, then $\Delta(S) = W(S)$ is the open unit disc. More generally Stout showed [10] that weighted periodic shifts S have open numerical ranges; therefore $\Delta(S) = W(S)$.

(b) There exist a number of Toeplitz operators with open numerical range. See the papers by E. M. Klein [7] and by J. K. Thukral [11].

(c) Let X be an operator lying in a C^*-subalgebra of $L(\mathcal{H})$ with no finite dimensional projections. Then for any real θ, $\overline{W}(\mathrm{Re}e^{\mathrm{i}\theta}X) = W_e(\mathrm{Re}e^{\mathrm{i}\theta}X)$. From the projection property for $W(\cdot)$ and $W_e(\cdot)$ we infer that $W_e(X) = \overline{W}(X)$.

(d) Let X be an essentially normal operator : $X^*X - XX^*$ is compact. It is known that $W_e(X) = \mathrm{co}Sp_e(X)$. Indeed, for such an operator the essential norm equals the essential spectral radius : $\|X\|_e = \rho_e(X)$. Denoting by $w_e(X)$ the essential numerical radius of X we deduce that $\|X\|_e = w_e(X) = \rho_e(X)$. Note that $e^{\mathrm{i}\theta}X + \mu I = Y$ is also an essentially normal operator for any $\theta \in \mathbf{R}$ and $\mu \in \mathbf{C}$. Let z be an extremal point of $W_e(X)$. With suitable θ and μ we have $e^{\mathrm{i}\theta}z + \mu = w_e(Y) = \max\{|y|,\ y \in W_e(Y)\}$, the maximum being attained at the single point $e^{\mathrm{i}\theta}z + \mu$. Since $\mathrm{co}Sp_e(Y) \subset W_e(Y)$ and $\rho_e(Y) = w_e(Y)$, this implies that $e^{\mathrm{i}\theta}z + \mu \in Sp_e(Y)$. Hence $z \in Sp_e(Y)$, so that $W_e(X) = \mathrm{co}Sp_e(X)$.

2. The pinching theorem

Recall that one way to define the essential numerical range of an operator A is:

$$W_e(A) = \{\lambda \mid \text{there is an orthonormal system } \{e_n\}_{n=1}^{\infty} \text{ with } \lim\langle e_n, Ae_n\rangle = \lambda\}.$$

It is then very easy to check that $W_e(A)$ is a compact convex set. Moreover $W_e(A)$ contains the open unit disc $\mathcal{D}$ if and only if there is a basis $\{e_n\}_{n=1}^{\infty}$ such that

$$\mathrm{co}\{\langle e_k, Ae_k\rangle \mid k > n\} \supset \mathcal{D}$$

for all n.

We state:

Theorem 2.1. *Let A be an operator with $W_e(A) \supset \mathcal{D}$ and let $\{A_n\}_{n=1}^{\infty}$ be a sequence of operators such that $\sup_n \|A_n\|_\infty < 1$. Then, we have a pinching*

$$\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_n.$$

(*If A and $\{A_n\}_{n=1}^{\infty}$ are real, then we may take a real pinching*).

We need two lemmas. The first one is Theorem 2.1 for a single strict contraction:

Lemma 2.2. *Let A be an operator with $W_e(A) \supset \mathcal{D}$ and let X be a strict contraction. Then there exists a projection E such that $A_E = X$.*

The second Lemma is a refined version of the first one:

Lemma 2.3. *Let $a \geq 1$ and $1 > \rho > 0$ be two constants. Let h be a norm one vector, let X be a strict contraction with $\|X\|_\infty < \rho$ and let B be an operator with $\|B\|_\infty \leq a$ and $W_e(B) \supset \mathcal{D}$. Then, there exist a number $\varepsilon > 0$, only depending on ρ and a, and a projection E such that:*

(i) $\dim E = \infty$ *and* $B_E = X$,

(ii) $\dim E^\perp = \infty$, $W_e(B_{E^\perp}) \supset \mathcal{D}$ *and* $\|Eh\| \geq \varepsilon$.

Proof of Theorem 2.1. The proof is organized in five steps:
Step 1. Some preliminaries are given.
Step 2. Proof of Lemma 2.2 in the special case when X is normal, diagonalizable.
Step 3. Proof of Lemma 2.2 in the general case.
Step 4. Proof of Lemma 2.3.
Step 5. Conclusion.

1. *Preliminaries*

We shall use a sequence $\{V_k\}_{k\geq 1}$ of orthogonal matrices acting on spaces of dimensions 2^k. This sequence is built up by induction:

$$V_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{then} \quad V_k = \frac{1}{\sqrt{2}}\begin{pmatrix} V_{k-1} & V_{k-1} \\ -V_{k-1} & V_{k-1} \end{pmatrix} \quad \text{for } k \geq 2.$$

Given a Hilbert space $\mathcal{G}$ and a decomposition

$$\mathcal{G} = \bigoplus_{j=1}^{2^k} \mathcal{H}_j \quad \text{with } \mathcal{H}_1 = \cdots = \mathcal{H}_{2^k} = \mathcal{H},$$

we may consider the unitary (orthogonal) operator on $\mathcal{G}$: $W_k = V_k \bigotimes I$, where I denotes the identity on $\mathcal{H}$,

Now, let $B : \mathcal{G} \to \mathcal{G}$ be an operator which, with respect to the above decomposition of $\mathcal{G}$, has a block diagonal matrix

$$B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_{2^k} \end{pmatrix}.$$

We observe that the block matrix representation of $W_k B W_k^*$ has its diagonal entries all equal to

$$\frac{1}{2^k}\left(B_1 + \ldots B_{2^k}\right).$$

So, the orthogonal operators W_k allow us to pass from a block diagonal matrix representation to a block matrix representation in which the diagonal entries are all equal.

2. *Proof of Lemma 2.2 when X is normal, diagonalizable.*

Let $\{\lambda_n(X)\}_{n\geq 1}$ be the eigenvalues of X repeated according to their multiplicities. Since $|\lambda_n(X)| < 1$ for all n and $W_e(A) \supset \mathcal{D}$, we may find a norm one vector e_1 such that $\langle e_1, Ae_1\rangle = \lambda_1(T)$. Let $F_1 = [\text{span}\{e_1, Ae_1, A^*e_1\}]^\perp$. As F_1 is of finite codimension, $W_e(A_{F_1}) \supset \mathcal{D}$. So, there exists a norm one vector $e_2 \in F_1$ such that $\langle e_2, Ae_2\rangle = \lambda_2(T)$. Next, we set $F_2 = [\text{span}\{e_1, Ae_1, A^*e_1, e_2, Ae_2, A^*e_2\}]^\perp, \ldots$. If we go on like this, we exhibit an orthonormal system $\{e_n\}_{n\geq 1}$ such that, setting $E = \text{span}\{e_n\}_{n\geq 1}$, we have $A_E = X$.

3. *Proof of Lemma 2.2 in the general case.*

The contraction $Y = (1/\|X\|_\infty)X$ can be dilated in a unitary

$$U = \begin{pmatrix} Y & -(I - YY^*)^{1/2} \\ (I - Y^*Y)^{1/2} & Y^* \end{pmatrix}$$

thus X can be dilated in a normal operator $N = \|X\|_\infty U$ with $\|N\|_\infty < \rho$. This permits to restrict to the case when X is a normal strict contraction. So, let X

be a normal operator with $\|X\|_\infty < \rho < 1$. We remark with the Berg-Weyl-von Neumann theorem [2], that X can be written as

$$X = D + K \tag{1}$$

where D is normal diagonalizable, $\|D\|_\infty = \|X\|_\infty < \rho$, and K is compact with an arbitrarily small norm. Let $K = \mathrm{Re}K + \mathrm{i}\mathrm{Im}K$ be the Cartesian decomposition of K. We can find an integer l, a real α and a real β such that decomposition (1) satisfies:
a) the operators αD, $\beta\mathrm{Re}K$, $\beta\mathrm{Im}K$ are dominated in norm by ρ,
b) there are positive integers m, n with $2^l = m + 2n$ and

$$X = \frac{1}{2^l}(m\alpha D + n\beta\mathrm{Re}K + n\beta\mathrm{i}\mathrm{Im}K). \tag{2}$$

More precisely we can take any l such that $[2^l/(2^l-2)].\|X\|_\infty < \rho$. Next, assuming $\|K\|_\infty < \rho/2^l$, we can take $m = 2^l - 2$, $n = 1$, $\alpha = 2^l/(2^l-2)$ and $\beta = 2^l$.

Let then T be the diagonal normal operator acting on the space

$$\mathcal{G} = \bigoplus_{j=1}^{2^l} \mathcal{H}_j \quad \text{with } \mathcal{H}_1 = \cdots = \mathcal{H}_{2^l} = \mathcal{H},$$

and defined by

$$T = \left(\bigoplus_{j=1}^{m} D_j\right) \bigoplus \left(\bigoplus_{j=m+1}^{m+n} R_j\right) \bigoplus \left(\bigoplus_{j=m+n+1}^{2^l} S_j\right)$$

where $D_j = \alpha D$, $S_j = \beta\mathrm{Re}K$ and $S_j = \beta\mathrm{i}\mathrm{Im}K$.

We note that $\|T\|_\infty < \rho < 1$ and that the operator $W_l T W_l^*$, represented in the preceding decomposition of $\mathcal{G}$, has its diagonal entries all equal to X by (2). Hence, applying the preceding step to T yields Lemma 2.2.

4. Proof of Lemma 2.3.

Let $a \geq 1$ and let $1 > \rho > 0$ be two constants. We take an arbitrary norm one vector h and any operator B satisfying to the assumptions of Lemma 2.3. We can show, using the same reasoning as that applied in the above Step 2, that we have an orthonormal system $\{f_n\}_{n\geq 0}$, with $f_0 = h$, such that:

a) $\langle f_{2j}, Bf_{2j}\rangle = 0$ for all $j \geq 1$.
b) $\{\langle f_{2j+1}, Bf_{2j+1}\rangle\}_{j\geq 0}$ is a dense sequence in $\mathcal{D}$.
c) If $F = \mathrm{span}\{f_j\}_{j\geq 0}$, then B_F is the normal operator

$$\sum_{j\geq 0} \langle f_j, Bf_j\rangle f_j \otimes f_j.$$

Setting $F_0 = \mathrm{span}\{f_{2j}\}_{j\geq 0}$ and $F_0' = \mathrm{span}\{f_{2j+1}\}_{j\geq 0}$, we then have:

a) With respect to the decomposition $F = F_0 \bigoplus F_0'$, B_F can be written

$$B_F = \begin{pmatrix} B_{F_0} & 0 \\ 0 & B_{F_0'} \end{pmatrix}.$$

b) $W_e(B_{F_0'}) \supset \mathcal{D}$ and $h \in F_0$.

We can then write a decomposition of F_0', $F_0' = \bigoplus_{j=1}^{\infty} F_j$ where for each index j, F_j commutes with B_F and $W_e(B_{F_j}) \supset \mathcal{D}$; so that the decomposition $F = \bigoplus_{j=0}^{\infty} F_j$ yields a representation of B_F as a block diagonal matrix,

$$B_F = \bigoplus_{j=0}^{\infty} B_{F_j}.$$

Since $W_e(B_{F_j}) \supset \mathcal{D}$ when $j \geq 1$, the same reasoning as in Step 3 entails that for any sequence $\{X_j\}_{j\geq 0}$ of strict contractions we have decompositions (†) $F_j = G_j \bigoplus G_j'$ allowing us to write, for $j \geq 1$,

$$B_{F_j} = \begin{pmatrix} X_j & * \\ * & * \end{pmatrix}.$$

Since $\|X\|_\infty < \rho < 1$ and $\|B\|_\infty \leq a$, we can find an integer l only depending on ρ and a, as well as strict contractions $X_1, \ldots, X_{2^l}$, such that

$$X = \frac{1}{2^l}\left(B_{F_0} + \sum_{j=1}^{2^l-1} X_j\right). \tag{3}$$

Considering decompositions (†) adapted to these X_j, we set

$$G = F_0 \bigoplus \left(\bigoplus_{j=1}^{2^l-1} G_j\right).$$

With respect to this decomposition,

$$B_G = \begin{pmatrix} B_{F_0} & & & \\ & X_1 & & \\ & & \ddots & \\ & & & X_{2^l-1} \end{pmatrix}.$$

Then we deduce from (3) that the block matrix $W_l B_G W_l^*$ has its diagonal entries all equal to X.

Summary: $h \in G$ and there exists a decomposition $G = \bigoplus_{j=1}^{2^l} E_j$, in which l depends only on ρ and a, such that $B_{E_j} = X$ for each j. Thus we have an integer j_0 such that, setting $E_{j_0} = E$, we have

$$B_E = X \quad \text{and} \quad \|Eh\| \geq \frac{1}{\sqrt{2^l}}.$$

Taking $\varepsilon = 1/\sqrt{2^l}$ ends the proof of Lemma 2.3.

4. Conclusion.

Fix a dense sequence $\{h_n\}$ in the unit sphere of $\mathcal{H}$ and set $a = \|A\|_\infty$. We claim that the statement (i) and (ii) of Lemma 2.3 ensure that there exists a sequence of mutually orthogonal projections $\{E_j\}$ such that, setting $F_n = \sum_{j\le n} E_j$, we have for all integers n:

$(*)$ $A_n = A_{E_n}$ and $W_e(A_{F_n^\perp}) \supset \mathcal{D}$ (so $\dim F_n^\perp = \infty$),

$(**)$ $\|F_n h_n\| \ge \varepsilon$.

In Lemma 2.3, set $a = \|A\|_\infty$. Replacing B by A, Lemma 2.3 proves $(*)$ and $(**)$ for $n = 1$. Suppose this holds for an $N \ge 1$. Let $\nu(N) \ge N+1$ be the first integer for which $F_N h_{\nu(N)} \ne 0$. Note that $\|A_{F_N^\perp}\|_\infty \le \|A\|_\infty$. We apply Lemma 2.3 to $B = A_{F_N^\perp}$, $X = A_{N+1}$ and $h = F_N h_{\nu(N)}/\|F_N h_{\nu(N)}\|$. We then deduce that $(*)$ and $(**)$ are still valid for $N+1$. Therefore $(*)$ and $(**)$ hold for all n. Denseness of $\{h_n\}$ and $(**)$ show that F_n strongly increases to the identity I so that $\sum_{j=1}^\infty E_j = I$ as required. $\square$

Corollary 2.2. *Let A be an operator with $W_e(A) \supset \mathcal{D}$. For any strict contraction X, there is an isometry V such that $X = V^*AV$.*

We use the strict inclusion notation $X \subset\subset Y$ for subsets X, Y of $\mathbf{C}$ to mean that there is an $\varepsilon > 0$ such that $\{x + z \mid x \in X,\ |z| < \varepsilon\} \subset Y$.

Theorem 2.3. *Let A be an operator and let $\{A_n\}_{n=1}^\infty$ be a sequence of normal operators. If $\cup_{n=1}^\infty W(A_n) \subset\subset W_e(A)$ then we have a pinching*

$$\mathcal{P}(A) = \bigoplus_{n=1}^{\infty} A_n.$$

(For self-adjoint operators, this result holds with the strict inclusion of $\mathbf{R}$.*)*

Sketch of proof. Let N be a normal operator with $W(N) \subset\subset W_e(A)$. If N is diagonalizable, reasonning as in the proof of Theorem 1, we deduce that N can be realized as a compression of A. If N is not diagonalizable we may assume that $0 \in W_e(A)$. Thanks to the Berg-Weyl-von Neumann Theorem and still reasonning as in the proof of Theorem 1 we again deduce that N is a compression of A. Finally, the strict containment assumption allows us to get the wanted pinching.

To finish this section, we mention that we can not drop the assumption that the strict contractions A_n of Theorem 1 are uniformly bounded in norm by a real < 1. This observation is equivalent to the fact that we can not delete the strict containment assumption in Theorem 2:

Let P be a halving projection ($\dim P = \dim P^\perp = \infty$), so $W_e(P) = [0, 1]$. Then the sequence $\{1 - 1/n^2\}_{n\ge1}$ can not be realized as the entries of the main diagonal of a matrix representation of P. To check that, we note that the positive operator

$I-P$ would be in the trace-class: a contradiction. (Recall that a positive operator with a summable diagonal is trace class.)

3. Related results

3.1. Open numerical ranges. Diagonal sets

Remark 3.1. Let A be an operator whose numerical range is open. There exists an infinite projection E such that $\mathrm{W}(A) = \mathrm{W}(A_E) = \mathrm{W}(A_{E^\perp})$.

Question 3.2. Given a compact convex subset Γ of C, does there exist a normal diagonalizable operator N such that $\Gamma = \Delta(N)$?

Question 3.3. Is there, for an arbitrary operator A, a compact operator L such that $\Delta(A+L) = \mathrm{int}\Delta(A)$?

We have very poor information concerning the topologigal properties of the boundary of a diagonal set. We may note that:

Proposition 3.4. *The set of hilbertian bases has a natural Polish space structure. The diagonal set of an operator is an analytical set.*

Proof. The *-strong topology confers the structure of a Polish space (compatible with the group structure) to the unitary group $\mathcal{U}$ of $\mathrm{L}(\mathcal{H})$. Fix a basis $\mathcal{E}$ and consider the map $U \longrightarrow U(\mathcal{E})$ between $\mathcal{U}$ and the set of bases. This correspondance confers the structure of a Polish space to the set of bases.

Let A be a contraction on $\mathcal{H}$. Let $\mathcal{D}$ denote the closed unit disc of $\mathbf{C}$. The subset $\mathcal{C} \subset \mathcal{D}^{\mathrm{N}}$ consisting of constant sequences is obviously a closed subset of $\mathcal{D}^{\mathrm{N}}$ endowed with the pointwise topology. The map

$$\phi\colon \{e_n\}_{n=1}^{\infty} \longrightarrow \{\langle e_n, Ae_n\rangle\}_{n=1}^{\infty}$$

from the set of hilbertian bases into $\mathcal{D}^{\mathrm{N}}$ is continuous. So, $\phi^{-1}\mathcal{C}$ is closed. Let $\nu(s)$ be the value of the constant sequence s. Then $\Delta(A) = \nu \circ \phi[\phi^{-1}\mathcal{C}]$ is an analytical set. $\diamondsuit$

D.A. Herrero [?] proved the following sufficient condition for an operator A to ensure that $\mathrm{W}(A)$ is open :

If h, Ah, A^*h are linearly independant for all h, $\|h\| = 1$, then $\mathrm{W}(A)$ *is open .* (Hc)

Proof of Herrero's criteria. We first note that A satisfies (Hc) if and only if A_θ satisfies (Hc) for any $\theta \in \mathbf{R}$, and also, if and only if $\lambda I + A$ satisfies (Hc) for any $\lambda \in \mathbf{C}$. We have to show that an operator A with a non open numerical range can not satisfy (Hc). With a suitable choice of θ and λ, we may assume that $B = \lambda I + A_\theta$ has the following property : $\mathrm{W}(B)$ lies in the left halfplane and has

0 as an extreme point. Thus, there exists a norm one vector h such that $\mathrm{Re}Bh = \max\{\langle f, Bf\rangle \mid \ \|f\| = 1\} = 0$. This entails that $\mathrm{Re}Bh = 0$. Consequentely nor B neither A satisfy (Hc).

3.2. Appendix on trace-class operators

For a trace class operator A with a nonzero trace, the diagonal set is empty. However we have the following result:

Proposition 3.5 *Let A be a trace class operator. There is a hilbertian basis $\{e_n\}_{n=1}^{\infty}$ and a sequence $\{a_n\}_{n=1}^{\infty}$ in R_+^* such that $\langle e_n, Ae_n\rangle = a_n \mathrm{Tr}\, A$ for all n. Equivalently,*

$$|\mathrm{Tr}\, A| = \sum_{n=1}^{\infty} |\langle e_n, Ae_n\rangle|.$$

Before giving the proof we make two remarks. 1) For any operator A the set of normalized vectors h such that $\langle e_n, Ae_n\rangle = 0$ is a closed subset of the unit sphere of $\mathcal{H}$. 2) Given a complete, separable metric space $\mathcal{P}$ and a dense sequence $\mathcal{X}$ in $\mathcal{P}$, one may explicitely construct, for any closed subspace $\mathcal{P}'$ in $\mathcal{P}$, a dense sequence $\mathcal{X}'$ in $\mathcal{P}'$.

Proof. First assume that $\mathrm{Tr}\, A = 0$. We make the following observation :
(O) B being a trace class operator with $\mathrm{Tr}\, B = 0$, there is a normalized $f \in \mathcal{H}$ such that $\langle f, Bf\rangle = 0$. Therefore $\mathrm{Tr}\, B_E = 0$, where E is the subspace orthogonal to f.

Indeed, let $\{g_n\}_{n=1}^{\infty}$ be a hilbertian basis. Suppose that for any integer k, 0 is not in the convex hull of $\{\langle g_n, Bg_n\rangle\}_{n=1}^{k}$. This implies that the sequence $\{\langle g_n, Ag_n\rangle\}_{n=1}^{\infty}$ lies in an open half-plane not containing 0. We deduce that $\mathrm{Tr}\, B \neq 0$, a contradiction. So, there is an integer k such that 0 lies in the convex hull of $\{\langle g_n, Bg_n\rangle\}_{n=1}^{k}$. By the Hausdorff-Toeplitz theorem we have a normalized $f \in \mathrm{span}\{g_n | 1 \leq n \leq k\}$ satisfying $\langle f, Bf\rangle = 0$ and (O) is proved.

To avoid transfinite induction and to get an explicit basis, we go on in the following way. Let $\mathcal{X} = \{x_n\}_{n=1}^{\infty}$ be a dense sequence in the unit sphere of $\mathcal{H}$. For any projection F, we set

$$\delta(F) = \sum_{n=1}^{\infty} 2^{-n} \|x_n - Fx_n\|.$$

The map δ satisfies the three following properties.

(i) $\delta(I) = 0$, where I is the identity operator
(ii) If $F_n \longrightarrow F$ in the weak operator topology, then $\delta(F_n) \longrightarrow \delta(F)$
(iii) $F \leq G$ implies $\delta(F) \geq \delta(G)$. Moreover $F \leq G$ and $\delta(F) = \delta(G)$ imply $F = G$.

We construct by induction an orthonormal sequence $\mathcal{E} = \{e_n\}_{n=1}^{\infty}$ in $\mathcal{H}$. Let $E_k = \operatorname{span}\{e_1, \dots e_k\}$. We take $e_{k+1} \in E_k^{\perp}$ such that

$$(\mathrm{C}_k) \qquad \delta(E_{k+1}) < 2^{-(k+1)} + \inf\{\delta(F) \mid F \in \mathcal{P}_k\},$$

where $\mathcal{P}_k$ is the set of projections F which can be written $F = E_k + f \otimes f$ with f such that $\langle f, Af\rangle = 0$. Note that by observation (O) applied to $B = A_{E_k^{\perp}}$, $\mathcal{P}_k$ is not empty. Note also that by the remarks preceding the proof, the choice of e_{k+1} can be explicitely made.

We claim that $\mathcal{E}$ is a basis. Assume it is not; so the projection $E = \lim E_k$ is not the full space. Since $\operatorname{Tr} AE = \lim \operatorname{Tr} AE_k$ we deduce that there exists a normalized f in $E^{\perp}$ with $\langle f, Af\rangle = 0$ and a x_p in $\mathcal{X}$ such that $\|f - x_p\| < 1/2$. Thus $\|x_p - Ex_p\| > \|f - x_p\|$. Therefore we have an $\eta > 0$ such that

$$2^{-p}\|x_p - Ex_p\| = 2^{-p}\|f - x_p\| + \eta. \tag{4}$$

We then take q large enough to have

$$(5) \quad 2^{-q} < \eta \qquad \text{and} \qquad (6) \quad \delta(E_r) - \delta(E) < \eta/2 \text{ for all } r \geq q.$$

From (4) we get

$$2^{-p}\|x_p - f\| + \eta \leq 2^{-p}\|x_p - E_q a_p\|,$$

hence

$$2^{-p}\|x_p - (E_q + f \otimes f)f\| + \eta \leq 2^{-p}\|x_p - E_q x_p\|,$$

therefore,

$$\delta(E_q + f \otimes f) \leq \delta(E_q) - \eta,$$

so, using (6)

$$\delta(E_q + f \otimes f) \leq \delta(E_{q+1}) - \eta/2,$$

thus, by (5),

$$\delta(E_{q+1}) - \delta(E_q + f \otimes f) > \eta/2 > 2^{-(q+1)};$$

But this contradicts (C_k). So $\mathcal{E}$ is necessarily a basis. Proposition 3.5 is proved in case of $\operatorname{Tr} A = 0$. The general case can be easily deduced. Without loss of generality, we may assume that $\operatorname{Tr} A = 1$. Fix a hilbertian basis $\{h_n\}_{n=1}^{\infty}$ and consider the positive, trace class operator $P = \sum_n 2^{-n} h_n \otimes h_n$. Then $A = (A - P) + P$ with $\operatorname{Tr}(A - P) = 0$. So, we have a hilbertian basis $\{e_n\}_{n=1}^{\infty}$ such that $\langle e_n, (A - P)e_n\rangle = 0$ for all n. Since $\ker P$ is reduced to 0, we have $\langle e_n, Ae_n\rangle > 0$ for all n. $\square$

Corollary 3.6. *Let H and K be two self-adjoint trace class operators. There is a hilbertian basis $\{e_n\}_{n=1}^{\infty}$ and a sequence $\{a_n\}_{n=1}^{\infty}$ in R_+^* such that $\langle e_n, Ae_n\rangle = a_n \operatorname{Tr} K$ and $\langle e_n, He_n\rangle = a_n \operatorname{Tr} H$ for all n.*

4. Addendum

We give a proof of the result of Chui-Smith-Smith-Ward.

Proof. We first note that if the essential numerical range of A is contained in a line then Proposition 1.1 is a rather elementary fact. Indeed, we then know that $A = \lambda I + \mu S + L$ for some scalars λ, μ, some self-adjoint operator S and some compact operator L (I stands for the identity). An obvious property of the numerical range,

$$W(\lambda I + \mu X) = \lambda + \mu W(X)$$

for all scalars λ, μ and all operators X, reduces the proof to the case of a self-adjoint operator A. Elementary spectral theory then shows that there is a compact self-adjoint operator K commuting with A such that $W(A+K)$ is either an open segment or a single point.

Now, we consider the more general and interesting case when the essential numerical range of A has a nonempty interior. We may assume, and we do, that 0 is in the interior of $W_e(A)$. We have the following property :

($\star$) If $\{F_j\}$ is a decreasing sequence of projections of finite codimension, then $\overline{W}(F_j A F_j)$ converge to $W_e(A)$ in the Hausdorff metric.

Recall that the Hausdorff distance between compact subsets $X, Y \subset \mathbf{C}$ is

$$\mathrm{dist}(X,Y) = \max_{x\in X} \min_{y\in Y} |x-y| + \max_{y\in Y} \min_{x\in X} |x-y|.$$

For any $\varepsilon > 0$, we know that we may find a compact operator L such that the Hausdorff distance of $\overline{W}(A+L)$ from $W_e(A)$ is smaller than ε [1, Proposition 2]. This allows us to assume that

$$\frac{1}{2}W(A) \subset\subset W_e(A), \tag{7}$$

where we use the notation $X \subset\subset Y$, for subsets X, Y in $\mathbf{C}$, to mean that there is a small ball centered at the origine, say B, such that $X + B \subset Y$. The construction of K will then result of the following claim:

(C) There is a decreasing sequence $\{E_n\}$ of projections of finite rank such that the operators

$$A_n = \sum_{k=1}^{n} \frac{1}{2^k} E_k A E_k$$

satisfy $W(A_n) \subset\subset W_e(A)$ for all n.

Indeed, assuming (C) proved, we notice that the operator

$$A_\infty = \sum_{k=1}^{\infty} \frac{1}{2^k} E_k A E_k$$

satisfies $W(A_\infty) \subset W_e(A)$. Furthermore, $A_\infty = A + K$ where $K =$

$$-\sum_{k=1}^{\infty} \frac{1}{2^k} \Big\{ E_k A(E_k - E_{k+1}) + (E_k - E_{k+1})AE_k - (E_k - E_{k+1})A(E_k - E_{k+1}) \Big\}$$

which is compact as a norm limit of finite rank operators. Consequently $\overline{W}(A + K) = W_e(A)$ as we wished.

We prove (C) by induction. Fix $N \geq 1$ and suppose that we have found projections $E_1 = I \geq \cdots \geq E_N$ as in (C) and such that

$$W(A_N) \subset\subset (1 - \frac{1}{2^{N+1}})W_e(A). \tag{8}$$

By (7), this is true for $N = 1$. We observe that

$$W_e(A_N) = (1 - \frac{1}{2^N})W_e(A). \tag{9}$$

Let $\{e_j\}$ be an orthonormal basis of E_N and consider

$$E_N^j = \overline{\mathrm{span}}\{e_k,\ k \geq j\}.$$

We claim that there exists an integer p such that setting $E_{N+1} = E_N^p$ we have

$$W(A_{N+1}) = W(A_N + \frac{1}{2^{N+1}} E_{N+1} A E_{N+1}) \subset\subset (1 - \frac{1}{2^{N+2}})W_e(A).$$

Let us denote by $z \frown Z$ the distance of $z \in \mathbf{C}$ from $Z \subset \mathbf{C}$. If the previous claim was not true, there would exist a sequence $\{x_j\}$ of unit vectors such that

$$\langle x_j, (A_N + \frac{1}{2^{N+1}} E_N^j A E_N^j) x_j \rangle \frown W_e \setminus (1 - \frac{1}{2^{N+2}})W_e(A) \longrightarrow 0 \tag{10}$$

as $j \to \infty$. By (6) and strong-$\lim_{j\to\infty} E_N^j = 0$, we should have $x_j \to 0$ weakly. This zero weak-limit and $(*)$ would then imply that, using (7),

$$\langle x_j, (A_N + \frac{1}{2^{N+1}} E_N^j A E_N^j) x_j \rangle \frown (1 - \frac{1}{2^N} + \frac{1}{2^{N+1}})W_e(A) \longrightarrow 0 \tag{11}$$

as $j \to \infty$. Note that $1 - 2^{-N} + 2^{-N-1} = 1 - 2^{-N-1} < 1 - 2^{-N-2}$ so that (11) contradicts (10). Therefore (C) holds and the proof is complete. $\square$

References

[1] J.H. Anderson and J.G. Stampfli, *Commutators and compression*, Israel J. Math **10** (1971), 433-441

[2] I.D. Berg, *A extension of the Weyl-von Neumann theorem to normal operators* Trans. Amer. Math. Soc. **160** (1971), 365-371

[3] I.D. Berg and B. Sims, *Denseness of operators which attain their numerical ranges*, J. Austral. Math. Soc. **36**, serie A (1984), 130-133

[4] C.K. Chui, P.W. Smith, R.R. Smith, J.D. Ward, *L-ideals and numerical range preservation*, Illinois J. Math. **21**, 2 (1977), 365-373.

[5] P. Fan, *On the diagonal of an operator*, Trans. Amer.Math. Soc. **283** (1984), 239-251

[6] R.A. Horn and C.R. Johnson, *Topics in matrix analysis*, Cambridge University Press, 1990

[7] E.M. Klein, *The numerical range of a Toeplitz operator*, Proc. Amer. Math. Soc. **35** (1972), 101-103

[8] J.S. Lancaster, *The boundary of the numerical range*, Proc. Amer. Math. Soc. **49**, 2 (1975), 393-398

[9] Q.F. Stout, *Shur products of operators and the essential numerical range*, Trans. Amer. Math. Soc **264**, 1 (1981), 39-47

[10] Q.F. Stout, *The numerical range of a weighted shift*, Proc. Amer. Math. Soc. **88**, 3 (1983), 495-502

[11] J.K. Thukral, *The numerical range of a Toeplitz operator with harmonic symbol*, J. Operator theory **34** (1995), 213-216