Interdisciplinary Information Sciences Vol. 16, No. 2 (2010) 139–248

ISSN 1340-9050 print/1347-6157 online
DOI 10.4036/iis.2010.139

Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization

Fumio HIAI*

Graduate School of Information Sciences, Tohoku University, Aoba-ku, Sendai 980-8579, Japan

Received February 26, 2010; final version accepted July 20, 2010

These lecture notes are concerned with several topics in matrix analysis covering the Löwner and Kraus theory on matrix/operator monotone and convex functions, the Kubo–Ando theory on operator means, spectral variation and majorization, and means for matrices. Matrix norm inequalities related to majorization and means for matrices are also discussed.

KEYWORDS: matrix, operator, inner product, norm, positive semidefinite, operator monotone, operator convex, divided difference, operator mean, eigenvalue, singular value, majorization, means

Preface

These lecture notes are largely based on my course at Graduate School of Information Sciences of Tohoku University during April–July of 2009. The aim of my lectures was to explain several important topics on matrix analysis from the point of view of functional analysis. These notes are also suitable for an introduction to functional analysis though the arguments are mostly restricted to the finite-dimensional situation.

The main topics covered in these notes are matrix/operator monotone and convex functions (the so-called Löwner and Kraus theory), operator means (the so-called Kubo–Ando theory), majorization for eigen/singular values of matrices and its applications to matrix norm inequalities, and means of matrices and related norm inequalities. These have been chosen from my knowledge and interest while there are many other important topics on the subject.

I have tried to make expositions as transparent as possible and also as self-contained as possible. To do so, some technical stuffs rather apart from matrix analysis are compiled in Appendices. The proof of the theorem of Kraus is also deferred into Appendices since it seems too much to include in the main body. A number of exercises are put in these lecture notes, which are supplements of my expositions as proofs omitted, examples, and further remarks. Concerning References I should mention that their list and citations are not so complete.

At the moment I am collaborating with D. Petz in writing a more comprehensive textbook on matrix analysis, hoping that some parts of these notes will be incorporated into the forthcoming book.

I express my gratitude to Professors T. Ando and H. Kosaki. Ando sent me his English translation of the German paper by Kraus, without which I could not understand the characterization of matrix convex functions due to Kraus. Kosaki gave me comments on Chapter 5, which were helpful to update the content of the chapter. I am thankful to Professor N. Obata, Editor-in-Chief of Interdisciplinary Information Sciences, who suggested me to submit these lecture notes as GSIS selected lectures, a newly launched section of the journal. Finally, this work was partially supported by Grant-in-Aid for Scientific Research (C)21540208.

1. Basics on Matrices

1.1 Basic definitions

For each $n \in \mathbb{N}$, $\mathbb{M}_n = \mathbb{M}_n(\mathbb{C})$ denotes the space of all $n \times n$ complex matrices, which is an n^2-dimensional complex vector space with the linear operations

$$\lambda A := [\lambda a_{ij}], \qquad A + B := [a_{ij} + b_{ij}]$$

for $A = [a_{ij}]_{i,j=1}^n, B = [b_{ij}]_{i,j=1}^n \in \mathbb{M}_n$ and for $\lambda \in \mathbb{C}$. For $i, j = 1, \dots, n$ let E_{ij} be the $n \times n$ matrix of (i, j)-entry equal to one and all other entries equal to zero. Then E_{ij}, $1 \le i, j \le n$, are called *matrix units* and form a basis of $\mathbb{M}_n$ as

$$A = [a_{ij}]_{i,j=1}^n = \sum_{i,j=1}^n a_{ij} E_{ij}.$$

2010 Mathematics Subject Classification: 15A42, 15A45, 15A60, 47A64.

* Corresponding author. E-mail: hiai@math.is.tohoku.ac.jp

The *product* AB of $A = [a_{ij}]_{i,j=1}^n$ and $B = [b_{ij}]_{i,j=1}^n$ is defined by

$$AB = [c_{ij}]_{i,j=1}^n \quad \text{with} \quad c_{ij} := \sum_{k=1}^n a_{ik} b_{kj}.$$

Moreover, the *adjoint matrix* A^* of A is defined as the conjugate transpose of A, i.e.,

$$A^* := \overline{A}^t = [\overline{a_{ji}}]_{i,j=1}^n \quad \text{where} \quad \overline{A} := [\overline{a_{ij}}]_{i,j=1}^n, \ A^t := [a_{ji}]_{i,j=1}^n.$$

Then $\mathbb{M}_n$ becomes a $*$-algebra:

$$(AB)C = A(BC), \quad (A+B)C = AC + BC, \quad A(B+C) = AB + AC,$$
$$(A+B)^* = A^* + B^*, \quad (\lambda A)^* = \overline{\lambda} A^*, \quad (A^*)^* = A, \quad (AB)^* = B^* A^*.$$

The identity of $\mathbb{M}_n$ is the $n \times n$ identity matrix I $(= I_n)$ that is the diagonal matrix of all diagonals equal to one. The most significant feature of matrices is noncommutativity: $AB \neq BA$. For example,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The vector space $\mathbb{C}^n$ of all n-dimensional vectors of complex numbers is a complex Hilbert space with the *Hermitian inner product*

$$\langle x, y \rangle := \sum_{i=1}^n \overline{\xi_i} \eta_i \quad \text{for} \quad x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \ y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}.$$

A matrix $A \in \mathbb{M}_n$ acts on $\mathbb{C}^n$ as a *linear operator* defined by

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = A \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \qquad \eta_i := \sum_{j=1}^n a_{ij} \xi_j, \quad 1 \le i \le n.$$

In fact, $A(x+y) = Ax + Ay$ and $A(\lambda x) = \lambda Ax$ for all $x, y \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$. The product AB corresponds to the composition of linear operators: $(AB)x = A(Bx)$, $x \in \mathbb{C}^n$. The adjoint A^* is determined via inner product: $\langle Ax, y \rangle = \langle x, A^* y \rangle$, $x, y \in \mathbb{C}^n$. The identity matrix corresponds to the identity operator: $Ix = x$, $x \in \mathbb{C}^n$.

1.2 Finite-dimensional Hilbert space

Let $\mathcal{H}$ be an abstract n-dimensional complex Hilbert space with inner product $\langle \cdot, \cdot \rangle$, i.e., for every $x, y, z \in \mathcal{H}$ and $\lambda \in \mathbb{C}$,

$$\begin{cases} \langle x, x \rangle \ge 0, \quad \langle x, x \rangle = 0 \iff x = 0, \\ \langle \lambda x, y \rangle = \overline{\lambda} \langle x, y \rangle, \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \langle x, y \rangle = \overline{\langle y, x \rangle}, \end{cases}$$

so that $\langle x, y \rangle$ is linear in y, i.e.,

$$\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle$$

while conjugate-linear in x, i.e.,

$$\langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle = \overline{\lambda_1} \langle x_1, y \rangle + \overline{\lambda_2} \langle x_2, y \rangle.$$

(This is the physics convention; in mathematics, $\langle x, y \rangle$ is linear in x and conjugate-linear in y.) The *norm* of $x \in \mathcal{H}$ is defined by

$$\|x\| := \langle x, x \rangle^{1/2}.$$

Then we have the *Schwarz inequality*

$$|\langle x, y \rangle| \le \|x\| \, \|y\|, \qquad x, y \in \mathcal{H} \tag{1.2.1}$$

and $\| \cdot \|$ indeed satisfies the properties of norm, i.e., for every $x, y \in \mathcal{H}$ and $\lambda \in \mathbb{C}$,

$$\begin{cases} \|x\| \ge 0, \quad \|x\| = 0 \iff x = 0, \\ \|\lambda x\| = |\lambda| \, \|x\|, \quad \|x + y\| \le \|x\| + \|y\|. \end{cases} \tag{1.2.2}$$

Exercise 1.2.1. Show the Schwarz inequality (1.2.1) and that the equality $|\langle x, y \rangle| = \|x\| \, \|y\|$ occurs if and only if x, y are linearly dependent. Also, show the properties (1.2.2) of norm.

Two vectors $x, y \in \mathcal{H}$ is said to be *orthogonal*, denoted by $x \perp y$, if $\langle x, y\rangle = 0$. A basis $\{e_1, \dots, e_n\}$ of $\mathcal{H}$ is said to be *orthonormal* if $\langle e_i, e_j\rangle = \delta_{ij}$ for $i, j = 1, \dots, n$. Such a basis is obtained by applying the *Gram–Schmidt procedure* to any linear basis. That is, for a linear basis $\{v_1, \dots, v_n\}$ of $\mathcal{H}$, we define

$$\begin{aligned}
e_1 &:= \frac{1}{\|v_1\|} v_1, \\
e_2 &:= \frac{1}{\|w_2\|} w_2 \quad \text{with} \quad w_2 := v_2 - \langle e_1, v_2\rangle e_1, \\
e_3 &:= \frac{1}{\|w_3\|} w_3 \quad \text{with} \quad w_3 := v_3 - \langle e_1, v_3\rangle e_1 - \langle e_2, v_3\rangle e_2, \\
&\vdots \\
e_n &:= \frac{1}{\|w_n\|} w_n \quad \text{with} \quad w_n := v_n - \langle e_1, v_n\rangle e_1 - \cdots - \langle e_{n-1}, v_n\rangle e_{n-1}.
\end{aligned}$$

Exercise 1.2.2. Show that $\{e_1, \dots, e_n\}$ constructed via the Gram–Schmidt procedure above is an orthonormal basis of $\mathcal{H}$.

Now, fix an orthonormal basis $\{e_1, \dots, e_n\}$ of $\mathcal{H}$. Then each $x \in \mathcal{H}$ is written as a unique linear combination of $\{e_1, \dots, e_n\}$ as

$$x = \sum_{i=1}^n \langle e_i, x\rangle e_i.$$

This is called the *Fourier expansion* of x, and $\langle e_i, x\rangle$, $1 \le i \le n$, are called the coordinates of x with respect to $\{e_1, \dots, e_n\}$. Since

$$\langle x, y\rangle = \sum_{i=1}^n \overline{\langle e_i, x\rangle}\langle e_i, y\rangle, \qquad x, y \in \mathcal{H}, \tag{1.2.3}$$

it follows that $\mathcal{H}$ is isomorphic to $\mathbb{C}^n$ as Hilbert spaces via $x \in \mathcal{H} \mapsto (\langle e_i, x\rangle)_{i=1}^n \in \mathbb{C}^n$.

Let $f : \mathcal{H} \to \mathbb{C}$ be a linear functional on $\mathcal{H}$. Set $\lambda_i := f(e_i)$ for $1 \le i \le n$. Then for every $x \in \mathcal{H}$,

$$f(x) = f\left(\sum_{i=1}^n \langle e_i, x\rangle e_i\right) = \sum_{i=1}^n \langle e_i, x\rangle f(e_i) = \sum_{i=1}^n \lambda_i \langle e_i, x\rangle = \left\langle \sum_{i=1}^n \overline{\lambda_i} e_i, x\right\rangle.$$

Hence, if we set $x_f := \sum_{i=1}^n \overline{\lambda_i} e_i$, then

$$f(x) = \langle x_f, x\rangle, \qquad x \in \mathcal{H},$$

and such an $x_f \in \mathcal{H}$ is uniquely determined. (This is the *Riesz representation theorem* in the finite-dimensional case.)

When $\mathcal{M}$ is a subspace of $\mathcal{H}$, the *orthogonal complement* $\mathcal{M}^\perp$ of $\mathcal{M}$ is the subspace of $\mathcal{H}$ defined by

$$\mathcal{M}^\perp := \{x \in \mathcal{H} : \langle x, y\rangle = 0 \text{ for all } y \in \mathcal{M}\}.$$

Choose an orthonormal basis $\{e_1, \dots, e_m\}$ of $\mathcal{M}$ where $m = \dim \mathcal{M}$. One can enlarge $\{e_1, \dots, e_k\}$ to an orthonormal basis $\{e_1, \dots, e_m, e_{m+1}, \dots, e_n\}$ of $\mathcal{H}$. Then $\{e_{m+1}, \dots, e_n\}$ is an orthonormal basis of $\mathcal{M}^\perp$ so that $\mathcal{H}$ has the *orthogonal decomposition*

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp,$$

that is, for every $x \in \mathcal{H}$ there exist unique $x_0 \in \mathcal{M}$ and $x_1 \in \mathcal{M}^\perp$ such that $x = x_0 + x_1$.

Let $B(\mathcal{H})$ denote the set of all linear operators on $\mathcal{H}$, which is a vector space with usual linear operations. The *product* AB of $A, B \in B(\mathcal{H})$ is defined as the composition. The *adjoint* A^* of $A \in B(\mathcal{H})$ is defined as

$$\langle x, Ay\rangle = \langle A^*x, y\rangle, \qquad x, y \in \mathcal{H}. \tag{1.2.4}$$

In fact, for each $x \in \mathcal{H}$, we have a linear functional $y \in \mathcal{H} \mapsto \langle x, Ay\rangle \in \mathbb{C}$ so that by the Riesz representation theorem, there exists a unique $A^*x \in \mathcal{H}$ for which (1.2.4) holds. Then it is easy to see that $x \mapsto A^*x$ is a linear operator on $\mathcal{H}$, i.e., $A^* \in B(\mathcal{H})$.

Exercise 1.2.3. Show that $B(\mathcal{H})$ becomes a $*$-algebra with the operations introduced above.

For each $A \in B(\mathcal{H})$ we associate an $n \times n$ matrix $[a_{ij}]$ given by

$$a_{ij} := \langle e_i, Ae_j\rangle, \qquad 1 \le i, j \le n, \tag{1.2.5}$$

that is,

$$Ae_j = \sum_{i=1}^{n} a_{ij} e_i, \qquad 1 \le i \le n.$$

When $x = \sum_{i=1}^{n} \xi_i e_i$ and $y = Ax = \sum_{i=1}^{n} \eta_i e_i$, we have

$$\eta_i = \langle e_i, Ax \rangle = \left\langle e_i, \sum_{j=1}^{n} \xi_j A e_j \right\rangle = \sum_{j=1}^{n} \xi_j \langle e_i, Ae_j \rangle = \sum_{j=1}^{n} a_{ij} \xi_j, \qquad 1 \le i \le n,$$

so that the equation $y = Ax$ is rewritten in terms of the coordinates of x and y as

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Proposition 1.2.4. *Let Φ be the map sending $A \in B(\mathcal{H})$ to $[a_{ij}] \in \mathbb{M}_n$ defined by* (1.2.5). *Then Φ is a $*$-isomorphism between $B(\mathcal{H})$ and $\mathbb{M}_n$, that is, Φ is a linear map from $B(\mathcal{H})$ onto $\mathbb{M}_n$ such that $\Phi(AB) = \Phi(A)\Phi(B)$ and $\Phi(A^*) = \Phi(A)^*$ for all $A, B \in B(\mathcal{H})$.*

Proof. Let $\Phi(A) = [a_{ij}]$ and $\Phi(B) = [b_{ij}]$. Since

$$\begin{aligned} (AB)e_j = A(Be_j) &= A\left(\sum_{k=1}^{n} b_{kj} e_k\right) = \sum_{k=1}^{n} b_{kj} A e_k \\ &= \sum_{k=1}^{n} b_{kj} \left(\sum_{i=1}^{n} a_{ik} e_i\right) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) e_i, \end{aligned}$$

we have

$$\Phi(AB) = \left[\sum_{k=1}^{n} a_{ik} b_{kj}\right]_{i,j=1}^{n} = [a_{ij}]\,[b_{ij}] = \Phi(A)\Phi(B).$$

Since

$$\langle e_i, A^* e_j \rangle = \overline{\langle A^* e_j, e_i \rangle} = \overline{\langle e_j, A e_i \rangle} = \overline{a_{ji}},$$

we have

$$\Phi(A^*) = [\overline{a_{ji}}]_{i,j=1}^{n} = [a_{ij}]^* = \Phi(A)^*.$$

The linearity of Φ is easy to check, so we omit the details. Finally, for every $[c_{ij}] \in \mathbb{M}_n$, define

$$Cx := \sum_{i=1}^{n} \left(\sum_{j=1}^{n} c_{ij} \xi_j\right) e_i \quad \text{for} \quad x = \sum_{i=1}^{n} \xi_i e_i \in \mathcal{H}.$$

Then it is immediate to see that $C \in B(\mathcal{H})$ and $\Phi(C) = [c_{ij}]$ so that Φ is surjective. □

Thus, $B(\mathcal{H})$ can be identified with $\mathbb{M}_n$ as $*$-algebras when $\mathcal{H}$ is n-dimensional. In particular, $B(\mathbb{C}^n) = \mathbb{M}_n$.

More generally, let $B(\mathcal{H}, \mathcal{K})$ denote the set of all linear operators between two finite-dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ of possibly different dimensions. For each $A \in B(\mathcal{H}, \mathcal{K})$ one can define the adjoint $A^* \in B(\mathcal{K}, \mathcal{H})$ similarly to (1.2.4):

$$\langle u, Ax \rangle_{\mathcal{K}} = \langle A^* u, x \rangle_{\mathcal{H}}, \qquad x \in \mathcal{H},\ u \in \mathcal{K}.$$

Fix orthonormal bases $\{e_1, \dots, e_n\}$ of $\mathcal{H}$ and $\{f_1, \dots, f_m\}$ of $\mathcal{K}$, and associate with A an $m \times n$ matrix $[a_{ij}]$ by

$$a_{ij} := \langle f_i, Ae_j \rangle_{\mathcal{K}}, \qquad 1 \le i \le m,\ 1 \le j \le n,$$

that is, $Ae_j = \sum_{i=1}^{m} a_{ij} f_i$, $1 \le i \le n$. Then $Ax = \sum_{i=1}^{m} \eta_i f_i$ for $x = \sum_{j=1}^{n} \xi_j e_j$ is rewritten in the matrix form:

$$\begin{bmatrix} \eta_1 \\ \vdots \\ \eta_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}.$$

Let $\mathbb{M}_{m,n}$ denote the set of all $m \times n$ complex matrices. Then $B(\mathcal{H}, \mathcal{K})$ can be identified with $\mathbb{M}_{m,n}$ as complex vector spaces when $\dim \mathcal{H} = n$ and $\dim \mathcal{K} = m$. In particular, $B(\mathbb{C}^n, \mathbb{C}^m) = \mathbb{M}_{m,n}$.

1.3 Basic notions of operators and matrices

Let I $(= I_{\mathcal{H}})$ denote the identity operator on a finite-dimensional Hilbert space $\mathcal{H}$. For $A \in B(\mathcal{H})$, the *kernel* and the *range* of A are

$$\ker A := \{x \in \mathcal{H} : Ax = 0\}, \qquad \operatorname{ran} A := \{Ax : x \in \mathcal{H}\},$$

respectively, which are subspaces of $\mathcal{H}$. The dimension formula familiar in linear algebra is

$$\dim \mathcal{H} = \dim(\ker A) + \dim(\operatorname{ran} A). \tag{1.3.1}$$

We say that $A \in \mathbb{M}_n$ is *invertible* if there is an $B \in B(\mathcal{H})$ such that $AB = BA = I$. In this case, such a $B \in B(\mathcal{H})$ is unique and we write $B = A^{-1}$, called the *inverse* of A. It is seen from (1.3.1) that $A \in B(\mathcal{H})$ is invertible if and only if $\ker A = \{0\}$ (i.e., A is injective), that is also equivalent to $\operatorname{ran} A = \mathcal{H}$ (i.e., A is surjective).

Lemma 1.3.1. (1) *For every $A \in B(\mathcal{H})$ and every $x, y \in \mathcal{H}$,*

$$\begin{aligned}\langle y, Ax\rangle = \frac{1}{4}\{&\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle \\ &+ i\langle x+iy, A(x+iy)\rangle - i\langle x-iy, A(x-iy)\rangle\}.\end{aligned} \tag{1.3.2}$$

(*This identity is called the polarization identity.*)

(2) *If $A \in B(\mathcal{H})$ and $\langle x, Ax\rangle = 0$ for all $x \in \mathcal{H}$, then $A = 0$.*

Proof. (1) is obtained by a direct computation, and (2) immediately follows from (1). □

An operator $A \in B(\mathcal{H})$ is said to be *normal* if $A^*A = AA^*$, and *self-adjoint* or *Hermitian* if $A = A^*$. If $\langle x, Ax\rangle \ge 0$ for all $x \in \mathcal{H}$, then A is said to be *positive semidefinite* or simply *positive*, and we write $A \ge 0$. Moreover, if $\langle x, Ax\rangle > 0$ for all $x \in \mathcal{H}$ with $x \ne 0$, then A is *positive definite* or *strictly positive*, and we write $A > 0$. The sets of all self-adjoint operators and of all positive semidefinite operators in $B(\mathcal{H})$ are denoted by $B(\mathcal{H})^{sa}$ and $B(\mathcal{H})^+$, respectively. If $A, B \in B(\mathcal{H})^{sa}$ and $B - A \ge 0$, i.e., $\langle x, Ax\rangle \le \langle x, Bx\rangle$ for all $x \in \mathcal{H}$, then we write $A \le B$.

Proposition 1.3.2.

(1) *$A \in B(\mathcal{H})$ is normal if and only if $\|Ax\| = \|A^*x\|$ for all $x \in \mathcal{H}$.*
(2) *$A \in B(\mathcal{H})$ is self-adjoint if and only if $\langle x, Ax\rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$.*
(3) *$A \le B$ is a partial order relation on $B(\mathcal{H})^{sa}$.*
(4) *If $A, B \in B(\mathcal{H})^{sa}$ and $A \le B$ then $C^*AC \le C^*BC$ for all $C \in B(\mathcal{H})$.*

Proof. (1) follows from Lemma 1.3.1 (2) since

$$\langle x, (A^*A - AA^*)x\rangle = \langle Ax, Ax\rangle - \langle A^*x, A^*x\rangle = \|Ax\|^2 - \|A^*x\|^2.$$

(2) is also seen from the same lemma since

$$\langle x, (A - A^*)x\rangle = \langle x, Ax\rangle - \langle Ax, x\rangle = \langle x, Ax\rangle - \overline{\langle x, Ax\rangle}.$$

(3) For $A, B \in B(\mathcal{H})^{sa}$, if $A \le B$ and $A \ge B$ then $\langle x, (A - B)x\rangle = 0$ and so $A = B$. Other properties for a partial order are immediate.

(4) Under the assumption of (4), $\langle x, C^*ACx\rangle = \langle Cx, ACx\rangle \le \langle Cx, BCx\rangle = \langle x, C^*BCx\rangle$ for all $x \in \mathcal{H}$. □

An operator $U \in B(\mathcal{H})$ is called a *unitary* if $U^*U = UU^* = I$, i.e., U^* is the inverse of U.

Proposition 1.3.3. *For $U \in B(\mathcal{H})$ the following are equivalent:*

(i) *U is a unitary;*
(ii) *$\langle Ux, Uy\rangle = \langle x, y\rangle$ for all $x, y \in \mathcal{H}$;*
(iii) *$\|Ux\| = \|x\|$ for all $x \in \mathcal{H}$.*

Proof. If U is a unitary then $\langle Ux, Uy\rangle = \langle x, U^*Uy\rangle = \langle x, y\rangle$. Hence (i) ⇒ (ii). (ii) ⇒ (iii) is obvious. If (iii) is satisfied, then $\langle x, (U^*U - I)x\rangle = \|Ux\|^2 - \|x\|^2 = 0$ for all $x \in \mathcal{H}$. By Lemma 1.3.1 (2), $U^*U = I$. In particular, U is injective and so invertible. Hence $U^* = U^{-1}$ and (iii) ⇒ (i). □

Clearly, the set of all normal operators on $\mathcal{H}$ includes the set of all unitaries on $\mathcal{H}$ and also $B(\mathcal{H})^{sa}$ $(\supset B(\mathcal{H})^+)$.

Another important notion of operators is that of orthogonal projections. Let $\mathcal{M}$ be a subspace of $\mathcal{H}$. For each $x \in \mathcal{H}$ take the orthogonal decomposition $x = x_0 + x_1$ with $x_0 \in \mathcal{M}$ and $x_1 \in \mathcal{M}^\perp$, and define $P_{\mathcal{M}}x := x_0$. Then it is immediate to see that $P_{\mathcal{M}}$ is a linear operator on $\mathcal{H}$ with $\operatorname{ran} P_{\mathcal{M}} = \mathcal{M}$. The operator $P_{\mathcal{M}}$ is called the *orthogonal projection* from $\mathcal{H}$ onto $\mathcal{M}$.

Proposition 1.3.4. *$P \in B(\mathcal{H})$ is the orthogonal projection onto a subspace of $\mathcal{H}$ if and only if $P^* = P = P^2$.*

Proof. For a subspace $\mathcal{M}$ of $\mathcal{H}$, $P_{\mathcal{M}}^2 = P_{\mathcal{M}}$ is clear. For every $x, y \in \mathcal{M}$ take the orthogonal decompositions $x = x_0 + x_1$ and $y = y_0 + y_1$ with $x_0, y_0 \in \mathcal{M}$ and $x_1, y_1 \in \mathcal{M}^\perp$. Then

$$\langle x, P_{\mathcal{M}} y\rangle = \langle x_0 + x_1, y_0\rangle = \langle x_0, y_0\rangle = \langle x_0, y_0 + y_1\rangle = \langle P_{\mathcal{M}} x, y\rangle,$$

implying that $P_{\mathcal{M}} = P_{\mathcal{M}}^*$. Conversely, assume that $P^* = P = P^2$ and set $\mathcal{M} := \operatorname{ran} P$. For every x write $x = x_0 + x_1$ with $x_0 \in \mathcal{M}$ and $x_1 \in \mathcal{M}^\perp$. Since $x_0 = Pz$ for some $z \in \mathcal{H}$, $Px_0 = P^2 z = Pz = x_0$. Moreover, $\langle y, Px_1\rangle = \langle Py, x_1\rangle = 0$ for all $y \in \mathcal{H}$ so that $Px_1 = 0$. Hence $Px = Px_0 + Px_1 = x_0 = P_{\mathcal{M}} x$, implying that $P = P_{\mathcal{M}}$. □

Next, let $\mathcal{H}$ and $\mathcal{K}$ be two finite-dimensional Hilbert spaces. For each $A \in B(\mathcal{H}, \mathcal{K})$ we have $A^* \in B(\mathcal{K}, \mathcal{H})$ as defined in the preceding section, and so $A^*A \in B(\mathcal{H})$ and $AA^* \in B(\mathcal{K})$. If $\|Ax\| \le \|x\|$ for all $x \in \mathcal{H}$, then A is called a *contraction*. Moreover, A is called an *isometry* if $\|Ax\| = \|x\|$ for all $x \in \mathcal{H}$.

Exercise 1.3.5. Let $A \in B(\mathcal{H}, \mathcal{K})$. Prove the following assertions:
(1) The following conditions are equivalent: (i) A is a contraction, (ii) $A^*A \le I_{\mathcal{H}}$, and (iii) $AA^* \le I_{\mathcal{K}}$.
(2) A is an isometry if and only if $A^*A = I_{\mathcal{H}}$. In this case, AA^* is the orthogonal projection from $\mathcal{K}$ onto $\operatorname{ran} A$.

Finally, note that all the materials in this section can be, in particular, applied to matrices in $\mathbb{M}_n = B(\mathbb{C}^n)$ and in $\mathbb{M}_{m,n} = B(\mathbb{C}^n, \mathbb{C}^m)$. We write $\mathbb{M}_n^{sa}$ and $\mathbb{M}_n^+$ for the sets of all $n \times n$ Hermitian matrices and of all $n \times n$ positive semidefinite matrices. An important fact in linear algebra is that $A \in \mathbb{M}_n$ is invertible if and only if $\det A \ne 0$, i.e., A has the non-zero determinant. All $n \times n$ unitary matrices form a group, the so-called unitary group of order n.

Exercise 1.3.6. Let $U \in \mathbb{M}_n$ and $u_1, \dots, u_n$ be n column vectors of U, i.e., $U = [u_1\, u_2 \cdots u_n]$. Prove that U is a unitary matrix if and only if $\{u_1, \dots, u_n\}$ is an orthonormal basis of $\mathbb{C}^n$.

1.4 Spectral decomposition and polar decomposition

Let $\mathcal{H}$ be an n-dimensional Hilbert space. For $A \in B(\mathcal{H})$ and $\lambda \in \mathbb{C}$, we say that λ is an *eigenvalue* of A if there is a non-zero vector $v \in \mathcal{H}$ such that $Av = \lambda v$, i.e., $v \in \ker(A - \lambda I)$. Such a vector v is called an *eigenvector* of A for the eigenvalue λ. Recall that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if $A - \lambda I$ is not invertible, which is equivalent to $\det(\lambda I - A) = 0$. Here, note that $\det B$ of $B \in B(\mathcal{H})$ can be defined by regarding B as an $n \times n$ matrix under some orthonormal basis of $\mathcal{H}$ (see Section 1.2). In fact, the definition of $\det B$ is independent of the choice of an orthonormal basis of $\mathcal{H}$. Since $\det(\lambda I - A)$ is a polynomial of degree n, A has exactly n eigenvalues with counting multiplicities.

Theorem 1.4.1. *Assume that $A \in B(\mathcal{H})$ is normal, i.e., $A^*A = AA^*$. Then there exist $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and $u_1, \dots, u_n \in \mathcal{H}$ such that $\{u_1, \dots, u_n\}$ is an orthonormal basis of $\mathcal{H}$ and $Au_i = \lambda_i u_i$ for all $i = 1, \dots, n$ (i.e., each λ_i is an eigenvalue of A and u_i is the corresponding eigenvector).*

Proof. Let us prove this by induction on $n = \dim \mathcal{H}$. The case $n = 1$ trivially holds. Suppose the assertion for dimension $n - 1$. Assume that $\dim \mathcal{H} = n$ and $A \in B(\mathcal{H})$ is normal. Choose a root λ_1 of $\det(\lambda I - A) = 0$. As explained before the theorem, λ_1 is an eigenvalue of A so that there is an eigenvector u_1 with $Au_1 = \lambda_1 u_1$. One may assume that u_1 is a unit vector, i.e., $\|u_1\| = 1$. Since A is normal, we have

$$\begin{aligned}(A - \lambda_1 I)^*(A - \lambda_1 I) &= (A^* - \overline{\lambda_1} I)(A - \lambda_1 I) \\ &= A^*A - \overline{\lambda_1} A - \lambda_1 A^* + \lambda_1 \overline{\lambda_1} I \\ &= AA^* - \overline{\lambda_1} A - \lambda_1 A^* + \lambda_1 \overline{\lambda_1} I \\ &= (A - \lambda_1 I)(A - \lambda_1 I)^*,\end{aligned}$$

that is, $A - \lambda_1 I$ is also normal. Therefore, by Proposition 1.3.2 (1),

$$\|(A^* - \overline{\lambda_1} I)u_1\| = \|(A - \lambda_1 I)^* u_1\| = \|(A - \lambda_1 I)u_1\| = 0$$

so that $A^* u_1 = \overline{\lambda_1} u_1$. Let $\mathcal{H}_1 := \{u_1\}^\perp$, the orthogonal complement of $\{u_1\}$. If $x \in \mathcal{H}_1$ then

$$\begin{aligned}\langle Ax, u_1\rangle &= \langle x, A^* u_1\rangle = \langle x, \overline{\lambda_1} u_1\rangle = \overline{\lambda_1}\langle x, u_1\rangle = 0, \\ \langle A^* x, u_1\rangle &= \langle x, Au_1\rangle = \langle x, \lambda_1 u_1\rangle = \lambda_1 \langle x, u_1\rangle = 0\end{aligned}$$

so that $Ax, A^*x \in \mathcal{H}_1$. Hence we have $A\mathcal{H}_1 \subset \mathcal{H}_1$ and $A^*\mathcal{H}_1 \subset \mathcal{H}_1$. So one can define $A_1 := A|_{\mathcal{H}_1} \in B(\mathcal{H}_1)$. Then $A_1^* = A^*|_{\mathcal{H}_1}$, which implies that A_1 is also normal. Since $\dim \mathcal{H}_1 = n - 1$, the induction hypothesis can be applied to obtain $\lambda_2, \dots, \lambda_n \in \mathbb{C}$ and $u_2, \dots, u_n \in \mathcal{H}_1$ such that $\{u_2, \dots, u_n\}$ is an orthonormal basis of $\mathcal{H}_1$ and $A_1 u_i = \lambda_i u_i$ for all $i = 2, \dots, n$. Then $\{u_1, u_2, \dots, u_n\}$ is an orthonormal basis of $\mathcal{H}$ and $Au_i = \lambda_i u_i$ for all $i = 1, 2, \dots, n$. Thus the assertion holds for dimension n as well. □

Here, let us introduce a convenient notation. For any $u, v \in \mathcal{H}$ define

$$(|u\rangle\langle v|)x := \langle v, x\rangle u, \qquad x \in \mathcal{H}.$$

Then we have $|u\rangle\langle v| \in B(\mathcal{H})$, whose range is the one-dimensional subspace $\mathbb{C}u$ generated by u if $u, v \neq 0$ (otherwise, $|u\rangle\langle v| = 0$). The physics symbol $|u\rangle\langle v|$ is often written as $u \otimes v$ and called the *Schatten form* in mathematics. In particular, $|u\rangle\langle u|$ for a unit vector u is the orthogonal projection onto $\mathbb{C}u$.

Exercise 1.4.2. Show the following properties:

$$(|u\rangle\langle v|)^* = |v\rangle\langle u|, \quad (|u_1\rangle\langle v_1|)(|u_2\rangle\langle v_2|) = \langle v_1, u_2\rangle |u_1\rangle\langle v_2|,$$
$$A(|u\rangle\langle v|) = |Au\rangle\langle v|, \quad (|u\rangle\langle v|)A = |u\rangle\langle A^*v| \quad \text{for all } A \in B(\mathcal{H}).$$

The conclusion of the above theorem is put together into the form

$$A = \sum_{i=1}^{n} \lambda_i |u_i\rangle\langle u_i|, \tag{1.4.1}$$

that is called the *Schmidt decomposition* of A. Now, let $\alpha_1, \ldots, \alpha_m$ be the distinct eigenvalues of A from $\lambda_1, \ldots, \lambda_n$. The set $\{\alpha_1, \ldots, \alpha_m\}$ is often called the *spectrum* of A and denoted by $\sigma(A)$. Define

$$P_j := \sum_{i\,:\,\lambda_i = \alpha_j} |u_i\rangle\langle u_i|, \qquad 1 \le j \le m.$$

Then P_j is the orthogonal projection onto the eigenspace $\ker(A - \alpha_j I)$ for $1 \le j \le m$ and

$$\sum_{j=1}^{m} P_j = \sum_{i=1}^{n} |u_i\rangle\langle u_i| = I.$$

Thus, $\{P_1, \ldots, P_m\}$ is a *partition of unity* or a *spectral measure*, and we have

$$A = \sum_{j=1}^{m} \alpha_j P_j, \tag{1.4.2}$$

which is called the *spectral decomposition* of A. Note that the normality is indeed necessary for $A \in \mathbb{M}_n$ to have the Schmidt decomposition (1.4.1) or the spectral decomposition (1.4.2). Also, note that the spectral decomposition of A is uniquely determined while the Schmidt decomposition of A is not unique if some eigenvalue of A is not simple.

Corollary 1.4.3. *Let $A \in B(\mathcal{H})$ be normal and $\sigma(A)$ be the spectrum of A. Then the following hold:*
(1) *A is self-adjoint if and only if $\sigma(A) \subset \mathbb{R}$.*
(2) *$A \ge 0$ if and only if $\sigma(A) \subset [0, \infty)$.*
(3) *$A > 0$ if and only if $\sigma(A) \subset (0, \infty)$.*
(4) *A is a unitary if and only if $\sigma(A) \subset \mathbb{T} := \{\zeta \in \mathbb{C} : |\zeta| = 1\}$.*
(5) *A is an orthogonal projection if and only if $\sigma(A) \subset \{0, 1\}$.*

Proof. With the spectral decomposition (1.4.2) we have

$$A^* = \sum_{j=1}^{m} \overline{\alpha_j} P_j, \quad A^*A = \sum_{j=1}^{m} |\alpha_j|^2 P_j, \quad A^2 = \sum_{j=1}^{M} \alpha_j^2 P_j.$$

Then $A^* - A = \sum_{j=1}^{m} (\overline{\alpha_j} - \alpha_j) P_j = 0$ if and only if $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ or $\sigma(A) \subset \mathbb{R}$. The arguments for other properties are similar. □

Let f be a complex-valued function on $D \subset \mathbb{C}$. When $A \in B(\mathcal{H})$ is normal with $\sigma(A) \subset D$, one can define $f(A) \in B(\mathcal{H})$ by

$$f(A) := \sum_{i=1}^{n} f(\lambda_i)\, |u_i\rangle\langle u_i| = \sum_{j=1}^{m} f(\alpha_j) P_j \tag{1.4.3}$$

via the decompositions (1.4.1) and (1.4.2). The correspondence $f \mapsto f(A)$ so defined for functions f whose domain contains the eigenvalues of A is called the *functional calculus* of A. In particular, when f is a function on an interval J in $\mathbb{R}$, $f(A)$ is defined for any $A \in B(\mathcal{H})^{sa}$ whose eigenvalues are contained in J. When $f(\zeta) = \zeta^k$ with $k \in \mathbb{N}$, it is obvious that $f(A) = A^k$ (the k-fold product of A). When $f(\zeta) \equiv 1$, $f(A) = I$. Hence, for any polynomial $p(\zeta)$, $p(A)$ coincides with the usual definition by substitution of A for ζ.

Example 1.4.4.
(1) Consider $f_+(t) := \max\{t, 0\}$ and $f_-(t) := \max\{-t, 0\}$ for $t \in \mathbb{R}$. For each $A \in B(\mathcal{H})^{sa}$ define $A_+ := f_+(A)$ and $A_- := f_-(A)$. Since $f_+(t), f_-(t) \ge 0$, $f_+(t) - f_-(t) = t$ and $f_+(t) f_-(t) = 0$, we have

$$A_+, A_- \ge 0, \quad A = A_+ - A_-, \quad A_+ A_- = 0.$$

These A_+ and A_- are called the *positive part* and the *negative part* of A, respectively, and $A = A_+ + A_-$ is called the *Jordan decomposition* of A.

(2) For any $A \geq 0$ and for any positive $p > 0$, define $A^p := f_p(A) \geq 0$ for the function $f_p(t) := t^p$, $t \geq 0$. This A^p is called the pth (fractional) power of $A \geq 0$. In particular, when $p = 1/2$, $A^{1/2}$ is called the *square-root* of $A \geq 0$. Here, it is worth noting that, unlike the nonnegative number case, $A^p \geq B^p$ does not generally follow from $A \geq B \geq 0$ and $p > 0$. For example, when $A \geq B \geq 0$, $A^{1/2} \geq B^{1/2}$ is always valid while $A^2 \geq B^2$ is not, as will be studied in Chapter 2.

(3) For each $A \in B(\mathcal{H})$, note that $A^*A \geq 0$ since $\langle A^*Ax, x\rangle = \|Ax\|^2 \geq 0$ for all $x \in \mathcal{H}$. So, define $|A| := (A^*A)^{1/2}$ that is called the *absolute value* of A. More generally, for each $A \in B(\mathcal{H}, \mathcal{K})$, since $A^*A \in B(\mathcal{H})$ and $AA^* \in B(\mathcal{K})$ are positive semidefinite, one can define the absolute values $|A| := (A^*A)^{1/2} \in B(\mathcal{H})$ and $|A^*| := (AA^*)^{1/2} \in B(\mathcal{K})$.

Corollary 1.4.5. *For each $A \in B(\mathcal{H})$, $A \geq 0$ if and only if $A = B^*B$ for some $B \in B(\mathcal{H})$, and $A > 0$ if and only if $A = B^*B$ for some invertible $B \in B(\mathcal{H})$.*

Proof. If $A = B^*B$ then $\langle Ax, x\rangle = \|Bx\|^2 \geq 0$ and hence $A \geq 0$. Conversely, if $A \geq 0$ then with the square root $B := A^{1/2}$ in Example 1.4.4 (2), we have $B \geq 0$ and $A = B^2 = B^*B$. The latter assertion is seen since $B \in B(\mathcal{H})$ is invertible if and only if $B^*B > 0$. □

Theorem 1.4.1 is rephrased as the *diagonalization* theorem for normal matrices in the following way.

Theorem 1.4.6. *For every normal matrix $A \in \mathbb{M}_n$, there exist $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and a unitary matrix $U \in \mathbb{M}_n$ such that*

$$A = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*, \tag{1.4.4}$$

where $\mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ stands for the diagonal matrix of diagonals $\lambda_1, \dots, \lambda_n$. Furthermore, $\lambda_1, \dots, \lambda_n$ are uniquely determined (up to permutations) as the eigenvalues of A with counting multiplicities.

Proof. By Theorem 1.4.1 there exist $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ and $u_1, \dots, u_n \in \mathbb{C}^n$ such that $\{u_1, \dots, u_n\}$ is an orthonormal basis of $\mathbb{C}^n$ and $Au_i = \lambda_i u_i$ for $1 \leq i \leq n$. Define $U := [u_1\, u_2\, \cdots\, u_n]$ that is a unitary thanks to Exercise 1.3.6. We then have

$$\begin{aligned} AU &= [Au_1\, Au_2\, \cdots\, Au_n] = [\lambda_1 u_1\, \lambda_2 u_2\, \cdots\, \lambda_n u_n] \\ &= [u_1\, u_2\, \cdots\, u_n]\,\mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = U\mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \end{aligned}$$

so that (1.4.4) is obtained. In this case, since

$$\det(\lambda I - A) = \det(\mathrm{Diag}(\lambda - \lambda_1, \dots, \lambda - \lambda_1)) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n),$$

it follows that $\lambda_1, \dots, \lambda_n$ are the roots of $\det(\lambda I - A) = 0$, or the eigenvalues of A with multiplicities. □

The formula (1.4.4) of diagonalization is rewritten as

$$A = [\lambda_1 u_1\, \cdots\, \lambda_n u_n]\begin{bmatrix} u_1^* \\ \vdots \\ u_n^* \end{bmatrix} = \sum_{i=1}^n \lambda_i u_i u_i^*,$$

that is the Schmidt decomposition (1.4.1). In fact, uv^* is nothing but $|u\rangle\langle v|$ in the case where $\mathcal{H} = \mathbb{C}^n$.

The following decomposition theorem is quite useful in operator/matrix analysis, which is the operator analog of the polar representation $\zeta = |\zeta|e^{i\theta}$ for $\zeta \in \mathbb{C}$.

Theorem 1.4.7. *For every $A \in B(\mathcal{H})$ there exists a unitary $U \in B(\mathcal{H})$ such that*

$$A = U|A|. \tag{1.4.5}$$

Proof. Notice that

$$\|\,|A|x\| = \langle x, |A|^2 x\rangle^{1/2} = \langle x, A^*Ax\rangle^{1/2} = \|Ax\|, \qquad x \in \mathcal{H}. \tag{1.4.6}$$

Set $\mathcal{K} := \{|A|x : x \in \mathcal{H}\}$ and $\mathcal{L} := \{Ax : x \in \mathcal{H}\}$, which are subspaces of $\mathcal{H}$. We define $U_0 : \mathcal{K} \to \mathcal{L}$ by

$$U_0(|A|x) := Ax, \qquad x \in \mathcal{H}. \tag{1.4.7}$$

The well-definedness of U_0 is guaranteed by (1.4.6). In fact, if $|A|x = |A|y$ then $\|Ax - Ay\| = \|A(x - y)\| = \|\,|A|(x - y)\| = 0$ so that $Ax = Ay$. Moreover, it is immediate to see that U_0 is linear. Hence by (1.4.6), U_0 is a linear isometry from $\mathcal{K}$ onto $\mathcal{L}$. This implies also that $\dim \mathcal{K} = \dim \mathcal{L}$ and so $\dim \mathcal{K}^\perp = \dim \mathcal{L}^\perp$. Hence one can choose orthogonal bases $\{u_1, \dots, u_k\}$ of $\mathcal{K}^\perp$ and $\{v_1, \dots, v_k\}$ of $\mathcal{L}^\perp$. Define $U_1 : \mathcal{K}^\perp \to \mathcal{L}^\perp$ by $U_1 u_j = v_j$ for $1 \leq j \leq k$ and extending it by linearity. Then U_1 is a linear isometry from $\mathcal{K}^\perp$ onto $\mathcal{L}^\perp$. Since $\mathcal{K} \oplus \mathcal{K}^\perp = \mathcal{L} \oplus \mathcal{L}^\perp = \mathcal{H}$, one can define a linear isometry U on $\mathcal{H}$ by

$$U(x_0 + x_1) = U_0 x_0 + U_1 x_1 \quad \text{for} \quad x_0 \in \mathcal{K},\ x_1 \in \mathcal{K}^\perp.$$

Then by Proposition 1.3.3, U is a unitary. By the definition (1.4.7) we have $A = U|A|$. □

The expression (1.4.5) is called the *polar decomposition* of A. Note that a unitary U is not unique unless A is invertible. If A is invertible, then so is $|A|$ and U is given by $U = A|A|^{-1}$. In this case, we directly check that $U^*U = |A|^{-1}A^*A|A|^{-1} = |A|^{-1}|A|^2|A|^{-1} = I$.

Exercise 1.4.8. Let $A \in B(\mathcal{H}, \mathcal{K})$ with finite-dimensional Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$. Show that there exist partial isometries $V \in B(\mathcal{H}, \mathcal{K})$ and $W \in B(\mathcal{H}, \mathcal{K})$ such that

$$A = V|A| = |A^*|W,$$

where $V \in B(\mathcal{H}, \mathcal{K})$ is called a partial isometry if $V^*V \in B(\mathcal{H})$ is an orthogonal projection (or equivalently, if $VV^* \in B(\mathcal{K})$ is an orthogonal projection).

It may be conceptually natural to consider normal, self-adjoint, and positive semidefinite operators/matrices as noncommutative counterparts of complex, real, and nonnegative numbers, respectively, and also unitary operators/matrices as complex numbers of modulus one. In this way, the noncommutative analysis is considered as the analysis over operators/matrices in place of real/complex numbers.

1.5 Norms and trace

The *operator norm* of $A \in B(\mathcal{H})$ is defined by

$$\begin{aligned}\|A\| &:= \sup\{\|Ax\| : x \in \mathcal{H},\ \|x\| \le 1\} \\ &= \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in \mathcal{H},\ x \ne 0\right\} \\ &= \sup\{|\langle x, Ay\rangle| : x, y \in \mathcal{H},\ \|x\|, \|y\| \le 1\}. \end{aligned} \tag{1.5.1}$$

This $\|\cdot\|$ on $B(\mathcal{H})$ indeed has the properties of norm, i.e., for every $A, B \in B(\mathcal{H})$ and $\lambda \in \mathbb{C}$,

$$\begin{cases} \|A\| \ge 0, \quad \|A\| = 0 \Longleftrightarrow A = 0, \\ \|\lambda A\| = |\lambda|\,\|A\|, \quad \|A + B\| \le \|A\| + \|B\|. \end{cases} \tag{1.5.2}$$

Exercise 1.5.1. Prove that the three expressions in (1.5.1) are equal. Also show the properties (1.5.2) and $\|AB\| \le \|A\|\,\|B\|$.

Significant properties of the operator norm are:

Proposition 1.5.2. *For every* $A \in B(\mathcal{H})$,

$$\|A^*\| = \|A\|, \qquad \|A^*A\| = \|A\|^2.$$

Proof. Since $|\langle x, A^*y\rangle| = |\langle A^*y, x\rangle| = |\langle y, Ax\rangle|$, the first identity follows from the last expression of (1.5.1). For the second identity, $\|A^*A\| \le \|A^*\|\,\|A\| = \|A\|^2$ by the first. Moreover, by the Schwarz inequality (1.2.1), we have

$$\|Ax\|^2 = |\langle x, A^*Ax\rangle| \le \|A^*Ax\|\,\|x\| \le \|A^*A\|\,\|x\|^2$$

so that $\|Ax\| \le \|A^*A\|^{1/2}\|x\|$, implying that $\|A\| \le \|A^*A\|^{1/2}$. Hence $\|A\|^2 \le \|A^*A\|$. □

When $\{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathcal{H}$, the *trace* $\mathrm{Tr}\,A$ of $A \in B(\mathcal{H})$ is defined as

$$\mathrm{Tr}\,A := \sum_{i=1}^{n} \langle e_i, Ae_i\rangle.$$

The definition is independent of the choice of an orthonormal basis, as we will see shortly. Obviously, Tr is a linear functional on $\mathbb{M}_n$, which is positive and faithful, i.e., for any $A \ge 0$, $\mathrm{Tr}\,A \ge 0$ and $\mathrm{Tr}\,A = 0$ only if $A = 0$. In fact, the faithfulness is shown since $\mathrm{Tr}\,A = \sum_{i=1}^{n} \|A^{1/2}e_i\|^2 = 0$ implies that $A^{1/2} = 0$ and so $A = 0$. A principal property of the trace is

$$\mathrm{Tr}\,AB = \mathrm{Tr}\,BA \quad \text{for all } A, B \in B(\mathcal{H}).$$

In fact, thanks to (1.2.3),

$$\begin{aligned}\mathrm{Tr}\,AB &= \sum_{i=1}^{n} \langle e_i, ABe_i\rangle = \sum_{i=1}^{n} \langle A^*e_i, Be_i\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{\langle e_j, A^*e_i\rangle}\langle e_j, Be_i\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} \overline{\langle e_i, B^*e_j\rangle}\langle e_i, Ae_j\rangle = \sum_{j=1}^{n} \langle e_j, BAe_j\rangle = \mathrm{Tr}\,BA.\end{aligned}$$

Now, let $\{f_1, \dots, f_n\}$ be another orthonormal basis of $\mathcal{H}$. Then we have a unitary U defined by $Ue_i = f_i$, $1 \le i \le n$, so that

$$\sum_{i=1}^{n}\langle f_i, Af_i\rangle = \sum_{i=1}^{n}\langle Ue_i, AUe_i\rangle = \mathrm{Tr}\, U^*AU = \mathrm{Tr}\, AUU^* = \mathrm{Tr}\, A,$$

which says that the definition of $\mathrm{Tr}\, A$ is actually independent of the choice of an orthonormal basis.

When $A \in \mathbb{M}_n = B(\mathbb{C}^n)$, the trace of A is nothing but the sum of the principal diagonal entries of A:

$$\mathrm{Tr}\, A = a_{11} + a_{22} + \cdots + a_{nn} \quad \text{for} \quad A = [a_{ij}]_{i,j=1}^n.$$

Proposition 1.5.3. *For every $A \in B(\mathcal{H})$, $\mathrm{Tr}\, A$ is the sum of the eigenvalues of A with counting multiplicities.*

Proof. Taking the matrix representation with respect to an orthonormal basis, we may prove the result when A is an $n \times n$ matrix. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A with multiplicities. Then

$$\det(tI - A) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n). \tag{1.5.3}$$

From the usual definition of determinant we notice that the t^{n-1} term in the left-hand side of (1.5.3) appears only from the product $(t - a_{11})(t - a_{22})\cdots(t - a_{nn})$. Hence the coefficient of t^{n-1} in the left-hand side is $-\mathrm{Tr}\, A$. On the other hand, that in the right-hand side of (1.5.3) is $-(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$ so that the assertion follows. □

By use of the trace, define the *Hilbert–Schmidt inner product* on $B(\mathcal{H})$ by

$$\langle A, B\rangle_{\mathrm{HS}} := \mathrm{Tr}\, A^*B, \qquad A, B \in B(\mathcal{H}).$$

It is immediately verified that $\langle\cdot\,,\,\cdot\rangle_{\mathrm{HS}}$ is an inner product on $B(\mathcal{H})$. The norm on $B(\mathcal{H})$ induced by $\langle\cdot\,,\,\cdot\rangle_{\mathrm{HS}}$ is

$$\|A\|_{\mathrm{HS}} := \langle A, A\rangle_{\mathrm{HS}}^{1/2} = (\mathrm{Tr}\, A^*A)^{1/2}, \qquad A \in B(\mathcal{H}),$$

which is called the *Hilbert–Schmidt norm* of A.

For every linear functional $\psi : B(\mathcal{H}) \to \mathbb{C}$, by the Riesz representation theorem, there exists a unique $D_\psi \in B(\mathcal{H})$ such that

$$\psi(X) = \langle D_\psi^*, X\rangle_{\mathrm{HS}} = \mathrm{Tr}\, D_\psi X, \qquad X \in B(\mathcal{H}).$$

The operator D_ψ is sometimes called the *Radon–Nikodym derivative* of ψ with respect to Tr. A linear functional ψ on $B(\mathcal{H})$ is said to be *positive* if $\psi(X) \geq 0$ for all $X \in B(\mathcal{H})^+$, and called a *state* if ψ is positive and $\psi(I) = 1$. The set of all states on $B(\mathcal{H})$ is a convex set, whose extreme points are called *pure states*.

Exercise 1.5.4. Let ω be a linear functional on $B(\mathcal{H})$. Prove:

(1) ω is a state if and only if $D_\omega \geq 0$ and $\mathrm{Tr}\, D_\omega = 1$. Such an operator on $\mathcal{H}$ is called a *density operator*.

(2) ω is a pure state if and only if there is a unit vector $u \in \mathcal{H}$ such that $\omega(X) = \langle u, Xu\rangle$ for $X \in B(\mathcal{H})$ (or equivalently, $D_\omega = |u\rangle\langle u|$).

(3) For each $X \in B(\mathcal{H})$, $X \in B(\mathcal{H})^+$ if and only if $\omega(X) \geq 0$ for all states ω on $B(\mathcal{H})$.

In the case $B(\mathcal{H}) = \mathbb{M}_n$, since

$$\langle A, B\rangle_{\mathrm{HS}} = \sum_{i,j=1}^{n} \overline{a_{ij}} b_{ij}, \qquad A = [a_{ij}],\ B = [b_{ij}],$$

we notice that $\langle\cdot\,,\,\cdot\rangle_{\mathrm{HS}}$ is actually the Hermitian inner product when $\mathbb{M}_n$ is regarded as $\mathbb{C}^{n^2}$. Hence, $(\mathbb{M}_n, \langle\cdot\,,\,\cdot\rangle_{\mathrm{HS}})$ is an n^2-dimensional Hilbert space and the matrix units E_{ij}, $1 \leq i, j \leq n$, form a canonical orthonormal basis.

Exercise 1.5.5. Prove the following inequalities for $A \in \mathbb{M}_n$:

$$\left.\begin{array}{l} \max\{\|a_1\|, \ldots, \|a_n\|\} \\ \max\{\|a_1'\|, \ldots, \|a_n'\|\} \end{array}\right\} \leq \|A\| \leq \|A\|_{\mathrm{HS}},$$

where $a_1, \ldots, a_n$ are the column vectors and $a_1', \ldots, a_n'$ are the row vectors of A, i.e., $A = [a_1\ \cdots\ a_n] = [a_1'\ \cdots\ a_n']^t$.

Exercise 1.5.6.

(1) Show that the following hold for all $A, X \in B(\mathcal{H})$:

$$\|A^*\|_{\mathrm{HS}} = \|A\|_{\mathrm{HS}}, \quad \|AX\|_{\mathrm{HS}} \leq \|A\|\,\|X\|_{\mathrm{HS}}, \quad \|XA\|_{\mathrm{HS}} \leq \|A\|\,\|X\|_{\mathrm{HS}}.$$

(2) For each $A \in B(\mathcal{H})$ define $L_A, R_A : B(\mathcal{H}) \to B(\mathcal{H})$ by the left and the right multiplications:

$$L_A X := AX, \qquad R_A X := XA, \qquad X \in B(\mathcal{H}),$$

which are obviously linear operators on $(B(\mathcal{H}), \langle\cdot\,,\,\cdot\rangle_{\mathrm{HS}})$. Prove:

(a) The operator norms of L_A and R_A are equal to $\|A\|$, i.e., $\|L_A\| = \|R_A\| = \|A\|$.

(b) $(L_A)^* = L_{A^*}$ and $(R_A)^* = R_{A^*}$.

(c) If $A \in B(\mathcal{H})^+$ then $L_A \geq 0$ and $R_A \geq 0$, i.e., $\langle L_A X, X\rangle_{\mathrm{HS}} \geq 0$ and $\langle R_A X, X\rangle_{\mathrm{HS}} \geq 0$ for all $X \in B(\mathcal{H})$, and vice versa.

(3) When $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{M}_2$, find the matrix representations of L_A and R_A on $\mathbb{M}_2$ with respect to the basis $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ of $\mathbb{M}_2$. When $A = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$, find what those are.

The *spectral radius* of $A \in B(\mathcal{H})$ is defined as

$$r(A) := \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Also, the *numerical range* of $A \in B(\mathcal{H})$ is

$$W(A) := \{\langle x, Ax\rangle : x \in \mathcal{H},\ \|x\| = 1\},$$

and the *numerical radius* of A is

$$w(A) := \max\{|\langle x, Ax\rangle| : x \in \mathcal{H},\ \|x\| = 1\}.$$

Proposition 1.5.7.

(1) *For any $A, B \in B(\mathcal{H})$, $\sigma(AB) = \sigma(BA)$ and hence $r(AB) = r(BA)$.*

(2) *The numerical radius $w(\cdot)$ is a norm on $B(\mathcal{H})$.*

(3) *For every $A \in B(\mathcal{H})$,*

$$r(A) \leq w(A) \leq \|A\| \leq 2w(A).$$

(4) *If $A \in B(\mathcal{H})$ is normal, then $r(A) = w(A) = \|A\|$.*

Proof. (1) It is enough to show that $\det(\lambda I - AB) = \det(\lambda I - BA)$ for $A, B \in \mathbb{M}_n$. Assume first that A is invertible. We then have

$$\det(\lambda I - AB) = \det(A^{-1}(\lambda I - AB)A) = \det(\lambda I - BA)$$

and hence $\sigma(AB) = \sigma(BA)$. When A is not invertible, choose a sequence $\{\varepsilon_k\}$ in $\mathbb{C} \setminus \sigma(A)$ with $\varepsilon_k \to 0$, and set $A_k := A - \varepsilon_k I$. Then

$$\det(\lambda I - AB) = \lim_{k\to\infty} \det(\lambda I - A_k B) = \lim_{k\to\infty} \det(\lambda I - BA_k) = \det(\lambda I - BA).$$

(2) It is obvious that $w(A) \geq 0$ and $w(A) = 0$ implies $A = 0$ by Lemma 1.3.1 (2). For every $A, B \in B(\mathcal{H})$ and $x \in \mathcal{H}$, $\|x\| = 1$, we have

$$|\langle x, (A+B)x\rangle| \leq |\langle x, Ax\rangle| + |\langle x, Bx\rangle| \leq w(A) + w(B)$$

and hence $w(A+B) \leq w(A) + w(B)$.

(3) For each $\lambda \in \sigma(A)$ choose a unit vector $v \in \mathcal{H}$ such that $Av = \lambda v$. Then $|\lambda| = |\langle v, Av\rangle| \leq w(A)$. Hence the first inequality holds. The second inequality follows from the Schwarz inequality. The last inequality will be proved after the proof of (4).

(4) Since $r(A) \leq w(A) \leq \|A\|$ was proved in (3), it suffices to show that $\|A\| \leq r(A)$ for normal A. By the Schmidt decomposition (1.4.1) we have

$$\begin{aligned} \|Ax\| &= \left\| \sum_{i=1}^n \lambda_i \langle u_i, x\rangle u_i \right\| = \left(\sum_{i=1}^n |\lambda_i \langle u_i, x\rangle|^2 \right)^{1/2} \\ &\leq \left(\max_{1\leq i\leq n} |\lambda_i| \right) \left(\sum_{i=1}^n |\langle u_i, x\rangle|^2 \right)^{1/2} = r(A)\,\|x\|, \end{aligned}$$

which implies that $\|A\| \leq r(A)$.

Finally, let us prove that $\|A\| \leq 2w(A)$. For any $A \in B(\mathcal{H})$ let

$$B := \frac{1}{2}(A + A^*), \qquad C := \frac{1}{2i}(A - A^*).$$

Then $B, C \in B(\mathcal{H})^{sa}$ and $A = B + iC$ (called the *Descartes decomposition* of A). Hence by (4), we have

$$\|A\| \leq \|B\| + \|C\| = w(B) + w(C).$$

Since $w(A) = w(A^*)$ as immediately seen, $w(B) \leq w(A)$ and $w(C) \leq w(A)$ by (2). Therefore, $\|A\| \leq 2w(A)$. □

Exercise 1.5.8. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Since A has only zero eigenvalues, $r(A) = 0$. Compute $w(A)$ and $\|A\|$.

Note that the spectral radius $r(\cdot)$ is not a norm on $\mathbb{M}_n$. For example, $r(A + A^*) = 1$ but $r(A) = r(A^*) = 0$ for A in the above exercise.

1.6 Tensor product and Schur product

To introduce tensor products of operators/matrices, we start with tensor product for Hilbert spaces. Let $\mathcal{H}$ and $\mathcal{K}$ be finite-dimensional Hilbert spaces with $\dim \mathcal{H} = n$ and $\dim \mathcal{K} = m$. The tensor product vector space $\mathcal{H} \otimes \mathcal{K}$ is abstractly defined as the quotient vector space of the free vector space over $\{x \otimes u : x \in \mathcal{H},\ u \in \mathcal{K}\}$ by the subspace spanned by

$$(x + y) \otimes u - x \otimes u - y \otimes u, \quad x \otimes (u + v) - x \otimes u - y \otimes v,$$
$$(\lambda x) \otimes u - \lambda(x \otimes u), \quad x \otimes (\lambda u) - \lambda(x \otimes u)$$

for all $x, y \in \mathcal{H}$, $u, v \in \mathcal{K}$ and $\lambda \in \mathbb{C}$. Fix orthonormal bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ of $\mathcal{H}$ and $\mathcal{K}$, respectively. Then $\{e_i \otimes f_j : 1 \le i \le n,\ 1 \le j \le m\}$ is a linear basis of $\mathcal{H} \otimes \mathcal{K}$. So $\mathcal{H} \otimes \mathcal{K}$ may be more conveniently introduced as the vector space over $\{e_i \otimes f_j : 1 \le i \le n,\ 1 \le j \le m\}$. We next introduce an inner product on $\mathcal{H} \otimes \mathcal{K}$ by

$$\left\langle \sum_{k=1}^{p} x_k \otimes u_k, \sum_{l=1}^{q} y_l \otimes v_l \right\rangle := \sum_{k=1}^{p} \sum_{l=1}^{q} \langle x_k, y_l \rangle \langle u_k, v_l \rangle \tag{1.6.1}$$

for $x_k, y_l \in \mathcal{H}$ and $u_k, v_l \in \mathcal{K}$. The well-definedness of (1.6.1) is seen by the universality of the tensor product as follows: Let $\mathcal{F}$ be the vector space of all conjugate-linear functionals on $\mathcal{H} \otimes \mathcal{K}$. For each $x \in \mathcal{H}$ and $u \in \mathcal{K}$, since $(y, v) \in \mathcal{H} \times \mathcal{K} \mapsto \langle x, y \rangle \langle u, v \rangle$ is bilinear, we have a unique linear functional $\overline{\varphi(x, u)}$ on $\mathcal{H} \otimes \mathcal{K}$ such that $\varphi(x, u)(y \otimes v) = \langle x, y \rangle \langle u, v \rangle$, $y \in \mathcal{H}$, $v \in \mathcal{K}$. It is easy to see that $(x, u) \in \mathcal{H} \times \mathcal{K} \mapsto \overline{\varphi(x, u)} \in \mathcal{F}$ is a bilinear map, so we have a unique linear map $F : \mathcal{H} \otimes \mathcal{K} \to \mathcal{F}$ such that $F(x \otimes u) = \overline{\varphi(x, u)}$, $x \in \mathcal{H}$, $u \in \mathcal{K}$. Then the form (1.6.1) is written as $\langle \zeta, \eta \rangle = \overline{F(\zeta)}(\eta)$ for $\zeta = \sum_{k=1}^{p} x_k \otimes u_k$ and $\eta = \sum_{l=1}^{q} y_l \otimes v_l$, so (1.6.1) is well-defined. Any $\zeta \in \mathcal{H} \otimes \mathcal{K}$ is written as $\zeta = \sum_{j=1}^{m} x_j \otimes f_j$ with $x_j \in \mathcal{H}$. Since

$$\langle \zeta, \zeta \rangle = \sum_{j,k=1}^{m} \langle x_j, x_k \rangle \langle f_j, f_k \rangle = \sum_{j=1}^{m} \|x_j\|^2,$$

$\langle \zeta, \zeta \rangle = 0$ implies that $x_j = 0$ for all j and so $z = 0$. The other properties of inner product for (1.6.1) are easy to see. Thus $\mathcal{H} \otimes \mathcal{K}$ becomes an nm-dimensional Hilbert space, called the *tensor product Hilbert space* of $\mathcal{H}$ and $\mathcal{K}$, and it is clear that $\{e_i \otimes f_j : 1 \le i \le n,\ 1 \le j \le m\}$ is an orthonormal basis. (Note that the completion procedure is further necessary to define the tensor product of infinite-dimensional Hilbert spaces but the completeness is automatic in the finite-dimensional case.)

For each $A \in B(\mathcal{H})$ and $B \in B(\mathcal{K})$, a linear operator $A \otimes B$, called the *tensor product* of A and B, on $\mathcal{H} \otimes \mathcal{K}$ is defined by

$$(A \otimes B)\left(\sum_{k=1}^{p} x_k \otimes u_k \right) := \sum_{k=1}^{p} Ax_k \otimes Bu_k$$

for $x_k \in \mathcal{H}$ and $u_k \in \mathcal{K}$, where the well-definedness is seen similarly to that of (1.6.1) above.

Now, consider $\mathcal{H} = \mathbb{C}^n$ and $\mathcal{K} = \mathbb{C}^m$ with $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ being the standard bases, respectively. Then $\mathbb{C}^n \otimes \mathbb{C}^m$ is identified with $\mathbb{C}^{nm}$ by arranging the basis

$$\{e_1 \otimes f_1, \ldots, e_1 \otimes f_m, e_2 \otimes f_1, \ldots, e_2 \otimes f_m, \quad \ldots\ldots \quad, e_n \otimes f_1, \ldots, e_n \otimes f_m\}.$$

For $A = [a_{ij}] \in \mathbb{M}_n$ $(= B(\mathbb{C}^n))$ and $B = [b_{kl}] \in \mathbb{M}_m$ $(= B(\mathbb{C}^m))$, the matrix representation of $A \otimes B$ with respect to the above basis is written in the block-matrix form as

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{bmatrix}, \tag{1.6.2}$$

since $\langle e_i \otimes f_k, (A \otimes B)(e_j \otimes f_l) \rangle = a_{ij} b_{kl}$. This form of $A \otimes B$ is often called the *Kronecker product* of A and B.

Proposition 1.6.1. *Let $A, A_1, A_2 \in B(\mathcal{H})$, $B, B_1, B_2 \in B(\mathcal{K})$ and $\alpha_1, \alpha_2 \in \mathbb{C}$. Then*

(1) $(\alpha_1 A_1 + \alpha_2 A_2) \otimes B = \alpha_1 (A_1 \otimes B) + \alpha_2 (A_2 \otimes B)$, $A \otimes (\alpha_1 B_1 + \alpha_2 B_2) = \alpha_1 (A \otimes B) + \alpha_2 (A \otimes B_2)$.

(2) $(A_1 \otimes B_1)(A_2 \otimes B_2) = A_1 A_2 \otimes B_1 B_2$, $(A \otimes B)^* = A^* \otimes B^*$.

(3) *If A, B are invertible, then so is $A \otimes B$ and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.*

(4) *If $A \ge 0$ and $B \ge 0$, then $A \otimes B \ge 0$.*

Proof. The proofs of (1)–(3) are left for exercises. To show (4), let $A \ge 0$ and $B \ge 0$. With the square-roots $A^{1/2} \ge 0$ and $B^{1/2} \ge 0$, we have

$$A \otimes B = (A^{1/2})^2 \otimes (B^{1/2})^2 = (A^{1/2} \otimes B^{1/2})^2 = (A^{1/2} \otimes B^{1/2})^*(A^{1/2} \otimes B^{1/2}) \geq 0$$

by (2) and Corollary 1.4.5. □

Exercise 1.6.2. Show (1)–(3) of the above proposition.

The *Schur product* or the *Hadamard product* $A \circ B$ of two $n \times n$ matrices A and B is defined by the entrywise product as

$$A \circ B := [a_{ij}b_{ij}]_{i,j=1}^n \quad \text{for} \quad A = [a_{ij}]_{i,j=1}^n,\ B = [b_{ij}]_{i,j=1}^n.$$

It is obvious that

$$A \circ B = B \circ A, \quad (A \circ B)^* = A^* \circ B^*, \quad (\lambda A + \mu B) \circ C = \lambda(A \circ C) + \mu(B \circ C)$$

for all $A, B, C \in \mathbb{M}_n$ and $\lambda, \mu \in \mathbb{C}$. An important property of this product is the following *Schur product theorem.*

Theorem 1.6.3. *If $A, B \in \mathbb{M}_n^+$ then $A \circ B \in \mathbb{M}_n^+$.*

Proof. Let $A, B \in \mathbb{M}_n^+$ and represent the tensor product $A \otimes B$ by an $n^2 \times n^2$ matrix as in (1.6.2) with respect to the basis $\{e_1 \otimes e_1, \dots, e_1 \otimes e_n, e_2 \otimes e_1, \dots, e_2 \otimes e_n, \dots\dots, e_n \otimes e_1, \dots, e_n \otimes e_n\}$. Then $A \circ B$ is realized as a principal submatrix of $A \otimes B$ corresponding to a sub-basis $\{e_1 \otimes e_1, e_2 \otimes e_2, \dots, e_n \otimes e_n\}$. We have $A \otimes B \geq 0$ by Proposition 1.6.1 (4). It is clear that a principal submatrix of a positive semidefinite matrix is also positive semidefinite. Hence $A \circ B \geq 0$ follows. □

Exercise 1.6.4. Prove the following:
(1) If $A, B \in \mathbb{M}_n^{sa}$ then $A \circ B \in \mathbb{M}_n^{sa}$.
(2) If $A_1 \geq A_2 \geq 0$ and $B_1 \geq B_2 \geq 0$ in $\mathbb{M}_n$ then $A_1 \circ B_1 \geq A_2 \circ B_2$.
(3) If $A > 0$ and $B > 0$ in $\mathbb{M}_n$ then $A \circ B > 0$.

1.7 Positive maps and completely positive maps

Let $\mathcal{H}$ and $\mathcal{K}$ be finite-dimensional Hilbert spaces, and let $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ be a linear map. The map Φ is said to be *positive* if $\Phi(A) \in B(\mathcal{K})^+$ for all $A \in B(\mathcal{H})^+$, and Φ is said to be *unital* if $\Phi(I_{\mathcal{H}}) = I_{\mathcal{K}}$.

Exercise 1.7.1. Show that if Φ is positive, then $\Phi(A^*) = \Phi(A)^*$ for all $A \in B(\mathcal{H})$; in particular, $\Phi(A) \in B(\mathcal{K})^{sa}$ for all $A \in B(\mathcal{H})^{sa}$.

The method using 2×2-block matrices will be useful in later discussions, so let us first prepare some basics on 2×2 block matrices. For two (finite-dimensional) Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ consider the *direct sum Hilbert space*

$$\mathcal{H}_1 \oplus \mathcal{H}_2 := \{x_1 \oplus x_2 : x_1 \in \mathcal{H}_1,\ x_2 \in \mathcal{H}_2\}$$

equipped with the inner product

$$\langle x_1 \oplus x_2, y_1 \oplus y_2 \rangle := \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle, \qquad x_1, y_1 \in \mathcal{H}_1,\ x_2, y_2 \in \mathcal{H}_2.$$

A general $\mathbf{A} \in B(\mathcal{H}_1 \oplus \mathcal{H}_2)$ is represented as a 2×2 block matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

with $A_{11} \in B(\mathcal{H}_1)$, $A_{12} \in B(\mathcal{H}_2, \mathcal{H}_1)$, $A_{21} \in B(\mathcal{H}_1, \mathcal{H}_2)$ and $A_{22} \in B(\mathcal{H}_2)$, which acts on $\mathcal{H}_1 \oplus \mathcal{H}_2$ in the matrix form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11}x_1 + A_{12}x_2 \\ A_{21}x_1 + A_{22}x_2 \end{bmatrix}$$

for $x_1 \oplus x_2 \in \mathcal{H}_1 \oplus \mathcal{H}_2$ and $y_1 \oplus y_2 = A(x_1 \oplus x_2)$. Then the product and the adjoint in $B(\mathcal{H}_1 \oplus \mathcal{H}_2)$ are computed in the conventional way for matrices. In particular, when $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$, one can identify $\mathcal{H} \oplus \mathcal{H}$ with the tensor product Hilbert space $\mathbb{C}^2 \otimes \mathcal{H}$ and then $B(\mathcal{H} \oplus \mathcal{H}) = B(\mathbb{C}^2 \otimes \mathcal{H})$ with

$$\mathbb{M}_2 \otimes B(\mathcal{H}) = \left\{ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} : A_{ij} \in B(\mathcal{H}) \right\}.$$

The next lemma will be of some use.

Lemma 1.7.2. *If $C \in B(\mathcal{H})$, then $\begin{bmatrix} I & C \\ C^* & I \end{bmatrix} \geq 0$ if and only if $\|C\| \leq 1$.*

Proof. This is seen as

$$\begin{bmatrix} I & C \\ C^* & I \end{bmatrix} \geq 0 \iff \left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} I & C \\ C^* & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle \geq 0 \quad \text{for all } x, y \in \mathcal{H}$$
$$\iff \langle x, x\rangle + 2\,\mathrm{Re}\langle x, Cy\rangle + \langle y, y\rangle \geq 0 \quad \text{for all } x, y \in \mathcal{H}$$
$$\iff |\langle x, Cy\rangle| \leq \frac{\langle x, x\rangle + \langle y, y\rangle}{2} \quad \text{for all } x, y \in \mathcal{H}$$
$$\iff \|C\| \leq 1. \qquad \square$$

When Φ is positive and unital, *Kadison's inequality*

$$\Phi(A)^2 \leq \Phi(A^2)$$

holds for every $A \in B(\mathcal{H})^{sa}$. The following is an extension of Kadison's inequality due to Choi [26]. The proof below is from Bhatia [14, Sect. 2.3] (also [7, Sect. 4]).

Theorem 1.7.3. *If Φ is positive and unital, then*

$$\Phi(A)^*\Phi(A) \leq \Phi(A^*A)$$

for every normal $A \in B(\mathcal{H})$.

Proof. Let $A = \sum_{j=1}^m \alpha_j P_j$ be the spectral decomposition. Then $A^* = \sum_j \overline{\alpha_j} P_j$ and $A^*A = \sum_j |\alpha_j|^2 P_j$. Under the identification $B(\mathbb{C}^2 \otimes \mathcal{K}) = \mathbb{M}_2 \otimes B(\mathcal{K})$, since $I = \Phi(I) = \sum_j \Phi(P_j)$, we can write

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A)^* \\ \Phi(A) & I \end{bmatrix} = \sum_{j=1}^m \begin{bmatrix} |\alpha_j|^2\Phi(P_j) & \overline{\alpha_j}\Phi(P_j) \\ \alpha_j\Phi(P_j) & \Phi(P_j) \end{bmatrix} = \sum_{j=1}^m \begin{bmatrix} |\alpha_j|^2 & \overline{\alpha_j} \\ \alpha_j & 1 \end{bmatrix} \otimes \Phi(P_j).$$

Since $\begin{bmatrix} |\alpha_j|^2 & \overline{\alpha_j} \\ \alpha_j & 1 \end{bmatrix} \in \mathbb{M}_2^+$ and $\Phi(P_j) \in B(\mathcal{K})^+$ by positivity of Φ, it follows that

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A)^* \\ \Phi(A) & I \end{bmatrix} \geq 0.$$

Hence it suffices to show that $\begin{bmatrix} B & C^* \\ C & I \end{bmatrix} \geq 0$ if and only if $C^*C \leq B$. To prove this, we may assume by continuity that B is invertible. Then $\begin{bmatrix} B & C^* \\ C & I \end{bmatrix} \geq 0$ is equivalent to

$$\begin{bmatrix} B^{-1/2} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} B & C^* \\ C & I \end{bmatrix}\begin{bmatrix} B^{-1/2} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & B^{-1/2}C^* \\ CB^{-1/2} & I \end{bmatrix} \geq 0.$$

By Lemma 1.7.2 this is also equivalent to $\|CB^{-1/2}\| \leq 1$, that is, $B^{-1/2}C^*CB^{-1/2} \leq I$ or $C^*C \leq B$. $\square$

The following is also due to Choi [26] with proof from [7, 14].

Theorem 1.7.4. *If Φ is positive and unital, then*

$$\Phi(A)^{-1} \leq \Phi(A^{-1})$$

for every invertible $A \in B(\mathcal{H})^+$.

Proof. Since $A \geq \delta I$ for some $\delta > 0$, we have $\Phi(A) \geq \delta\Phi(I) = \delta I$ so that $\Phi(A)$ is invertible too. Let $A = \sum_{j=1}^m \alpha_j P_j$ be the spectral decomposition with $\alpha_j > 0$. Then $A^{-1} = \sum_{j=1}^m \alpha_j^{-1} P_j$ and hence

$$\begin{bmatrix} \Phi(A^{-1}) & I \\ I & \Phi(A) \end{bmatrix} = \sum_{j=1}^m \begin{bmatrix} \alpha_j^{-1} & 1 \\ 1 & \alpha_j \end{bmatrix} \otimes \Phi(P_j) \geq 0$$

since $\begin{bmatrix} \alpha_j^{-1} & 1 \\ 1 & \alpha_j \end{bmatrix} \geq 0$. Multiplying $\begin{bmatrix} I & 0 \\ 0 & \Phi(A)^{-1/2} \end{bmatrix}$ from both sides of the above gives

$$\begin{bmatrix} \Phi(A^{-1}) & \Phi(A)^{-1/2} \\ \Phi(A)^{-1/2} & I \end{bmatrix} \geq 0.$$

By the argument in the proof of Theorem 1.7.3 we have $\Phi(A)^{-1} \leq \Phi(A^{-1})$. $\square$

Theorem 1.7.5. *If Φ is positive and unital, then $\|\Phi\| = 1$, where*

$$\|\Phi\| := \sup_{X \neq 0} \frac{\|\Phi(X)\|}{\|X\|}.$$

Proof. Since $\Phi(I) = I$, $\|\Phi\| \geq 1$. We need to show that $\|\Phi(A)\| \leq 1$ for all $A \in B(\mathcal{H})$ with $\|A\| \leq 1$. For any unitary $U \in B(\mathcal{H})$, by Theorem 1.7.3 we have

$$\Phi(U)^*\Phi(U) \leq \Phi(U^*U) = \Phi(I) = I$$

so that $\|\Phi(U)\| \leq 1$. For every $A \in B(\mathcal{H})$ with $\|A\| \leq 1$, take the polar decomposition $A = U|A|$ with a unitary U and the spectral decomposition $|A| = \sum_{j=1}^m \alpha_j P_j$. Since $0 \leq \alpha_j \leq 1$, $\alpha_j = \cos\theta_j$ with $0 \leq \theta_j \leq \pi/2$. Define $V_1 := \sum_j e^{i\theta_j} P_j$ and $V_2 := \sum_j e^{-i\theta_j} P_j$. Then V_1 and V_2 are unitaries and $|A| = (V_1 + V_2)/2$ so that $A = (UV_1 + UV_2)/2$. Hence

$$\|\Phi(A)\| \leq \frac{\|\Phi(UV_1)\| + \|\Phi(UV_2)\|}{2} \leq 1,$$

as required. □

Proposition 1.7.6. *If Φ is positive, then $\|\Phi\| = \|\Phi(I)\|$.*

Proof. For each $\varepsilon > 0$ define a linear map $\Phi_\varepsilon : B(\mathcal{H}) \to B(\mathcal{K})$ by

$$\Phi_\varepsilon(A) := \Phi(A) + \varepsilon(\mathrm{Tr}\, A)I_{\mathcal{K}}, \qquad A \in B(\mathcal{H}).$$

Then Φ_ε is positive and $\Phi_\varepsilon(I)$ is invertible. So, further define a linear map $\Psi_\varepsilon : B(\mathcal{H}) \to B(\mathcal{K})$ by

$$\Psi_\varepsilon(A) := \Phi_\varepsilon(I)^{-1/2}\Phi_\varepsilon(A)\Phi_\varepsilon(I)^{-1/2}, \qquad A \in B(\mathcal{H}),$$

which is positive and unital. By Theorem 1.7.5, $\|\Psi_\varepsilon\| = 1$. Hence, if $\|A\| \leq 1$ then

$$\|\Phi_\varepsilon(A)\| = \|\Phi_\varepsilon(I)^{1/2}\Psi_\varepsilon(A)\Phi_\varepsilon(I)^{1/2}\| \leq \|\Phi_\varepsilon(I)^{1/2}\|^2 = \|\Phi_\varepsilon(I)\|.$$

Letting $\varepsilon \searrow 0$ gives $\|\Phi(A)\| \leq \|\Phi(I)\|$. Therefore, $\|\Phi\| = \|\Phi(I)\|$. □

Let $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ and $\tilde{\Phi} : B(\tilde{\mathcal{H}}) \to B(\tilde{\mathcal{K}})$ be two linear maps. Then the tensor product map

$$\Phi \otimes \tilde{\Phi} : B(\mathcal{H}) \otimes B(\tilde{\mathcal{H}}) = B(\mathcal{H} \otimes \tilde{\mathcal{H}}) \longrightarrow B(\mathcal{K}) \otimes B(\tilde{\mathcal{K}}) = B(\mathcal{K} \otimes \tilde{\mathcal{K}})$$

is defined by setting

$$(\Phi \otimes \tilde{\Phi})(A \otimes \tilde{A}) := \Phi(A) \otimes \tilde{\Phi}(\tilde{A}), \qquad A \in B(\mathcal{H}),\ \tilde{A} \in B(\tilde{\mathcal{H}}),$$

and by extending it by linearity. Here it might be expected that if both Φ and $\tilde{\Phi}$ are positive then so is $\Phi \otimes \tilde{\Phi}$. However, it is not true in general. This inconvenience tells us that the notion of positivity for linear maps is not very suitable in the noncommutative setting for matrices (operators). This is the reason why we need a stronger notion of positivity instead of simple positivity.

For each $n \in \mathbb{N}$ the map Φ is said to be *n-positive* if $\mathrm{id}_n \otimes \Phi : \mathbb{M}_n \otimes B(\mathcal{H})$ is positive, where id_n denotes the identity map on $\mathbb{M}_n$. In a slightly more concrete notation with block matrices, Φ is n-positive if and only if

$$\begin{bmatrix} \Phi(A_{11}) & \cdots & \Phi(A_{1n}) \\ \vdots & \ddots & \vdots \\ \Phi(A_{n1}) & \cdots & \Phi(A_{nn}) \end{bmatrix} \geq 0 \quad \text{in } \mathbb{M}_n \otimes B(\mathcal{K})$$

whenever

$$\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \geq 0 \quad \text{in } \mathbb{M}_n \otimes B(\mathcal{H}).$$

Clearly, 1-positivity means the usual positivity, and n-positivity implies m-positivity with $m < n$. Furthermore, Φ is said to be *completely positive* (often abbreviated as *CP*) if it is n-positive for every $n \in \mathbb{N}$. Since

$$\Phi \otimes \tilde{\Phi} = (\mathrm{id}_{B(\mathcal{K})} \otimes \tilde{\Phi})(\Phi \otimes \mathrm{id}_{B(\tilde{\mathcal{H}})}),$$

it is obvious that $\Phi \otimes \tilde{\Phi}$ is positive (indeed, completely positive) if both Φ and $\tilde{\Phi}$ are completely positive. Thus, the complete positivity is a satisfactory notion for linear maps between noncommutative $*$-algebras. For example, in quantum physics and quantum probability, a system is usually given by the $*$-algebra $B(\mathcal{H})$ over a Hilbert space $\mathcal{H}$, and the composite system of the two $B(\mathcal{H})$ and $B(\tilde{\mathcal{H}})$ is described by the tensor product $B(\mathcal{H}) \otimes B(\tilde{\mathcal{H}})$, so the positivity of the tensor product of linear maps is essential.

Proposition 1.7.7. *If Φ is 2-positive and unital, then it is a Schwarz map in the sense that*

$$\Phi(A)^*\Phi(A) \le \Phi(A^*A), \qquad A \in B(\mathcal{H}).$$

Proof. For every $A \in B(\mathcal{H})$, since $\begin{bmatrix} A^*A & A^* \\ A & I \end{bmatrix} \ge 0$ in $\mathbb{M}_2 \otimes B(\mathcal{H})$, we have

$$\begin{bmatrix} \Phi(A^*A) & \Phi(A)^* \\ \Phi(A) & I \end{bmatrix} \ge 0.$$

As in the proof of Theorem 1.7.3, this implies that $\Phi(A)^*\Phi(A) \le \Phi(A^*A)$. □

In this way, we have the following implications for unital linear maps:

$$\text{CP} \Rightarrow \quad \cdots \quad \Rightarrow \text{3-positivity} \Rightarrow \text{2-positivity} \Rightarrow \text{Schwarz map} \Rightarrow \text{positivity}.$$

Example 1.7.8. Let Φ be the transpose map on $\mathbb{M}_n$, $n \ge 2$, i.e.,

$$\Phi(A) = A^t = [a_{ji}]_{ij} \quad \text{for } A = [a_{ij}]_{ij}.$$

Then it is obvious that Φ is a positive and unital linear map. But it is not a Schwarz map. Indeed, assume that Φ is a Schwarz map. Then $(A^t)^*A^t \le (A^*A)^t$, i.e., $\overline{A}A^t \le A^t\overline{A}$ for all $A \in \mathbb{M}_n$. Replacing $\overline{A}$ with A we must have $AA^* \le A^*A$. But of course, this does not hold in general. In particular, the transpose map on $\mathbb{M}_2$ is positive but not 2-positive.

There are nice characterizations and representations for completely positive maps, which are summarized in the next theorem without proofs. For the details see [14, Chapter 2] and [66, §11.7] for example.

Theorem 1.7.9. *Let $\Phi : B(\mathcal{H}) \to B(\mathcal{K})$ be a linear map with $n = \dim \mathcal{H}$ and $m = \dim \mathcal{K}$. Then the following conditions are equivalent:*

(i) *Φ is completely positive.*

(ii) *For the matrix units E_{ij}, $1 \le i, j \le n$, of $B(\mathcal{H})$,*

$$\begin{bmatrix} \Phi(E_{11}) & \cdots & \Phi(E_{1n}) \\ \vdots & \ddots & \vdots \\ \Phi(E_{n1}) & \cdots & \Phi(E_{nn}) \end{bmatrix} \ge 0 \quad \textit{in } \mathbb{M}_n \otimes B(\mathcal{K}).$$

(iii) *There are operators $V_i : \mathcal{H} \to \mathcal{K}$, $1 \le i \le r$, such that*

$$\Phi(A) = \sum_{i=1}^{r} V_i A V_i^*, \qquad A \in B(\mathcal{H}),$$

where r can be chosen at most nm.

(iv) *For any $A_1, \ldots, A_k \in B(\mathcal{H})$ and $B_1, \ldots, B_k \in B(\mathcal{K})$ with any $k \in \mathbb{N}$,*

$$\sum_{i,j=1}^{k} B_i^* \Phi(A_i^* A_j) B_j \ge 0$$

holds.

(v) *There are a Hilbert space $\tilde{K}$, a representation (or $*$-homomorphism) $\pi : B(\mathcal{H}) \to B(\tilde{\mathcal{K}})$, and an operator $V : \mathcal{K} \to \tilde{\mathcal{K}}$ such that*

$$\Phi(A) = V^*\pi(A)V, \qquad A \in B(\mathcal{H}),$$

where $\dim \tilde{\mathcal{K}}$ *can be chosen at most n^2m.*

The above characterization (ii) is due to Choi, and the block matrix there is often called the *Choi matrix*. The representation in (iii) is called the *Kraus representation*. The representation in (v) is famous as the *Stinespring representation*. These representations are quite useful to treat completely positive maps.

When Φ is the transpose map on $\mathbb{M}_2$, the Choi matrix of Φ is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which is not positive semidefinite since the determinant is -1. This re-proves the last statement in Remark 1.7.8.

From (ii) we see that if Φ is n-positive where $n = \dim \mathcal{H}$, then it is completely positive. This can be slightly generalized in such a way that Φ is k-positive with $k := \min\{\dim \mathcal{H}, \dim \mathcal{K}\}$, then it is completely positive. If a positive linear map Φ has the range included in a commutative $*$-subalgebra (in particular, Φ is a positive linear functional), then it is completely positive. More generally, we may consider a linear map $\Phi : \mathcal{A} \to \mathcal{B}$, where $\mathcal{A} \subset B(\mathcal{H})$ and $\mathcal{B} \subset B(\mathcal{K})$ are $*$-subalgebras, and the positivity and complete positivity are similarly defined for Φ. If $\mathcal{A}$ or $\mathcal{B}$ is commutative and Φ is positive, then Φ is completely positive. A typical example of completely positive maps is a *conditional expectation*, i.e., a positive linear map E from a $*$-algebra $\mathcal{A}$ ($\subset B(\mathcal{H})$) onto a unital (i.e., $I_{\mathcal{B}} = I_{\mathcal{A}}$) $*$-subalgebra $\mathcal{B}$ of $\mathcal{A}$ such that $E(B_1 A B_2) = B_1 E(A) B_2$ for all $A \in \mathcal{A}$ and $B_1, B_2 \in \mathcal{B}$.

Exercise 1.7.10. Show the following statements.

(1) The linear maps

$$\mathrm{Tr}_{\mathcal{K}} : B(\mathcal{H}) \otimes B(\mathcal{K}) \to B(\mathcal{H}), \quad \mathrm{Tr}_{\mathcal{H}} : B(\mathcal{H}) \otimes B(\mathcal{K}) \to B(\mathcal{K})$$

are uniquely determined by

$$\mathrm{Tr}_{\mathcal{K}}(A \otimes B) = (\mathrm{Tr}\, B)A, \quad \mathrm{Tr}_{\mathcal{H}}(A \otimes B) = (\mathrm{Tr}\, A)B, \qquad A \in B(\mathcal{H}),\ B \in B(\mathcal{K}).$$

These $\mathrm{Tr}_{\mathcal{K}}$ and $\mathrm{Tr}_{\mathcal{H}}$ are called the *partial traces*.

(2) $(\dim \mathcal{K})^{-1}\, \mathrm{Tr}_{\mathcal{K}}$ and $(\dim \mathcal{H})^{-1}\, \mathrm{Tr}_{\mathcal{H}}$ are conditional expectations from $B(\mathcal{H}) \otimes B(\mathcal{K})$ onto the $*$-subalgebras $B(\mathcal{H}) = B(\mathcal{H}) \otimes I_{\mathcal{K}}$ and $B(\mathcal{K}) = I_{\mathcal{H}} \otimes B(\mathcal{K})$, respectively.

(3) $\mathrm{Tr}_{\mathcal{K}}$ and $\mathrm{Tr}_{\mathcal{H}}$ are completely positive. (This follows from (2) and a general result mentioned above, however prove this directly.)

2. Operator Monotone and Operator Convex Functions

2.1 Definitions of operator monotonicity and convexity

Throughout this chapter, let $\mathcal{H}$ be a finite-dimensional Hilbert space. The following is the famous *Löwner–Heinz inequality*, which provides essential examples of operator monotone functions.

Theorem 2.1.1. *For every $A, B \in B(\mathcal{H})^+$, $A \geq B$ implies $A^p \geq B^p$ for all $p \in [0,1]$ with the convention $A^0 := I$.*

Proof. The following proof is due to Pedersen. Assume first that $A \geq B > 0$, and set

$$\Delta := \{p \in \mathbb{R} : A^p \geq B^p\}.$$

Since A^p, B^p are continuous in p, Δ is a closed set. Clearly, $0, 1 \in \Delta$. Hence, to prove that $\Delta \supset [0,1]$, it suffices to show that if $p, q \in \Delta$ then $(p+q)/2 \in \Delta$. So, assume that $A^p \geq B^p$ and $A^q \geq B^q$. Then $A^{-p/2} B^p A^{-p/2} \leq A^{-p/2} A^p A^{-p/2} = I$ so that by Proposition 1.5.2

$$\|B^{p/2} A^{-p/2}\|^2 = \|(B^{p/2} A^{-p/2})^* (B^{p/2} A^{-p/2})\| = \|A^{-p/2} B^p A^{-p/2}\| \leq 1.$$

Hence $\|B^{p/2} A^{-p/2}\| \leq 1$ and similarly $\|B^{q/2} A^{-q/2}\| \leq 1$. Therefore, using (3), (1) and (4) of Proposition 1.5.7 we have

$$\begin{aligned} 1 &\geq \|(B^{p/2} A^{-p/2})^* (B^{q/2} A^{-q/2})\| = \|A^{-p/2} B^{(p+q)/2} A^{-q/2}\| \\ &\geq r(A^{-p/2} B^{(p+q)/2} A^{-q/2}) = r(A^{(q-p)/4} A^{-(p+q)/4} B^{(p+q)/2} A^{-q/2}) \\ &= r(A^{-(p+q)/4} B^{(p+q)/2} A^{-q/2} A^{(q-p)/4}) = r(A^{-(p+q)/4} B^{(p+q)/2} A^{-(p+q)/4}) \\ &= \|A^{-(p+q)/4} B^{(p+q)/2} A^{-(p+q)/4}\|. \end{aligned}$$

This implies that $I \geq A^{-(p+q)/4} B^{(p+q)/2} A^{-(p+q)/4}$ and so $A^{(p+q)/2} \geq B^{(p+q)/2}$, i.e., $(p+q)/2 \in \Delta$. Hence the assertion follows when A, B are invertible.

When $A \geq B \geq 0$, for any $\varepsilon > 0$ we have $A + \varepsilon I \geq B + \varepsilon I > 0$ so that $(A + \varepsilon I)^p \geq (B + \varepsilon I)^p$ for all $p \in [0,1]$. Letting $\varepsilon \searrow 0$ gives the assertion. □

The inequality in the above theorem does not hold when $p > 1$, as the next example shows.

Example 2.1.2. Consider $A := \begin{bmatrix} 3/2 & 0 \\ 0 & 3/4 \end{bmatrix}$ and $B := \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$. Then $A \geq B \geq 0$ is immediately checked.

Since B is an orthogonal projection, for each $p > 0$ we have $B^p = B$ and

$$A^p - B^p = \begin{bmatrix} (3/2)^p - 1/2 & -1/2 \\ -1/2 & (3/4)^p - 1/2 \end{bmatrix}.$$

Compute

$$\det(A^p - B^p) = \left(\frac{3}{8}\right)^p \left(3^p - \frac{2^p + 4^p}{2}\right).$$

If $A^p \geq B^p$ then we must have $\det(A^p - B^p) \geq 0$ so that $(2^p + 4^p)/2 \leq 3^p$, which is not satisfied when $p > 1$. Hence $A^p \geq B^p$ does not hold for any $p > 1$.

Definition 2.1.3. Let J be an interval (whichever closed or open) of $\mathbb{R}$ and f be a real-valued function on J.

(1) It is said that f is *matrix monotone of degree n* or *n-monotone* if, for every $A, B \in \mathbb{M}_n^{sa}$ with $\sigma(A), \sigma(B) \subset J$,

$$A \geq B \quad \text{implies} \quad f(A) \geq f(B).$$

If f is n-monotone for every $n \in \mathbb{N}$ (or the above property holds for every $A, B \in B(\mathcal{H})$ with arbitrary $\mathcal{H}$)), then f is said to be *operator monotone*.

(2) It is said that f is *matrix convex of degree n* or *n-convex* if

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda)f(B), \qquad \lambda \in (0,1), \tag{2.1.1}$$

for all $A, B \in \mathbb{M}_n^{sa}$ with $\sigma(A), \sigma(B) \subset J$. Note that when f is a continuous function on J, the mid-point convexity, i.e., $f((A+B)/2) \leq (f(A) + f(B))/2$ for all A, B as above is enough for f to be n-convex. If f is n-convex for every $n \in \mathbb{N}$ (or the above convexity property holds for every $A, B \in B(\mathcal{H})$ with arbitrary $\mathcal{H}$), then f is said to be *operator convex*. Also, f is said to be *n-concave* or *operator concave* if $-f$ is n-convex or operator convex, respectively.

Exercise 2.1.4. (1) Show that the square function t^2 on $\mathbb{R}$ is operator convex, i.e.,

$$\left(\frac{A+B}{2}\right)^2 \leq \frac{A^2+B^2}{2} \quad \text{for all } A, B \in \mathbb{M}_n^{sa}.$$

(2) For $A, B \in \mathbb{M}_n^{sa}$ and $\varepsilon > 0$, check that

$$\frac{(A+\varepsilon B)^3 + (A-\varepsilon B)^3}{2} - \left(\frac{(A+\varepsilon B)+(A-\varepsilon B)}{2}\right)^3 = \varepsilon^2(AB^2 + BAB + B^2A).$$

To show that t^3 on $[0,\infty)$ is not operator convex, find an example of $A > 0$ and $B \geq 0$ such that $AB^2 + BAB + B^2A \not\geq 0$.

2.2 Divided differences

In this and the next sections we prepare a certain differentiation technique, which will play a key role in Section 2.4.

Let f be a real-valued function on an open interval (a, b), where $-\infty \leq a < b \leq \infty$. Let $x_1, x_2, \dots$ be distinct points in (a, b). We have an important notion of divided differences defined as follows:

$$f^{[0]}(x_1) := f(x_1), \qquad f^{[1]}(x_1, x_2) := \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

and recursively for $n = 2, 3, \dots,$

$$f^{[n]}(x_1, x_2, \dots, x_{n+1}) := \frac{f^{[n-1]}(x_1, x_2, \dots, x_n) - f^{[n-1]}(x_2, x_3, \dots, x_{n+1})}{x_1 - x_{n+1}}.$$

We call $f^{[1]}$, $f^{[2]}$ and $f^{[n]}$ the *first*, the *second*, and the *nth divided differences*, respectively, of f. By induction on n it is easy to see that

$$f^{[n]}(x_1, x_2, \dots, x_{n+1}) = \sum_{k=1}^{n+1} \frac{f(x_k)}{\prod_{1 \leq j \leq n+1,\, j \neq k}(x_k - x_j)}. \tag{2.2.1}$$

This expression shows that $f^{[n]}(x_1, x_2, \dots, x_{n+1})$ is symmetric in the arguments $x_1, x_2, \dots, x_{n+1}$, i.e., invariant under permutations of the arguments.

Exercise 2.2.1. Verify the expression (2.2.1).

Exercise 2.2.2. When $f(x) = x^k$ with $k \in \mathbb{N}$, verify that

$$f^{[n]}(x_1, x_2, \dots, x_{n+1}) = \sum_{\substack{l_1, l_2, \dots, l_{n+1} \geq 0 \\ l_1 + l_2 + \cdots + l_{n+1} = k-n}} x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n} x_{n+1}^{l_{n+1}}.$$

(Hence, $f^{[k]} \equiv 1$ and $f^{[n]} \equiv 0$ if $n > k$.)

Lemma 2.2.3. *Let $\lambda_1, \lambda_2, \dots$ be distinct points in (a, b), and define polynomials*

$$p_0(x) := 1, \quad p_k(x) := \prod_{j=1}^{k} (x - \lambda_j) \quad \textit{for } k \geq 1.$$

Then

$$p_k^{[n]}(\lambda_1, \lambda_2, \ldots, \lambda_{n+1}) = \begin{cases} 1 & \textit{if } n = k, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Let $k \geq 0$ be arbitrarily fixed. By induction on n it is easy to see that

$$p_k^{[n]}(\lambda_m, \lambda_{m+1}, \ldots, \lambda_{m+n}) = 0 \quad \text{if } 1 \leq m \leq m+n \leq k.$$

Hence $p_k^{[n]}(\lambda_1, \ldots, \lambda_{n+1}) = 0$ for all $n < k$. By (2.2.1) we have

$$p_k^{[k]}(\lambda_1, \ldots, \lambda_{k+1}) = \frac{p_k(\lambda_{k+1})}{\prod_{j=1}^{k}(\lambda_{k+1} - \lambda_j)} = 1.$$

This implies that

$$p_k^{[k+1]}(\lambda_1, \ldots, \lambda_{k+2}) = \frac{p_k^{[k]}(\lambda_1, \ldots, \lambda_k, \lambda_{k+1}) - p_k^{[k]}(\lambda_1, \ldots, \lambda_k, \lambda_{k+2})}{\lambda_{k+1} - \lambda_{k+2}} = 0,$$

and recursively $p_k^{[n]}(\lambda_1, \ldots, \lambda_{n+1}) = 0$ for all $n > k$. □

The divided differences of f can be defined without restriction of the x_i's being distinct when f is sufficiently many times differentiable, as the next lemma shows.

Lemma 2.2.4. *Let $n \in \mathbb{N}$ and assume that f is C^n on (a, b). Then*

(1) *The nth divided difference $f^{[n]}(x_1, x_2, \ldots, x_{n+1})$ is extended to a symmetric continuous function on the whole $(a, b)^{n+1}$.*

(2) *For every $\lambda_1, \lambda_2 \ldots, \lambda_{n+1} \in (a, b)$ there exists a point ξ in the smallest interval containing the λ_i's such that*

$$f^{[n]}(\lambda_1, \ldots, \lambda_{n+1}) = \frac{f^{(n)}(\xi)}{n!}.$$

In particular,

$$f^{[n]}(x, x, \ldots, x) = \frac{f^{(n)}(x)}{n!}, \qquad x \in (a, b).$$

Proof. First we prove (2) when $\lambda_1, \ldots, \lambda_{n+1}$ are distinct. From the symmetry of $f^{[n]}$ we may assume that $\lambda_1 < \lambda_2 < \cdots < \lambda_{n+1}$. With the polynomials p_k, $0 \leq k \leq n$, given in Lemma 2.2.3, we consider the C^n function

$$h(x) := f(x) - \sum_{k=0}^{n} f^{[k]}(\lambda_1, \lambda_2, \ldots, \lambda_{k+1}) p_k(x), \qquad x \in (a, b).$$

By Lemma 2.2.3 we have

$$\begin{aligned} h^{[m]}(\lambda_1, \ldots, \lambda_{m+1}) &= f^{[m]}(\lambda_1, \ldots, \lambda_{m+1}) - \sum_{k=0}^{n} f^{[k]}(\lambda_1, \ldots, \lambda_{k+1}) p_k^{[m]}(\lambda_1, \ldots, \lambda_{m+1}) \\ &= f^{[m]}(\lambda_1, \ldots, \lambda_{m+1}) - f^{[m]}(\lambda_1, \ldots, \lambda_{m+1}) = 0 \end{aligned}$$

for all $m = 0, 1, \ldots, n$. Thanks to the expression (2.2.1) this implies that $h(\lambda_k) = 0$ for all $k = 1, \ldots, n+1$. Hence Rolle's theorem implies that there are $\xi_1, \ldots, \xi_n$ such that $\lambda_1 < \xi_1 < \lambda_2 < \xi_2 < \cdots < \lambda_n < \xi_n < \lambda_{n+1}$ and $h'(\xi_k) = 0$ for all $k = 1, \ldots, n$. Repeating this argument yields a $\xi \in (\lambda_1, \lambda_{n+1})$ such that $h^{(n)}(\xi) = 0$. Since

$$h^{(n)}(\xi) = f^{(n)}(\xi) - f^{[n]}(\lambda_1, \ldots, \lambda_{n+1})\, n!,$$

we have

$$f^{[n]}(\lambda_1, \ldots, \lambda_{n+1}) = \frac{f^{(n)}(\xi)}{n!}.$$

Secondly we prove (1) by induction on n. The initial case $n = 0$ trivially holds since the statement is just the C^1 of $f^{[0]}(x_1) = f(x_1)$. Let us prove the statement for n under the assumption of that for $n-1$, where $n \geq 1$. Assume now that f is C^n on (a, b), so by the induction hypothesis, $f^{[n-1]}(x_1, \ldots, x_n)$ is a symmetric continuous function on $(a, b)^n$. We have to extend $f^{[n]}(x_1, \ldots, x_{n+1})$ originally defined for distinct $x_1, \ldots, x_{n+1}$ in (a, b) to the whole $(a, b)^{n+1}$. For each $(x_1, \ldots, x_{n+1}) \in (a, b)^{n+1}$ choose a sequence $\{(x_1^{(k)}, \ldots, x_{n+1}^{(k)})\}$ in $(a, b)^{n+1}$ such that $x_1^{(k)}, \ldots, x_{n+1}^{(k)}$ are distinct and $x_j^{(k)} \to x_j$ as $k \to \infty$ for all $j = 1, \ldots, n+1$. Assume that $x_i \neq x_j$ for some $i, j \in \{1, \ldots, n+1\}$. Since $x_i^{(k)} \neq x_j^{(k)}$ for all k, we have

$$
\begin{aligned}
&f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})\\
&\quad=\frac{f^{[n-1]}(x_1^{(k)},\dots,x_{j-1}^{(k)},x_{j+1}^{(k)},\dots,x_n^{(k)})-f^{[n-1]}(x_1^{(k)},\dots,x_{i-1}^{(k)},x_{i+1}^{(k)},\dots,x_{n+1}^{(k)})}{x_i^{(k)}-x_j^{(k)}}\\
&\quad\longrightarrow\frac{f^{[n-1]}(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)-f^{[n-1]}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_{n+1})}{x_i-x_j}
\end{aligned}
$$

as $k\to\infty$. Hence one can define

$$
\begin{aligned}
&f^{[n]}(x_1,\dots,x_{n+1})\\
&\quad:=\lim_{k\to\infty}f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})\\
&\quad=\frac{f^{[n-1]}(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)-f^{[n-1]}(x_1,\dots,x_{i-1},x_{i+1},\dots,x_{n+1})}{x_i-x_j}
\end{aligned}
\tag{2.2.2}
$$

independently of the choice of the sequence $\{(x_1^{(k)},\dots,x_{n+1}^{(k)})\}$ and also of the choice of (i,j) with $x_i\neq x_j$. Moreover, for any sequence $\{(\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)})\}$ converging to $(x_1,\dots,x_{n+1})$ with $\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)}$ not necessarily distinct, it is seen from (2.2.2) and the continuity of $f^{[n-1]}$ that

$$f^{[n]}(x_1,\dots,x_{n+1})=\lim_{k\to\infty}f^{[n]}(\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)}).$$

Hence $f^{[n]}(x_1,\dots,x_{n+1})$ is continuous on $(a,b)^{n+1}\setminus\Delta$, where $\Delta:=\{(x,x,\dots,x):x\in(a,b)\}$.

Next assume that $x_1=\cdots=x_{n+1}=x$. Then $x_j^{(k)}\to x$ as $k\to\infty$ for all j. By (1) for distinct λ_i's proved above, for each k there is a $\xi^{(k)}$ in the smallest interval containing $x_1^{(k)},\dots,x_{n+1}^{(k)}$ such that

$$f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})=\frac{f^{(n)}(\xi^{(k)})}{n!},$$

Since $\xi^{(k)}\to x$ and f is C^n, it follows that

$$f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})\longrightarrow\frac{f^{(n)}(x)}{n!}\quad\text{as } k\to\infty.$$

Hence one can define

$$f^{[n]}(x,\dots,x):=\lim_{k\to\infty}f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})=\frac{f^{(n)}(x)}{n!}.$$

For any sequence $\{x^{(k)}\}$ with $x^{(k)}\to x$, we have

$$f^{[n]}(x^{(k)},\dots,x^{(k)})=\frac{f^{(n)}(x^{(k)})}{n!}\longrightarrow\frac{f^{(n)}(x)}{n!}=f^{[n]}(x,\dots,x).$$

Let $\{(\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)})\}$ be any sequence in $(a,b)^{n+1}\setminus\Delta$ converging to $(x,\dots,x)$. From the continuity of $f^{[n-1]}$, for each k one can choose $(x_1^{(k)},\dots,x_{n+1}^{(k)})$ such that $x_1^{(k)},\dots,x_{n+1}^{(k)}$ are distinct, $|x_j^{(k)}-\tilde{x}_j^{(k)}|<1/k$ for $1\le j\le n+1$ and $|f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})-f^{[n]}(\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)})|<1/k$. We then have

$$\lim_{k\to\infty}f^{[n]}(\tilde{x}_1^{(k)},\dots,\tilde{x}_{n+1}^{(k)})=\lim_{k\to\infty}f^{[n]}(x_1^{(k)},\dots,x_{n+1}^{(k)})=\frac{f^{(n)}(x)}{n!}$$

by (1) for distinct λ_i's again. Hence $f^{[n]}(x_1,\dots,x_{n+1})$ is continuous at $(x,\dots,x)$. Thus the statement of (1) for n is proved.

Finally we prove (2) for general λ_i's. Choose a sequence $\{(\lambda_1^{(k)},\dots,\lambda_{n+1}^{(k)})\}$ in $(a,b)^{n+1}$ converging to $(\lambda_1,\dots,\lambda_{n+1})$ with distinct $\lambda_1^{(k)},\dots,\lambda_{n+1}^{(k)}$. For each k there is a $\xi^{(k)}$ in the smallest interval containing $\lambda_1^{(k)},\dots,\lambda_{n+1}^{(k)}$ such that $f^{[n]}(\lambda_1^{(k)},\dots,\lambda_{n+1}^{(k)})=f^{(n)}(\xi^{(k)})/n!$. Let ξ be a limit point of $\{\xi^{(k)}\}$. Then it is clear that ξ is in the smallest interval containing $\lambda_1,\dots,\lambda_{n+1}$. By taking the limit thanks to (1) proved above, we have $f^{[n]}(\lambda_1,\dots,\lambda_{n+1})=f^{(n)}(\xi)/n!$. □

2.3 Fréchet derivatives of matrix functional calculus

Let f be a real-valued function on (a,b), and we denote by $\mathbb{M}_n^{sa}(a,b)$ the set of all $A\in\mathbb{M}_n^{sa}$ with $\sigma(A)\subset(a,b)$, i.e., $aI<A<bI$. It is clear that $\mathbb{M}_n^{sa}(a,b)$ is an open subset of the real Banach space $\mathbb{M}_n^{sa}$ with the Hilbert–Schmidt norm $\|\cdot\|_{\mathrm{HS}}$. In this section we discuss the differentiability property of the matrix functional calculus $A\in\mathbb{M}_n^{sa}(a,b)\mapsto f(A)\in\mathbb{M}_n^{sa}$. Let us first introduce the notion of Fréchet differentiability. The matrix functional calculus $A\mapsto f(A)$ is said to be *Fréchet differentiable* at $A_0\in\mathbb{M}_n^{sa}(a,b)$ if there exists a $Df(A_0)\in B(\mathbb{M}_n^{sa},\mathbb{M}_n^{sa})$, the space of linear maps from $\mathbb{M}_n^{sa}$ into itself, such that

$$\frac{\|f(A_0+X)-f(A_0)-Df(A_0)(X)\|_{\mathrm{HS}}}{\|X\|_{\mathrm{HS}}}\longrightarrow 0\quad\text{as } \|X\|_{\mathrm{HS}}\to 0 \text{ for } X\in\mathbb{M}_n^{sa}.$$

Then $Df(A_0)$ is called the *Fréchet derivative* of the matrix function $f(A)$ at A_0. This notion is inductively extended to the general higher degree. To do this, we denote by $B((\mathbb{M}_n^{sa})^m,\mathbb{M}_n^{sa})$ the set of all m-multilinear maps from $(\mathbb{M}_n^{sa})^m:=\mathbb{M}_n^{sa}\times\cdots\times\mathbb{M}_n^{sa}$ (m times) to $\mathbb{M}_n^{sa}$, and introduce the norm of $\Phi\in B((\mathbb{M}_n^{sa})^m,\mathbb{M}_n^{sa})$ as

$$\|\Phi\|:=\sup\left\{\frac{\|\Phi(X_1,\dots,X_m)\|_{\mathrm{HS}}}{\|X_1\|_{\mathrm{HS}}\cdots\|X_m\|_{\mathrm{HS}}}:X_1,\dots,X_m\in\mathbb{M}_n^{sa}\setminus\{0\}\right\}. \tag{2.3.1}$$

Now assume that $m\in\mathbb{N}$ with $m\ge 2$ and the $(m-1)$th Fréchet derivative $D^{m-1}f(A)$ exists for all $A\in\mathbb{M}_n^{sa}(a,b)$ in a neighborhood of $A_0\in\mathbb{M}_n^{sa}(a,b)$. We say that $f(A)$ is *m times Fréchet differentiable* at A_0 if $D^{m-1}f(A)$ is one more Fréchet differentiable at A_0, i.e., there exists a $D^mf(A_0)\in B(\mathbb{M}_n^{sa},B((\mathbb{M}_n^{sa})^{m-1},\mathbb{M}_n^{sa}))=B((\mathbb{M}_n^{sa})^m,\mathbb{M}_n^{sa})$ such that

$$\frac{\|D^{m-1}f(A_0+X)-D^{m-1}f(A_0)-D^mf(A_0)(X)\|}{\|X\|_{\mathrm{HS}}}\longrightarrow 0\quad\text{as } \|X\|_{\mathrm{HS}}\to 0 \text{ for } X\in\mathbb{M}_n^{sa},$$

with respect the norm (2.3.1) of $B((\mathbb{M}_n^{sa})^{m-1},\mathbb{M}_n^{sa})$. Then $D^mf(A_0)$ is called the *mth Fréchet derivative* of $f(A)$ at A_0. Note that the norms of $\mathbb{M}_n^{sa}$ and $B((\mathbb{M}_n^{sa})^m,\mathbb{M}_n^{sa})$ are irrelevant to the definition of Fréchet derivatives since the norms on a finite-dimensional vector space are all equivalent; we used the Hilbert–Schmidt norm just for convenience sake.

The following theorem is essentially due to Daleckii and Krein [28], where the higher derivatives of the function $t\mapsto f(A+tX)$ were obtained for self-adjoint operators in an infinite-dimensional Hilbert space while the derivatives treated in [28] are Gâteaux derivatives weaker than Fréchet derivatives. Our proof below is an extension of that in [13] for the case $m=1$. The proof is based on a useful criterion of the existence of the mth Fréchet derivative via Taylor's theorem, which is mentioned in Appendix A.1.

Theorem 2.3.1. *Let $m\in\mathbb{N}$ and assume that f is C^m on (a,b). Then the following hold true:*

(1) *$f(A)$ is m times Fréchet differentiable at every $A\in\mathbb{M}_n^{sa}(a,b)$.*

(2) *If $A=U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$ is the diagonalization of $A\in\mathbb{M}_n^{sa}(a,b)$, then the mth Fréchet derivative $D^mf(A)$ is given as*

$$D^mf(A)(X_1,\dots,X_m)=U\Bigg[\sum_{k_1,\dots,k_{m-1}=1}^{n}f^{[m]}(\lambda_i,\lambda_{k_1},\dots,\lambda_{k_{m-1}},\lambda_j)$$
$$\times\sum_{\sigma\in S_m}(U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}\Bigg]_{i,j=1}^{n}U^*$$

for all $X_1,\dots,X_m\in\mathbb{M}_n^{sa}$, where S_m is the permutations on $\{1,\dots,m\}$.

(3) *The map $A\mapsto D^mf(A)$ is a norm-continuous map from $\mathbb{M}_n^{sa}(a,b)$ to $B((\mathbb{M}_n^{sa})^m,\mathbb{M}_n^{sa})$.*

(4) *The Taylor formula holds: for every $A\in\mathbb{M}_n(a,b)$,*

$$f(A+X)=f(A)+\sum_{k=1}^{m}\frac{1}{k!}D^kf(A)(X,\dots,X)+o(\|X\|_{\mathrm{HS}}^m)\quad\textit{as } \|X\|_{\mathrm{HS}}\to 0 \textit{ for } X\in\mathbb{M}_n^{sa}.$$

(5) *For every $A\in\mathbb{M}_n^{sa}(a,b)$ and every $X_1,\dots,X_m\in\mathbb{M}_n^{sa}$,*

$$D^mf(A)(X_1,\dots,X_m)=\frac{\partial^m}{\partial t_1\cdots\partial t_m}f(A+t_1X_1+\cdots+t_mX_m)\Big|_{t_1=\cdots=t_m=0}.$$

Proof. The proof is by induction on m. To perform an induction procedure, one can take the initial case $m=0$. The statements (1)–(5) for the case $m=0$ reduce to the obvious fact that if f is continuous on (a,b) then the map $A\in\mathbb{M}_n^{sa}(a,b)\mapsto f(A)\in\mathbb{M}_n^{sa}$ is norm-continuous. In fact, the induction argument below can work also when $m=1$, under the above interpretation for the initial case.

Let $m\ge 1$ and assume that, for $r=0,1,\dots,m-1$, the statements (1)–(5) hold for the order r if f is C^r on (a,b). Upon this assumption let us prove (1)–(5) for m if f is C^m on (a,b). The proof is divided into several steps. In the following let $A\in\mathbb{M}_n^{sa}(a,b)$ with the diagonalization $A=U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$ and $X_1,\dots,X_m\in\mathbb{M}_n^{sa}$.

Step 1. When $f(x)=x^k$, it is easily verified by a direct computation that $D^mf(A)$ exists and

$$D^mf(A)(X_1,\dots,X_m)=\sum_{\substack{l_0,l_1,\dots,l_m\ge 0\\ l_0+l_1+\cdots+l_m=k-m}}\sum_{\sigma\in S_m}A^{l_0}X_{\sigma(1)}A^{l_1}X_{\sigma(2)}A^{l_2}\cdots A^{l_{m-1}}X_{\sigma(m)}A^{l_m} \tag{2.3.2}$$

(in particular, this is zero if $m>k$). The above expression is further written as

$$\sum_{\substack{l_0,l_1,\dots,l_m\ge 0\\ l_0+l_1+\cdots+l_m=k-m}} \sum_{\sigma\in S_m} U\Bigg[\sum_{k_1,\dots,k_{m-1}=1} \lambda_i^{l_0}\lambda_{k_1}^{l_1}\cdots\lambda_{k_{m-1}}^{l_{m-1}}\lambda_j^{l_m}$$

$$\times (U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}\Bigg]_{i,j=1}^{n} U^*$$

$$= U\Bigg[\sum_{k_1,\dots,k_{m-1}=1}^{n}\Bigg(\sum_{\substack{l_0,l_1,\dots,l_m\ge 0\\ l_0+l_1+\cdots+l_m=k-m}} \lambda_i^{l_0}\lambda_{k_1}^{l_1}\cdots\lambda_{k_{m-1}}^{l_{m-1}}\lambda_j^{l_m}\Bigg)$$

$$\times \sum_{\sigma\in S_m}(U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}\Bigg]_{i,j=1}^{n} U^*$$

$$= U\Bigg[\sum_{k_1,\dots,k_{m-1}=1}^{n} f^{[m]}(\lambda_i,\lambda_{k_1},\dots,\lambda_{k_{m-1}},\lambda_j)$$

$$\times \sum_{\sigma\in S_m}(U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}\Bigg]_{i,j=1}^{n} U^*$$

by Exercise 2.2.2. Hence it follows that $D^m f(A)$ exists and the expression in (2) is valid for all polynomials f. Let us extend this for all C^m functions f on (a,b) by a continuity argument.

Step 2. Let f be C^m on (a,b) and denote the right-hand side of the expression in (2) by $\mathcal{D}^m f(A)(X_1,\dots,X_m)$; then $\mathcal{D}^m f(A)$ is a symmetric (under permutation of arguments) m-multilinear map from $(\mathbb{M}_n^{sa})^m$ to $\mathbb{M}_n^{sa}$ since $f^{[m]}$ is symmetric and

$$\overline{\sum_{\sigma\in S_m}(U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}}$$
$$= \sum_{\sigma\in S_m}(U^*X_{\sigma(1)}U)_{jk_{m-1}}(U^*X_{\sigma(2)}U)_{k_{m-1}k_{m-2}}\cdots(U^*X_{\sigma(m-1)}U)_{k_2k_1}(U^*X_{\sigma(m)}U)_{k_1i}$$

for $X_1,\dots,X_m\in\mathbb{M}_n^{sa}$. By Lemma 2.2.4 (2) we have

$$|f^{[m]}(\lambda_i,\lambda_{k_1},\dots,\lambda_{k_{m-1}},\lambda_j)| \le \max_{x\in\Sigma(A)}\frac{|f^{(m)}(x)|}{m!}$$

for all $i,k_1,\dots,k_{m-1},j$, where $\Sigma(A)$ denotes the smallest interval containing the eigenvalues of A. Therefore,

$$\|\mathcal{D}^m(A)(X_1,\dots,X_m)\|_{\mathrm{HS}}$$
$$\le \max_{x\in\Sigma(A)}\frac{|f^{(m)}(x)|}{m!}\Bigg\{\sum_{i,j=1}^{n}\Bigg(\sum_{k_1,\dots,k_{m-1}=1}^{n}\sum_{\sigma\in S_m}$$
$$|(U^*X_{\sigma(1)}U)_{ik_1}(U^*X_{\sigma(2)}U)_{k_1k_2}\cdots(U^*X_{\sigma(m-1)}U)_{k_{m-2}k_{m-1}}(U^*X_{\sigma(m)}U)_{k_{m-1}j}|\Bigg)^2\Bigg\}^{1/2}$$
$$\le \max_{x\in\Sigma(A)}\frac{|f^{(m)}(x)|}{m!}\Bigg\{\sum_{i,j=1}^{n}\Bigg(\sum_{k_1,\dots,k_{m-1}=1}^{n}\sum_{\sigma\in S_m}\|X_{\sigma(1)}\|_{\mathrm{HS}}\|X_{\sigma(2)}\|_{\mathrm{HS}}\cdots\|X_{\sigma(m)}\|_{\mathrm{HS}}\Bigg)^2\Bigg\}^{1/2}$$
$$\le \max_{x\in\Sigma(A)}|f^{(m)}(x)|\cdot n^m\|X_{\sigma(1)}\|_{\mathrm{HS}}\|X_{\sigma(2)}\|_{\mathrm{HS}}\cdots\|X_{\sigma(m)}\|_{\mathrm{HS}}.$$

This implies that the norm of $\mathcal{D}^m f(A)$ on $(\mathbb{M}_n^{sa})^m$ is bounded as

$$\|\mathcal{D}^m f(A)\| \le n^m\max_{x\in\Sigma(A)}|f^{(m)}(x)|. \tag{2.3.3}$$

Step 3. For each $A\in\mathbb{M}_n^{sa}(a,b)$ and each $X\in\mathbb{M}_n^{sa}$ we write

$$\widetilde{R}_f(A,X) := f(A+X)-\sum_{k=0}^{m-1}\frac{1}{k!}D^k f(A)(X^{(k)}),$$

$$R_f(A,X) := \widetilde{R}_f(A,X)-\frac{1}{m!}\mathcal{D}^m f(A)(X^{(m)}),$$

where $X^{(k)}$ denotes k times $X,\dots,X$ and $\frac{1}{k!}D^k f(A)(X^{(k)})$ for $k=0$ means $f(A)$. Here, the existence of $D^k f(A)$ for $1\le k\le m-1$ and for $A\in\mathbb{M}_n^{sa}(a,b)$ is guaranteed by the induction hypothesis. We show that

$$\frac{\|R_f(B,X)\|_{\mathrm{HS}}}{\|X\|_{\mathrm{HS}}^m} \longrightarrow 0 \quad \text{as } (B,X) \in \mathbb{M}_n^{sa}(a,b) \times \mathbb{M}_n^{sa},\ (B,X) \to (A,0),$$

that is, for every $A \in \mathbb{M}_n^{sa}(a,b)$ and every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|R_f(A+Y,X)\|_{\mathrm{HS}} \le \varepsilon \|X\|_{\mathrm{HS}}^m$$

for all $X, Y \in \mathbb{M}_n^{sa}$ with $\|X\|_{\mathrm{HS}}, \|Y\|_{\mathrm{HS}} \le \delta$. To prove this, choose a $\delta_0 > 0$ such that

$$a_0 := \min \Sigma(A) - 2\delta_0 > a, \qquad b_0 := \max \Sigma(A) + 2\delta_0 < b.$$

Let $X, Y \in \mathbb{M}_n^{sa}$ with $\|X\|_{\mathrm{HS}}, \|Y\|_{\mathrm{HS}} \le \delta_0$. Since $\|X+Y\| \le \|X+Y\|_{\mathrm{HS}} \le 2\delta_0$, it is clear that $\Sigma(A+Y+X) \subset [a_0, b_0] \subset (a,b)$. Since f is C^m on (a,b), one can choose a sequence $\{f_i\}$ of polynomials such that $f_i^{(k)} \to f^{(k)}$ as $i \to \infty$ uniformly on $[a_0, b_0]$ for all $k = 0, 1, \ldots, m$. For each $Z \in \mathbb{M}_n^{sa}$ define a real-valued function

$$\phi_i(t) := \langle Z, f_i(A+Y+tX)\rangle_{\mathrm{HS}}, \qquad t \in [0,1].$$

Then the ϕ_i's are C^m on $[0,1]$ and

$$\phi_i^{(k)}(t) = \langle Z, D^k f_i(A+Y+tX)(X^{(k)})\rangle_{\mathrm{HS}},$$

$$\phi_i^{(m)}(t) = \langle Z, D^m f_i(A+Y+tX)(X^{(m)})\rangle_{\mathrm{HS}} = \langle Z, \mathcal{D}^m f_i(A+Y+tX)(X^{(m)})\rangle_{\mathrm{HS}},$$

since $D^m f_i = \mathcal{D}^m f_i$ for polynomials f_i as shown in Step 1. By Taylor's theorem applied to $\phi_i - \phi_j$, there exists a $\theta_0 \in (0,1)$ such that

$$\phi_i(1) - \phi_j(1) = \sum_{k=0}^{m-1} \frac{\phi_i^{(k)}(0) - \phi_j^{(k)}(0)}{k!} + \frac{\phi_i^{(m)}(\theta_0) - \phi_j^{(m)}(\theta_0)}{m!},$$

that is,

$$\langle Z, \widetilde{R}_{f_i}(A+Y,X) - \widetilde{R}_{f_j}(A+Y,X)\rangle_{\mathrm{HS}} = \frac{1}{m!}\langle Z, \mathcal{D}^m(f_i - f_j)(A+Y+\theta_0 X)(X^{(m)})\rangle_{\mathrm{HS}}.$$

Since

$$|\langle Z, \mathcal{D}^m(f_i - f_j)(A+Y+\theta_0 X)(X^{(m)})\rangle_{\mathrm{HS}}| \le \sup_{\theta \in [0,1]} \|\mathcal{D}^m(f_i - f_j)(A+Y+\theta X)\| \cdot \|X\|_{\mathrm{HS}}^m \|Z\|_{\mathrm{HS}},$$

we have

$$\|\widetilde{R}_{f_i}(A+Y,X) - \widetilde{R}_{f_j}(A+Y,X)\|_{\mathrm{HS}} \le \sup_{\theta \in [0,1]} \|\mathcal{D}^m(f_i - f_j)(A+Y+\theta X)\| \cdot \|X\|_{\mathrm{HS}}^m. \tag{2.3.4}$$

For any $\varepsilon > 0$ choose an $i_0 \in \mathbb{N}$ such that

$$\sup_{t \in [a_0,b_0]} |f_i^{(m)}(t) - f_j^{(m)}(t)| \le \frac{\varepsilon}{3n^m} \quad \text{for all } i, j \ge i_0,$$

$$\sup_{t \in [a_0,b_0]} |f_{i_0}^{(m)}(t) - f^{(m)}(t)| \le \frac{\varepsilon}{3n^m}.$$

Since $\Sigma(A+Y+\theta X) \subset [a_0, b_0]$ for all $\theta \in [0,1]$, it follows from (2.3.3) that

$$\sup_{\theta \in [0,1]} \|\mathcal{D}^m(f_i - f_j)(A+Y+\theta X)\| \le \frac{\varepsilon}{3} \quad \text{for all } i, j \ge i_0, \tag{2.3.5}$$

$$\|\mathcal{D}^m(f_{i_0} - f)(A+Y)\| \le \frac{\varepsilon}{3}. \tag{2.3.6}$$

By using (2) for $1 \le k \le m-1$ (the induction hypothesis) and Lemma 2.2.4 (2), we notice that

$$\|D^k(f_i - f)(A+Y)(X^{(k)})\|_{\mathrm{HS}} \longrightarrow 0 \quad \text{as } i \to \infty, \quad 1 \le k \le m-1$$

as well as $f_i(A+Y+X) \to f(A+Y+X)$ and $f_i(A+Y) \to f(A+Y)$ as $i \to \infty$. Therefore,

$$\|\widetilde{R}_{f_i}(A+Y,X) - \widetilde{R}_f(A+Y,X)\|_{\mathrm{HS}} \longrightarrow 0 \quad \text{as } i \to \infty.$$

So, letting $i \to \infty$ and $j = i_0$ in (2.3.4) and by applying (2.3.5) we obtain

$$\|\widetilde{R}_f(A+Y,X) - \widetilde{R}_{f_{i_0}}(A+Y,X)\|_{\mathrm{HS}} \le \frac{\varepsilon}{3}\|X\|_{\mathrm{HS}}^m. \tag{2.3.7}$$

Again by Taylor's theorem applied to ϕ_{i_0}, there exists a $\theta_1 \in (0,1)$ such that

$$\phi_{i_0}(1) = \sum_{k=0}^{m-1} \frac{\phi^{(k)}(0)}{k!} + \frac{\phi^{(m)}(\theta_1)}{m!},$$

that is,

$$\langle Z, \widetilde{R}_{f_{i_0}}(A+Y,X)\rangle_{\mathrm{HS}} = \frac{1}{m!}\langle Z, \mathcal{D}^m f_{i_0}(A+Y+\theta_1 X)(X^{(m)})\rangle_{\mathrm{HS}}.$$

Therefore,

$$\begin{aligned}|\langle Z, R_{f_{i_0}}(A+Y,X)\rangle_{\mathrm{HS}}| &= \frac{1}{m!}|\langle Z, \mathcal{D}^m f_{i_0}(A+Y+\theta_1 X)(X^{(m)}) - \mathcal{D}^m f_{i_0}(A+Y)(X^{(m)})\rangle_{\mathrm{HS}}| \\ &\le \sup_{\theta\in[0,1]} \|\mathcal{D}^m f_{i_0}(A+Y+\theta X) - \mathcal{D}^m f_{i_0}(A+Y)\| \cdot \|X\|_{\mathrm{HS}}^m \|Z\|_{\mathrm{HS}}\end{aligned}$$

for every $Z \in \mathbb{M}_n^{sa}$ and hence

$$\|R_{f_{i_0}}(A+Y,X)\|_{\mathrm{HS}} \le \sup_{\theta\in[0,1]} \|\mathcal{D}^m f_{i_0}(A+Y+\theta X) - \mathcal{D}^m f_{i_0}(A+Y)\| \cdot \|X\|_{\mathrm{HS}}^m.$$

For a polynomial f_{i_0} it is rather easy to see (left to Exercise 2.3.2 below) that the map $D^m f_{i_0} = \mathcal{D}^m f_{i_0}$ from $\mathbb{M}_n^{sa}$ into $B((\mathbb{M}_n^{sa})^m, \mathbb{M}_n^{sa})$ is norm-continuous. Hence there exists a $\delta \in (0,\delta_0)$, independently of Y with $\|Y\|_{\mathrm{HS}} \le \delta_0$, such that

$$\sup_{\theta\in[0,1]} \|\mathcal{D}^m f_{i_0}(A+Y+\theta X) - \mathcal{D}^m f_{i_0}(A+Y)\| \le \frac{\varepsilon}{3}$$

whenever $X \in \mathbb{M}_n^{sa}$ satisfies $\|X\|_{\mathrm{HS}} \le \delta$. So we have

$$\|R_{f_{i_0}}(A+Y,X)\|_{\mathrm{HS}} \le \frac{\varepsilon}{3}\|X\|_{\mathrm{HS}}^m. \tag{2.3.8}$$

Combining (2.3.6)–(2.3.8) implies that

$$\begin{aligned}\|R_f(A+Y,X)\|_{\mathrm{HS}} &\le \|\widetilde{R}_f(A+Y,X) - \widetilde{R}_{f_{i_0}}(A+Y,X)\|_{\mathrm{HS}} + \|R_{f_{i_0}}(A+Y,X)\|_{\mathrm{HS}} \\ &\qquad + \|\mathcal{D}^m f_{i_0}(A+Y)(X^{(m)}) - \mathcal{D}^m f(A+Y)(X^{(m)})\|_{\mathrm{HS}} \\ &\le \frac{\varepsilon}{3}\|X\|_{\mathrm{HS}}^m + \frac{\varepsilon}{3}\|X\|_{\mathrm{HS}}^m + \frac{\varepsilon}{3}\|X\|_{\mathrm{HS}}^m = \varepsilon\|X\|_{\mathrm{HS}}^m\end{aligned}$$

whenever $X, Y \in \mathbb{M}_n^{sa}$ satisfy $\|X\|_{\mathrm{HS}}, \|Y\|_{\mathrm{HS}} \le \delta$ ($\le \delta_0$), as required.

Step 4. Next, we show that the map $A \mapsto \mathcal{D}^m f(A)$ from $\mathbb{M}_n^{sa}(a,b)$ into $B((\mathbb{M}_n^{sa})^m, \mathbb{M}_n^{sa})$ is norm-continuous. Let $A \in \mathbb{M}_n^{sa}(a,b)$ and δ_0 be as in Step 3, and choose a sequence $\{f_i\}$ of polynomials as above. Then for any $\varepsilon > 0$ one can choose an $i_0 \in \mathbb{N}$ such that (2.3.6) holds for all $Y \in \mathbb{M}_N^{sa}$ with $\|Y\|_{\mathrm{HS}} \le \delta_0$. Moreover, by Exercise 2.3.2 below, there exists a $\delta_1 \in (0,\delta_0)$ such that $\|\mathcal{D}^m f_{i_0}(A+Y) - \mathcal{D}^m f_{i_0}(A)\| \le \varepsilon/3$ for all $Y \in \mathbb{M}_n^{sa}$ with $\|Y\|_{\mathrm{HS}} \le \delta_1$. Therefore, if $\|Y\|_{\mathrm{HS}} \le \delta_1$ then

$$\begin{aligned}&\|\mathcal{D}^m f(A+Y) - \mathcal{D}^m f(A)\| \\ &\quad \le \|\mathcal{D}^m (f - f_{i_0})(A+Y)\| + \|\mathcal{D}^m f_{i_0}(A+Y) - \mathcal{D}^m f_{i_0}(A)\| + \|\mathcal{D}^m (f_{i_0} - f)(A)\| \\ &\quad \le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,\end{aligned}$$

which yields the continuity of $A \mapsto \mathcal{D}^m f(A)$.

Step 5. We finally present the proofs of the statements (1)–(5) for m.

(1) By Steps 3 and 4 we can apply a general criterion in Lemma A.1.1 of Appendix A.1 to see that $D^m f(A)$ exists and $D^m f(A) = \mathcal{D}^m f(A)$ for every $A \in \mathbb{M}_n^{sa}(a,b)$.

(2) follows since $D^m f(A) = \mathcal{D}^m f(A)$.

(3) is contained in Step 4.

(4) is contained in Step 3; just let $Y = 0$ there.

(5) Since Fréchet differentiability implies Gâtaux (or directional) differentiability, one can differentiate $f(A + t_1X_1 + \cdots + t_mX_m)$ as

$$\begin{aligned}&\frac{\partial^m}{\partial t_1 \cdots \partial t_m} f(A + t_1X_1 + \cdots + t_mX_m)\Big|_{t_1=\cdots=t_m=0} \\ &\quad = \frac{\partial^m}{\partial t_1 \cdots \partial t_{m-1}} Df(A + t_1X_1 + \cdots + t_{m-1}X_{m-1})(X_m)\Big|_{t_1=\cdots=t_{m-1}=0} \\ &\quad = \cdots = D^m f(A)(X_1, \ldots, X_m).\end{aligned}$$

□

Theorem 2.3.1 in the case $m = 1$ says that if f is C^1 on (a, b), then $f(A)$ is Fréchet differentiable at every $A \in \mathbb{M}_n^{sa}(a, b)$ and the Fréchet derivative $Df(A)$ at A is written as

$$Df(A)(X) = U\Big([f^{[1]}(\lambda_i, \mu_j)]_{i,j=1}^n \circ (U^*XU)\Big)U^*, \tag{2.3.9}$$

where $A = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*$ is the diagonalization and $\circ$ denotes the Schur product (see Section 1.6). In particular, when $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ is in $\mathbb{M}_n^{sa}(a, b)$,

$$Df(A)(X) = [f^{[1]}(\lambda_i, \lambda_j)]_{i,j=1}^n \circ X.$$

The assertion for the case $m = 2$ says that if f is C^2 on (a, b), then the second Fréchet derivative $D^2f(A)$ at $A = \mathrm{Diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{M}_n^{sa}(a, b)$ is written as

$$D^2f(A)(X, Y) = \left[\sum_{k=1}^n f^{[2]}(\lambda_i, \lambda_k, \lambda_j)(X_{ik}Y_{kj} + Y_{ik}X_{kj})\right]_{i,j=1}^n.$$

Exercise 2.3.2. Let $m \in \mathbb{N}$ and f be a polynomial. By using the expression (2.3.2), show that the map $A \mapsto D^mf(A)$ from $\mathbb{M}_n^{sa}$ into $B((\mathbb{M}_n^{sa})^m, \mathbb{M}_n^{sa})$ is norm-continuous with respect to the norms $\|A\|_{\mathrm{HS}}$ and $\|\Phi\|$ of $\Phi \in B((\mathbb{M}_n^{sa})^m, \mathbb{M}_n^{sa})$ in (2.3.1).

Note that the Taylor formula (4) follows if the map $A \mapsto f(A)$ is C^m (i.e., (1) and (3) hold). But in our proof of Theorem 2.3.1, we first proved a stronger form of (4) as well as (3), and (1) was obtained from those two.

2.4 Characterizations of *n*-monotone and *n*-convex functions

The next theorem is due to Löwner [62] and the proof below is also based on the exposition in [30].

Theorem 2.4.1. *Assume that f is a 2-monotone function on (a, b). Then f is C^1 on (a, b), and moreover $f' > 0$ and f' is convex on (a, b) unless f is a constant.*

Proof. The proof of the theorem is divided into several steps. Assume that f is 2-monotone on (a, b) and is not a constant. Let $\xi_1 < \eta_1 < \xi_2 < \eta_2$ be arbitrary in (a, b).

Step 1. Let $A := \begin{bmatrix} \xi_1 & 0 \\ 0 & \xi_2 \end{bmatrix}$. We show that there exist $\alpha_1, \alpha_2 > 0$ such that $B := A + Q$ has the eigenvalues η_1, η_2 if we set $Q := \begin{bmatrix} \alpha_1 & \sqrt{\alpha_1\alpha_2} \\ \sqrt{\alpha_1\alpha_2} & \alpha_2 \end{bmatrix}$. In fact, since

$$\det(\lambda I - B) = \lambda^2 - (\xi_1 + \xi_2 + \alpha_1 + \alpha_2)\lambda + \xi_1\xi_2 + \xi_1\alpha_2 + \xi_2\alpha_1,$$

the required condition is

$$\begin{cases} \xi_1 + \xi_2 + \alpha_1 + \alpha_2 = \eta_1 + \eta_2, \\ \xi_1\xi_2 + \xi_1\alpha_2 + \xi_2\alpha_1 = \eta_1\eta_2. \end{cases}$$

This can be explicitly solved as

$$\alpha_1 = \frac{(\eta_1 - \xi_1)(\eta_2 - \xi_1)}{\xi_2 - \xi_1}, \qquad \alpha_2 = \frac{(\xi_2 - \eta_1)(\eta_2 - \xi_2)}{\xi_2 - \xi_1},$$

which are positive numbers.

Step 2. We prove that

$$\det\begin{bmatrix} f^{[1]}(\xi_1, \eta_1) & f^{[1]}(\xi_1, \eta_2) \\ f^{[1]}(\xi_2, \eta_1) & f^{[2]}(\xi_2, \eta_2) \end{bmatrix} \ge 0. \tag{2.4.1}$$

Choose $\alpha_1, \alpha_2 > 0$ as in Step 1. Since $Q \ge 0$ so that $A \le B$, we have $f(A) \le f(B)$. There is a unitary $U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$ such that $B = U\widetilde{B}U^*$ with $\widetilde{B} := \begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix}$. Note that $U\widetilde{B} - AU = QU$. Since

$$U\widetilde{B} - AU = \begin{bmatrix} (\eta_1 - \xi_1)u_{11} & (\eta_2 - \xi_1)u_{12} \\ (\eta_1 - \xi_2)u_{21} & (\eta_2 - \xi_2)u_{22} \end{bmatrix}$$

and

$$QU = \begin{bmatrix} \alpha_1 u_{11} + \sqrt{\alpha_1\alpha_2}u_{21} & \alpha_1 u_{12} + \sqrt{\alpha_1\alpha_2}u_{22} \\ \sqrt{\alpha_1\alpha_2}u_{11} + \alpha_2 u_{21} & \sqrt{\alpha_1\alpha_2}u_{12} + \alpha_2 u_{22} \end{bmatrix},$$

it follows that

$$\begin{cases} (\eta_1 - \xi_1)u_{11} = \alpha_1 u_{11} + \sqrt{\alpha_1\alpha_2}u_{21}, & (\eta_2 - \xi_1)u_{12} = \alpha_1 u_{12} + \sqrt{\alpha_1\alpha_2}u_{22}, \\ (\eta_1 - \xi_2)u_{21} = \sqrt{\alpha_1\alpha_2}u_{11} + \alpha_2 u_{21}, & (\eta_2 - \xi_2)u_{22} = \sqrt{\alpha_1\alpha_2}u_{12} + \alpha_2 u_{22}. \end{cases} \tag{2.4.2}$$

On the other hand,

$$\begin{aligned} &Uf(\widetilde{B}) - f(A)U \\ &= U\begin{bmatrix} f(\eta_1) & 0 \\ 0 & f(\eta_2) \end{bmatrix} - \begin{bmatrix} f(\xi_1) & 0 \\ 0 & f(\xi_2) \end{bmatrix} U \\ &= \begin{bmatrix} (f(\eta_1) - f(\xi_1))u_{11} & (f(\eta_2) - f(\xi_1))u_{12} \\ (f(\eta_1) - f(\xi_2))u_{21} & (f(\eta_2) - f(\xi_2))u_{22} \end{bmatrix} \\ &= \begin{bmatrix} f^{[1]}(\xi_1,\eta_1)\sqrt{\alpha_1}(\sqrt{\alpha_1}u_{11} + \sqrt{\alpha_2}u_{21}) & f^{[1]}(\xi_1,\eta_2)\sqrt{\alpha_1}(\sqrt{\alpha_1}u_{12} + \sqrt{\alpha_2}u_{22}) \\ f^{[1]}(\xi_2,\eta_1)\sqrt{\alpha_2}(\sqrt{\alpha_1}u_{11} + \sqrt{\alpha_2}u_{21}) & f^{[1]}(\xi_2,\eta_2)\sqrt{\alpha_2}(\sqrt{\alpha_1}u_{12} + \sqrt{\alpha_2}u_{22}) \end{bmatrix} \end{aligned}$$

thanks to (2.4.2). Hence

$$\begin{aligned} &\det(Uf(\widetilde{B}) - f(A)U) \\ &= \det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[1]}(\xi_1,\eta_2) \\ f^{[1]}(\xi_2,\eta_1) & f^{[2]}(\xi_2,\eta_2) \end{bmatrix}\sqrt{\alpha_1\alpha_2}(\sqrt{\alpha_1}u_{11} + \sqrt{\alpha_2}u_{21})(\sqrt{\alpha_1}u_{12} + \sqrt{\alpha_2}u_{22}). \end{aligned}$$

Thanks to (2.4.2) again we have

$$\begin{aligned} \det U &= \det\begin{bmatrix} \frac{\sqrt{\alpha_1}(\sqrt{\alpha_1}u_{11}+\sqrt{\alpha_2}u_{21})}{\eta_1-\xi_1} & \frac{\sqrt{\alpha_1}(\sqrt{\alpha_1}u_{12}+\sqrt{\alpha_2}u_{22})}{\eta_2-\xi_1} \\ \frac{\sqrt{\alpha_2}(\sqrt{\alpha_1}u_{11}+\sqrt{\alpha_2}u_{21})}{\eta_1-\xi_2} & \frac{\sqrt{\alpha_2}(\sqrt{\alpha_1}u_{12}+\sqrt{\alpha_2}u_{22})}{\eta_2-\xi_2} \end{bmatrix} \\ &= \det\begin{bmatrix} \frac{1}{\eta_1-\xi_1} & \frac{1}{\eta_2-\xi_1} \\ \frac{1}{\eta_1-\xi_2} & \frac{1}{\eta_2-\xi_2} \end{bmatrix}\sqrt{\alpha_1\alpha_2}(\sqrt{\alpha_1}u_{11} + \sqrt{\alpha_2}u_{21})(\sqrt{\alpha_1}u_{12} + \sqrt{\alpha_2}u_{22}). \end{aligned}$$

Noting that

$$\gamma := \det\begin{bmatrix} \frac{1}{\eta_1-\xi_1} & \frac{1}{\eta_2-\xi_1} \\ \frac{1}{\eta_1-\xi_2} & \frac{1}{\eta_2-\xi_2} \end{bmatrix} > 0,$$

we have

$$\begin{aligned} \det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[1]}(\xi_1,\eta_2) \\ f^{[1]}(\xi_2,\eta_1) & f^{[1]}(\xi_2,\eta_2) \end{bmatrix} &= \det(Uf(\widetilde{B}) - f(A)U)\cdot\gamma(\det U)^{-1} \\ &= \gamma\det(Uf(\widetilde{B}) - f(A)U)\det U^* \\ &= \gamma\det(Uf(\widetilde{B})U^* - f(A)) \\ &= \gamma\det(f(B) - f(A)) \ge 0. \end{aligned}$$

Step 3. We show that f is absolutely continuous on any closed subinterval of (a,b). To prove this, it suffices to show that f is Lipschitz continuous on any closed interval $[c,d] \subset (a,b)$. Fix $\xi_1 \in (a,c)$ and $\eta_2 \in (d,b)$. For every $\eta_1 < \xi_2$ in $[c,d]$, it follows from (2.4.1) that

$$\frac{f(\eta_2) - f(\xi_1)}{\eta_2 - \xi_1}\cdot\frac{f(\xi_2) - f(\eta_1)}{\xi_2 - \eta_1} \le \frac{f(\eta_1) - f(\xi_1)}{\eta_1 - \xi_1}\cdot\frac{f(\eta_2) - f(\xi_2)}{\eta_2 - \xi_2} \le \frac{(f(\eta_2) - f(\xi_1))^2}{(c - \xi_1)(\eta_2 - d)},$$

since $f(\xi_1) \le f(\eta_1) \le f(\xi_2) \le f(\eta_2)$. We may assume that $f(\eta_2) - f(\xi_1) > 0$; otherwise, f is obviously a constant on $[\xi_1,\eta_2] \supset [c,d]$. We then have the Lipschitz condition

$$f(\xi_2) - f(\eta_1) \le M(\xi_2 - \eta_1), \qquad \eta_1,\xi_2 \in [c,d],\ \eta_1 < \xi_2,$$

where $M := (\eta_2 - \xi_1)(f(\eta_2) - f(\xi_1))/(c - \xi_1)(\eta_2 - d)$.

Step 4. We prove that

$$\det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[2]}(\xi_1,\eta_1,\eta_2) \\ f^{[2]}(\xi_1,\xi_2,\eta_1) & f^{[3]}(\xi_1,\xi_2,\eta_1,\eta_2) \end{bmatrix} \ge 0. \tag{2.4.3}$$

In fact, the above determinant is equal to

$$\begin{aligned}
&\det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[2]}(\xi_1,\eta_1,\eta_2) \\ \frac{f^{[1]}(\xi_2,\eta_1)-f^{[1]}(\xi_1,\eta_1)}{\xi_2-\xi_1} & \frac{f^{[2]}(\xi_2,\eta_1,\eta_2)-f^{[2]}(\xi_1,\eta_1,\eta_2)}{\xi_2-\xi_1} \end{bmatrix} \\
&= \frac{1}{\xi_2-\xi_1}\det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[2]}(\xi_1,\eta_1,\eta_2) \\ f^{[1]}(\xi_2,\eta_1) & f^{[2]}(\xi_2,\eta_1,\eta_2) \end{bmatrix} \\
&= \frac{1}{\xi_2-\xi_1}\det\begin{bmatrix} f^{[1]}(\xi_1,\xi_2) & \frac{f^{[1]}(\xi_1,\eta_2)-f^{[1]}(\xi_1,\eta_1)}{\eta_2-\eta_1} \\ f^{[1]}(\xi_2,\eta_1) & \frac{f^{[1]}(\xi_2,\eta_2)-f^{[1]}(\xi_2,\eta_1)}{\eta_2-\eta_1} \end{bmatrix} \\
&= \frac{1}{(\xi_2-\xi_1)(\eta_2-\eta_1)}\det\begin{bmatrix} f^{[1]}(\xi_1,\eta_1) & f^{[1]}(\xi_1,\eta_2) \\ f^{[1]}(\xi_2,\eta_1) & f^{[1]}(\xi_2,\eta_2) \end{bmatrix} \ge 0
\end{aligned}$$

thanks to (2.4.1).

Step 5. Assume that f is smooth (or at least C^3) on (a,b). We then prove that

$$\det\begin{bmatrix} f'(x) & \frac{f''(x)}{2!} \\ \frac{f''(x)}{2!} & \frac{f^{(3)}(x)}{3!} \end{bmatrix} \ge 0 \quad \text{for all } x \in (a,b), \tag{2.4.4}$$

and moreover, $f'(x) > 0$ and $f^{(3)}(x) \ge 0$ for all $x \in (a,b)$. In fact, (2.4.4) immediately follows from (2.4.3) by letting $\xi_j, \eta_j \to x$, $j = 1,2$. Also, letting $\eta_j \to \xi_j$ in (2.4.1) yields

$$\det\begin{bmatrix} f'(\xi_1) & f^{[1]}(\xi_1,\xi_2) \\ f^{[1]}(\xi_1,\xi_2) & f'(\xi_2) \end{bmatrix} \ge 0 \quad \text{for all } \xi_1,\xi_2 \in (a,b).$$

If $f'(\xi_1) = 0$ for some $\xi_1 \in (a,b)$, then we must have $f^{[1]}(\xi_1,\xi_2) = 0$ so that $f(\xi_2) = f(\xi_1)$ for all $\xi_2 \in (a,b)$, contradicting the assumption that f is not a constant on (a,b). Hence $f'(x) > 0$ for all $x \in (a,b)$ and so $f^{(3)}(x) \ge 0$ for all $x \in (a,b)$ thanks to (2.4.4).

Step 6. To finish the proof of the theorem, let us employ the regularization technique described in Appendix A.2. For any $\varepsilon > 0$ small enough, let f_ε be the regularization of f defined in (A.2.1). If $A, B \in \mathbb{M}_2^{sa}(a+\varepsilon, b-\varepsilon)$ and $A \le B$, then

$$f_\varepsilon(A) = \int_{-1}^{1} \varphi(t) f(A - \varepsilon t I)\,dt \le \int_{-1}^{1} \varphi(t) f(B - \varepsilon t I)\,dt = f_\varepsilon(B).$$

Hence f_ε is 2-monotone on $(a+\varepsilon, b-\varepsilon)$, so it follows from Step 5 that f_ε' is nonnegative and convex on $(a+\varepsilon, b-\varepsilon)$. Lemma A.2.1 (4) and Step 3 imply that $f_\varepsilon'(x)$ converges as $\varepsilon \searrow 0$ to $f'(x)$ almost everywhere on any closed interval $[c,d] \subset (a,b)$. Choose $\delta > 0$ with $2\delta < \min\{c-a, b-d\}$. For any $\varepsilon \in (0,\delta]$ and $x, y \in [c,d]$, since $a+\varepsilon < c-\delta$ and $d+\delta < b-\varepsilon$, we have

$$\frac{f_\varepsilon'(c) - f_\varepsilon'(c-\delta)}{\delta} \le \frac{f_\varepsilon'(x) - f_\varepsilon'(y)}{x-y} \le \frac{f_\varepsilon'(d+\delta) - f_\varepsilon'(d)}{\delta}$$

thanks to the convexity of f_ε' on $(a+\varepsilon, b-\varepsilon)$. Hence

$$|f_\varepsilon'(x) - f_\varepsilon'(y)| \le K_\varepsilon |x-y|, \qquad x, y \in [c,d],$$

where

$$K_\varepsilon := \max\left\{\left|\frac{f_\varepsilon'(c) - f_\varepsilon'(c-\delta)}{\delta}\right|, \left|\frac{f_\varepsilon'(d+\delta) - f_\varepsilon'(d)}{\delta}\right|\right\}.$$

Here, changing c, d and δ arbitrarily small, we may assume that $f_\varepsilon'(x)$ converges as $\varepsilon \searrow 0$ to $f'(x)$ at the four points $c-\delta$, c, d, and $d+\delta$. Then it follows that $K := \sup\{K_\varepsilon : \varepsilon \in (0,\delta]\} < +\infty$. Hence $\{f_\varepsilon' : \varepsilon \in (0,\delta]\}$ is equicontinuous on $[c,d]$. From this we see that f_ε' uniformly converges as $\varepsilon \searrow 0$ to a continuous function g on $[c,d]$ so that $f'(x) = g(x)$ almost everywhere on $[c,d]$. But in this case, we obtain

$$f(x) = f(c) + \int_c^x g(t)\,dt, \qquad x \in [c,d],$$

which implies that $f'(x) = g(x) = \lim_{\varepsilon\searrow 0} f'_\varepsilon(x)$ for all $x \in [c,d]$. Hence f is C^1 and f' is nonnegative and convex on (a,b). From the argument in Step 5 it also follows that $f' > 0$ on (a,b) if f is not a constant. □

The next theorem concerned with 2-convex functions is due to Kraus [57], whose proof is based on the essentially same idea as that of the preceding proof but much more involved according to the stronger conclusion of C^2. We prefer to transfer the proof into Appendix A.3.

Theorem 2.4.2. *Assume that f is a 2-convex function on (a,b). Then f is C^2 on (a,b), and moreover $f^{[2]}(x,y,z) > 0$ for all distinct x,y,z in (a,b) unless f is linear on (a,b).*

In fact, the theorem holds under a weaker assumption that f is conditionally 2-convex on (a,b) in the sense that $f((1-\lambda)A + \lambda B) \le (1-\lambda)f(A) + \lambda f(B)$ for all $\lambda \in (0,1)$ and for all $A,B \in \mathbb{M}_2^{sa}$ with $\sigma(A),\sigma(B) \subset (a,b)$ such that $A \le B$.

In the rest of this section we present the characterizations of n-monotone functions (due to Löwner [62]) and of n-convex functions (due to Kraus [57]) in terms of divided difference matrices. Key roles in our proofs are played by the formulas in Theorem 2.3.1 (2) for the first and the second Fréchet derivatives of matrix functional calculus.

Theorem 2.4.3. *Let $n \in \mathbb{N}$ with $n \ge 2$ and let f be a real-valued function on (a,b). Then the following are equivalent:*
(i) *f is n-monotone on (a,b);*
(ii) *f is C^1 on (a,b) and $[f^{[1]}(\lambda_i,\lambda_j)]_{i,j=1}^n \ge 0$ for every choice of $\lambda_1 < \cdots < \lambda_n$ from (a,b).*

Proof. (i) ⇒ (ii). Since f is at least 2-monotone, the C^1 of f was proved in Theorem 2.4.1. For every choice of $\lambda_1 < \cdots < \lambda_n$ from (a,b) and for any $\xi_1,\dots,\xi_n \in \mathbb{C}$, define $A := \mathrm{Diag}(\lambda_1,\dots,\lambda_n)$ and $X := [\overline{\xi_i}\xi_j]_{i,j=1}^n$ (≥ 0). Since $A + \varepsilon X \in \mathbb{M}_n^{sa}(a,b)$ and $A + \varepsilon X \ge A$ for all $\varepsilon > 0$ small enough, we have $f(A+\varepsilon X) \ge f(A)$ for such an $\varepsilon > 0$. Since f is C^1 on (a,b) as mentioned above, we can apply the formula in Theorem 2.3.1 (2) for $m = 1$ to obtain

$$[f^{[1]}(\lambda_i,\lambda_j)\overline{\xi_i}\xi_j]_{i,j=1}^n = [f^{[1]}(\lambda_i,\lambda_j)] \circ X = Df(A)(X) = \lim_{\varepsilon\searrow 0} \frac{f(A+\varepsilon X) - f(A)}{\varepsilon} \ge 0,$$

which implies that $\sum_{i,j=1}^n f^{[1]}(\lambda_i,\lambda_j)\overline{\xi_i}\xi_j \ge 0$. Hence we have $[f^{[1]}(\lambda_i,\lambda_j)] \ge 0$.

(ii) ⇒ (i). For each $A,B \in \mathbb{M}_n^{sa}(a,b)$ with $A \le B$, define $A_t := (1-t)A + tB$ for $t \in [0,1]$. Note that $A_t \in \mathbb{M}_n^{sa}(a,b)$ and $A_t \le A_{t'}$ for all $t,t' \in [0,1]$ with $t \le t'$. For every $t \in [0,1]$ we have

$$\frac{d}{dt} f(A_t) = Df(A_t)(B-A) = U\Big([f^{[1]}(\lambda_i,\lambda_j)]_{i,j=1}^n \circ U^*(B-A)U\Big)U^*$$

thanks to Theorem 2.3.1 (2) for $m = 1$, where we take the diagonalization $A_t = U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$ (of course, depending on t). Hence (ii) and the Schur product theorem (Theorem 1.6.3) imply that $\frac{d}{dt}f(A_t) \ge 0$ for all $t \in [0,1]$ so that

$$f(B) - f(A) = \int_0^1 \frac{d}{dt} f(A_t)\,dt \ge 0. \qquad \square$$

Theorem 2.4.4. *Let $n \in \mathbb{N}$ with $n \ge 2$ and let f be a real-valued function on (a,b). Then the following are equivalent:*
(i) *f is n-convex on (a,b);*
(ii) *f is conditionally n-convex on (a,b), i.e., (2.1.1) holds for every $A,B \in \mathbb{M}_n^{sa}(a,b)$ such that $A \le B$;*
(iii) *f is C^2 on (a,b) and $[f^{[2]}(\lambda_1,\lambda_i,\lambda_j)]_{i,j=1}^n \ge 0$ for any choice of $\lambda_1,\dots,\lambda_n$ from (a,b).*

Proof. (i) ⇒ (ii) is obvious.

(ii) ⇒ (iii). Since (ii) implies the same condition for $n = 2$, the C^2 of f was given in Theorem 2.4.2 while the proof is in Appendix A.3. For any $\lambda_1,\dots,\lambda_n \in (a,b)$ and $\xi_1,\dots,\xi_n \in \mathbb{C}$, define $A := \mathrm{Diag}(\lambda_1,\dots,\lambda_n)$ and $X := [\overline{\xi_i}\xi_j]_{i,j=1}^n$. There is a $\delta > 0$ such that $A + tX \in \mathbb{M}_n^{sa}(a,b)$ for all $t \in (-\delta,\delta)$. For every $s,t \in (-\delta,\delta)$ with $s < t$, since $(A+tX) - (A+sX) = (t-s)X$ is positive semidefinite and of rank 1, (ii) implies that

$$f\Big(A + \frac{s+t}{2}X\Big) = f\Big(\frac{(A+sX)+(A+tX)}{2}\Big) \le \frac{f(A+sX)+f(A+tX)}{2},$$

which implies that $t \in (-\delta,\delta) \mapsto \omega(f(A+tX))$ is a convex function for each state ω on $\mathbb{M}_n$. Thanks to Theorem 2.3.1 (5) for $m = 2$ we have

$$\omega(D^2f(A)(X,X)) = \frac{d^2}{dt^2}\,\omega(f(A+tX))\Big|_{t=0} \ge 0$$

so that $D^2f(A)(X,X) \ge 0$ by Exercise 1.5.4 (3). By Theorem 2.3.1 (2) for $m = 2$ this means that

$$\left[\sum_{k=1}^n f^{[2]}(\lambda_i,\lambda_k,\lambda_j)\overline{\xi_i}|\xi_k|^2\xi_j\right]_{i,j=1}^n \ge 0,$$

that is,

$$\sum_{i,j=1}^{n}\sum_{k=1}^{n} f^{[2]}(\lambda_i,\lambda_k,\lambda_j)\overline{\xi_i}|\xi_k|^2\xi_j\overline{\zeta_i}\zeta_j \ge 0, \qquad \zeta_1,\dots,\zeta_n \in \mathbb{C}.$$

Replacing ζ_i by ζ_i/ξ_i under the assumption that $\xi_i \neq 0$ for all i, we have

$$\sum_{i,j=1}^{n}\sum_{k=1}^{n} f^{[2]}(\lambda_i,\lambda_k,\lambda_j)|\xi_k|^2\overline{\zeta_i}\zeta_j \ge 0, \qquad \zeta_1,\dots,\zeta_n \in \mathbb{C}.$$

Letting $\xi_1 = 1$ and $\xi_k \to 0$ for $k > 1$ gives

$$\sum_{i,j=1}^{n} f^{[2]}(\lambda_i,\lambda_1,\lambda_j)\overline{\zeta_i}\zeta_j \ge 0, \qquad \zeta_1,\dots,\zeta_n \in \mathbb{C}.$$

(iii) $\Rightarrow$ (i). Let $A,B \in \mathbb{M}_n^{sa}(a,b)$. Define $A_t := (1-t)A + tB$ for $t \in [0,1]$ and $X := B - A$. By (iii) and Theorem 2.3.1 (2) for $m = 2$ we have

$$\frac{d^2}{dt^2} f(A_t) = D^2 f(A_t)(X,X) = U\left[2\sum_{k=1}^{n} f^{[2]}(\lambda_i,\lambda_k,\lambda_j)\overline{(U^*XU)_{ki}}(U^*XU)_{kj}\right]_{i,j=1}^{n} U^*,$$

where $A_t = U\mathrm{Diag}(\lambda_1,\dots,\lambda_n)U^*$ is the diagonalization of A_t (depending on t). It follows from (iii) that $[f^{[2]}(\lambda_i,\lambda_k,\lambda_j)]_{i,j=1}^{n} \ge 0$ for all $k = 1,\dots,n$. Therefore, $\frac{d^2}{dt^2} f(A_t) \ge 0$ for all $t \in [0,1]$. Hence for every state ω on $\mathbb{M}_n$, $t \in [0,1] \mapsto \omega(f(A_t))$ is a convex function so that

$$\omega(f(A_t)) \le (1-t)\omega(f(A_0)) + t\omega(f(A_1)) = \omega((1-t)f(A) + tf(B)),$$

implying that $f(A_t) \le (1-t)f(A) + tf(B)$ for $0 \le t \le 1$ by Exercise 1.5.4 (3). □

Remark 2.4.5. It is worth emphasizing that the n-convexity of a function f on (a,b) was originally defined in [57] in the conditional sense (as stated in Theorem 2.4.2), and its equivalence with the unconditional sense (Definition 2.1.3 (2)) was shown in [57]. However, the definition in all the literatures later on is given in the unconditional sense. It is also seen from the above proof that when the C^2 of f is known, the conditional 2-convexity can be reduced to that under the condition that $A \le B$ and $B - A$ is of rank 1.

Combining Theorems 2.4.3 and 2.4.4 we have:

Corollary 2.4.6. *Let f be a real-valued function on (a,b) and $n \ge 2$. If f is n-convex, then $f^{[1]}(\lambda,\cdot)$ is $(n-1)$-monotone for every $\lambda \in (a,b)$. If $f^{[1]}(\lambda,\cdot)$ is n-monotone for every $\lambda \in (a,b)$, then f is n-convex. Hence, f is operator convex if and only if f is C^1 on (a,b) and $f^{[1]}(\lambda,\cdot)$ is operator monotone for every $\lambda \in (a,b)$. (The last result will be improved in Corollary 2.7.8.)*

Proof. If f is n-convex, then Theorem 2.4.4 implies that f is C^2 and $[f^{[2]}(\lambda_1,\lambda_i,\lambda_j)]_{i,j=2}^{n} \ge 0$ for every $\lambda_1,\dots,\lambda_n \in (a,b)$. Hence $f^{[1]}(\lambda_1,\cdot)$ is $(n-1)$-monotone by Theorem 2.4.3. Conversely, assume that $f^{[1]}(\lambda,\cdot)$ is n-monotone for every $\lambda \in (a,b)$. Since $f^{[1]}(\lambda,\cdot)$ is C^1 for every λ by Theorem 2.4.3, it follows that f itself is C^1. For any $\varepsilon > 0$ small enough, let f_ε be the regularization of f defined in (A.2.1) of Appendix A.2. For every $\lambda, x \in (a+\varepsilon, b-\varepsilon)$ we notice that

$$\begin{aligned} f_\varepsilon^{[1]}(\lambda,x) &= \frac{f_\varepsilon(\lambda) - f_\varepsilon(x)}{\lambda - x} = \int_{-1}^{1} \varphi(t)\frac{f(\lambda-\varepsilon t) - f(x-\varepsilon t)}{(\lambda-\varepsilon t) - (x-\varepsilon t)}\,dt \\ &= \int_{-1} \varphi(t) f^{[1]}(\lambda-\varepsilon t, x-\varepsilon t)\,dt. \end{aligned}$$

Note that the above expression is valid for the case $\lambda = x$ as well since $f_\varepsilon' = (f')_\varepsilon$ by Lemma A.2.1 (4). If $A \ge B$ in $\mathbb{M}_n^{sa}(a+\varepsilon, b-\varepsilon)$, then

$$f_\varepsilon^{[1]}(\lambda,A) = \int_{-1}^{1} \varphi(t) f^{[1]}(\lambda-\varepsilon t, A-\varepsilon tI)\,dt \ge \int_{-1}^{1} \varphi(t) f^{[1]}(\lambda-\varepsilon t, B-\varepsilon tI)\,dt = f_\varepsilon^{[1]}(\lambda,B).$$

Hence $f_\varepsilon^{[1]}(\lambda,\cdot)$ is n-monotone on $(a+\varepsilon,b-\varepsilon)$ for every $\lambda \in (a+\varepsilon,b-\varepsilon)$. Theorem 2.4.3 implies that $[f_\varepsilon^{[2]}(\lambda_1,\lambda_i,\lambda_j)]_{i,j=1}^{n} \ge 0$ for every $\lambda_1,\dots,\lambda_n \in (a+\varepsilon,b-\varepsilon)$. Hence f_ε is n-convex on $(a+\varepsilon,b-\varepsilon)$ by Theorem 2.4.4. Since $f(x) = \lim_{\varepsilon\searrow 0} f_\varepsilon(x)$ for all $x \in (a,b)$, f is n-convex on (a,b). □

Löwner's original proof of Theorem 2.4.3 is rather algebraic. The idea of the above proofs of Theorems 2.4.3 and 2.4.4 using the Taylor formula is due to Daleckii, which was briefly but clearly explained in a survey paper of Davis [29]. Theorem 2.4.3 is actually the first step of Löwner's theorem, stated in the following without proof. The proof is done by induction on n and by taking account of the larger degree versions of the determinants (2.4.3) and (2.4.4). Details are found in [30].

Theorem 2.4.7. *Let f be an n-monotone function on (a,b), where $n \ge 2$. Then f is C^{2n-3} on (a,b) and $f^{(2n-3)}$ is convex on (a,b).*

2.5 Hansen and Pedersen's characterization

The aim of this section is to characterize operator convex and monotone functions based on Hansen and Pedersen's method in [35] using 2×2 block matrices. See the first part of Section 1.7 for the basics on 2×2 block matrices.

The following simple lemma is needed in the proof below.

Lemma 2.5.1. (1) *Assume that $A \in B(\mathcal{H})$ is normal and $U \in B(\mathcal{H})$ is a unitary. Then for every function f on $\sigma(A)$, $f(U^*AU) = U^*f(A)U$.*

(2) *For every $X \in B(\mathcal{H})$ and every function f on $\sigma(X^*X)$, $Xf(X^*X) = f(XX^*)X$.*

Proof. (1) Take the spectral decomposition $A = \sum_{j=1}^m \alpha_j P_j$ as in (1.4.2). Since $U^*AU = \sum_{j=1}^m \alpha_j U^*P_jU$ is the spectral decomposition of U^*AU,

$$f(U^*AU) = \sum_{j=1}^m f(\alpha_j)U^*P_jU = U^*f(A)U.$$

(2) Since $\sigma(X^*X) = \sigma(XX^*)$ by Proposition 1.5.7 (1), $f(XX^*)$ is defined as well as $f(X^*X)$. Since $X(X^*X)^k = (XX^*)^kX$ for $k \in \mathbb{N}$, the assertion holds if f is a polynomial. Let f be an arbitrary function on $\sigma(X^*X) = \{\alpha_1, \dots, \alpha_m\}$. Define the so-called *Lagrange interpolation polynomial*

$$p(t) := \sum_{j=1}^m f(\alpha_j) \prod_{1 \le i \le m,\, i \ne j} \frac{t - \alpha_i}{\alpha_j - \alpha_i},$$

which is a polynomial such that $p(\alpha_j) = f(\alpha_j)$ for $1 \le j \le m$. Hence we have

$$Xf(X^*X) = Xp(X^*X) = p(XX^*)X = f(XX^*)X. \qquad \square$$

Theorem 2.5.2. *Let $0 < \alpha \le \infty$ and f be a real-valued function on $[0,\alpha)$. Then the following conditions are equivalent, where $\mathcal{H}$ is arbitrary and not fixed:*

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

(ii) *f is operator convex on $(0,\alpha)$ and $f(+0) \le f(0) \le 0$, where the existence of $f(+0) := \lim_{t \searrow 0} f(t)$ and $f(+0) \le f(0)$ are automatic from the operator convexity of f on $(0,\alpha)$;*

(iii) *$f(t)/t$ is operator monotone on $(0,\alpha)$ and $f(+0) \le f(0) \le 0$, where the existence of $f(+0)$ and $f(+0) \le 0$ are automatic from the the operator monotonicity of $f(t)/t$ on $(0,\infty)$;*

(iv) *$f(X^*AX) \le X^*f(A)X$ for every $A \in B(\mathcal{H})^{sa}$ with $\sigma(A) \subset [0,\alpha)$ and every $X \in B(\mathcal{H})$ with $\|X\| \le 1$;*

(v) *$f(X^*AX + Y^*BY) \le X^*f(A)X + Y^*f(B)Y$ for every $A,B \in B(\mathcal{H})^{sa}$ with $\sigma(A), \sigma(B) \subset [0,\alpha)$ and every $X,Y \in B(\mathcal{H})$ with $X^*X + Y^*Y \le I$;*

(vi) *$f(PAP) \le Pf(A)P$ for every $A \in B(\mathcal{H})^{sa}$ with $\sigma(A) \subset [0,\alpha)$ and every orthogonal projection $P \in B(\mathcal{H})$.*

Proof. (i) $\Rightarrow$ (iv). For A, X as in (iv) define $\mathbf{A}, \mathbf{U}, \mathbf{V} \in B(\mathcal{H} \oplus \mathcal{H})$ by

$$\mathbf{A} := \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{U} := \begin{bmatrix} X & (I - XX^*)^{1/2} \\ (I - X^*X)^{1/2} & -X^* \end{bmatrix},$$
$$\mathbf{V} := \begin{bmatrix} X & -(I - XX^*)^{1/2} \\ (I - X^*X)^{1/2} & X^* \end{bmatrix}.$$

Since $X(I - X^*X)^{1/2} = (I - XX^*)^{1/2}X$ by Lemma 2.5.1 (2), $\mathbf{U}^*\mathbf{U} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ so that $\mathbf{U}$ is a unitary and similarly for $\mathbf{V}$.

Compute

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} X^*AX & X^*A(I - XX^*)^{1/2} \\ (I - XX^*)^{1/2}AX & (I - XX^*)^{1/2}A(I - XX^*)^{1/2} \end{bmatrix},$$
$$\mathbf{V}^*\mathbf{A}\mathbf{V} = \begin{bmatrix} X^*AX & -X^*A(I - XX^*)^{1/2} \\ -(I - XX^*)^{1/2}AX & (I - XX^*)^{1/2}A(I - XX^*)^{1/2} \end{bmatrix}.$$

Hence (i) together with Lemma 2.5.1 (1) implies that

$$
\begin{aligned}
&\begin{bmatrix} f(X^*AX) & 0 \\ 0 & f((I-XX^*)^{1/2}A(I-XX^*)^{1/2}) \end{bmatrix} \\
&= f\left(\begin{bmatrix} X^*AX & 0 \\ 0 & (I-XX^*)^{1/2}A(I-XX^*)^{1/2} \end{bmatrix}\right) \\
&= f\left(\frac{\mathbf{U}^*\mathbf{A}\mathbf{U}+\mathbf{V}^*\mathbf{A}\mathbf{V}}{2}\right) \\
&\le \frac{f(\mathbf{U}^*\mathbf{A}\mathbf{U})+f(\mathbf{V}^*\mathbf{A}\mathbf{V})}{2} \\
&= \frac{1}{2}\mathbf{U}^*\begin{bmatrix} f(A) & 0 \\ 0 & f(0)I \end{bmatrix}\mathbf{U} + \frac{1}{2}\mathbf{V}^*\begin{bmatrix} f(A) & 0 \\ 0 & f(0)I \end{bmatrix}\mathbf{V} \\
&\le \frac{1}{2}\mathbf{U}^*\begin{bmatrix} f(A) & 0 \\ 0 & 0 \end{bmatrix}\mathbf{U} + \frac{1}{2}\mathbf{V}^*\begin{bmatrix} f(A) & 0 \\ 0 & 0 \end{bmatrix}\mathbf{V} \\
&= \begin{bmatrix} X^*f(A)X & 0 \\ 0 & (I-XX^*)^{1/2}f(A)(I-XX^*)^{1/2} \end{bmatrix}.
\end{aligned}
$$

Comparing the $(1,1)$-blocks gives $f(X^*AX) \le X^*f(A)X$.

(iv) $\Rightarrow$ (v). For A, B, X and Y as in (v) define

$$\mathbf{A} := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \quad \mathbf{X} := \begin{bmatrix} X & 0 \\ Y & 0 \end{bmatrix}.$$

Since $\mathbf{X}^*\mathbf{X} = \begin{bmatrix} X^*X+Y^*Y & 0 \\ 0 & 0 \end{bmatrix} \le \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$, $\|\mathbf{X}\| \le 1$. Also, $\mathbf{A} = \mathbf{A}^*$ and $\sigma(\mathbf{A}) = \sigma(A) \cup \sigma(B) \subset [0,\alpha)$. Since $\mathbf{X}^*\mathbf{A}\mathbf{X} = \begin{bmatrix} X^*AX+Y^*BY & 0 \\ 0 & 0 \end{bmatrix}$, we have

$$\begin{bmatrix} f(X^*AX+Y^*BY) & 0 \\ 0 & f(0)I \end{bmatrix} = f(\mathbf{X}^*\mathbf{A}\mathbf{X}) \le \mathbf{X}^*f(\mathbf{A})\mathbf{X} = \begin{bmatrix} X^*f(A)X+Y^*f(B)Y & 0 \\ 0 & 0 \end{bmatrix}$$

and hence $f(X^*AX+Y^*BY) \le X^*f(A)X+Y^*f(B)Y$.

(v) $\Rightarrow$ (vi). Let $X = P$ and $Y = 0$ in (v).

(vi) $\Rightarrow$ (i). For $A, B \in B(\mathcal{H})^{sa}$ with $\sigma(A), \sigma(B) \subset [0,\alpha)$ and $0 < \lambda < 1$, define

$$\mathbf{A} := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}, \quad \mathbf{U} := \begin{bmatrix} \sqrt{\lambda}\,I & -\sqrt{1-\lambda}\,I \\ \sqrt{1-\lambda}\,I & \sqrt{\lambda}\,I \end{bmatrix}, \quad \mathbf{P} := \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Then $\mathbf{A} = \mathbf{A}^*$ with $\sigma(\mathbf{A}) \subset [0,\alpha)$, $\mathbf{U}$ is a unitary and $\mathbf{P}$ is an orthogonal projection. Since

$$\mathbf{P}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{P} = \begin{bmatrix} \lambda A+(1-\lambda)B & 0 \\ 0 & 0 \end{bmatrix},$$

(vi) implies that

$$
\begin{aligned}
\begin{bmatrix} f(\lambda A+(1-\lambda)B) & 0 \\ 0 & f(0)I \end{bmatrix} &= f(\mathbf{P}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{P}) \\
&\le \mathbf{P}f(\mathbf{U}^*\mathbf{A}\mathbf{U})\mathbf{P} = \mathbf{P}\mathbf{U}^*f(\mathbf{A})\mathbf{U}\mathbf{P} \\
&= \begin{bmatrix} \lambda f(A)+(1-\lambda)f(B) & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}
$$

so that $f(\lambda A+(1-\lambda)B) \le \lambda f(A)+(1-\lambda)f(B)$ and $f(0) \le 0$.

Thus, (i), (iv), (v), and (vi) are equivalent. In the rest we prove that (i), (ii), and (iii) are equivalent.

(i) $\Rightarrow$ (ii). The operator convexity of f on $(0,\infty)$ is contained in (i), which of course implies the usual convexity of f on $(0,\infty)$. The latter implies that the limit $f(+0)$ exists and $f(+0) \le f(0)$.

(ii) $\Rightarrow$ (i). Define a function f_0 on $[0,\alpha)$ by $f_0(0) := f(+0)$ and $f_0(t) := f(t)$ for $t \in (0,\alpha)$. Then f_0 is continuous on $[0,\alpha)$ since it is convex on $(0,\alpha)$ in the usual sense. Let $A, B \in B(\mathcal{H})$ with $\sigma(A), \sigma(B) \subset [0,\alpha)$. When $\varepsilon > 0$ is small so that $\sigma(A+\varepsilon I), \sigma(B+\varepsilon I) \subset [0,\alpha)$, we have for every $\lambda \in (0,1)$

$$f(\lambda(A+\varepsilon I)+(1-\lambda)(B+\varepsilon I)) \le \lambda f(A+\varepsilon I)+(1-\lambda)f(B+\varepsilon I).$$

Letting $\varepsilon \searrow 0$ yields

$$f_0(\lambda A + (1-\lambda)B) \le \lambda f_0(A) + (1-\lambda) f_0(B),$$

that is, f_0 is operator convex on $[0,\alpha)$. Now let A and P be as in (vi). Let Q_0 be the orthogonal projection onto the kernel of A and $\tilde{Q}_0$ be that onto the kernel of PAP. Then $Q_1 := I - Q$ is the orthogonal projection onto the range of A and $\tilde{Q}_1 := I - \tilde{Q}_0$ is that onto the range of PAP. One can write

$$f(PAP) = f_0(PAP) + \alpha\tilde{Q}_0,$$
$$Pf(A)P = P(f_0(A) + \alpha Q_0)P = Pf_0(A)P + \alpha PQ_0P,$$

where $\alpha := f(0) - f(+0) \ge 0$. From (i) $\Rightarrow$ (vi) applied to f_0 we have $f_0(PAP) \le Pf_0(A)P$. Furthermore, taking account of the orthogonal decomposition $\mathbb{C}^n = P\mathbb{C}^n \oplus (I-P)\mathbb{C}^n$, we have

$$\begin{aligned} f(PAP) &= Pf(PAP)P + f(0)(I-P) \le Pf(PAP)P \\ &= Pf_0(PAP)P + \alpha P\tilde{Q}_0P \le Pf_0(A)P + \alpha P\tilde{Q}_0P. \end{aligned}$$

So, to see that f satisfies (vi) (hence (i)), it suffices to prove that $P\tilde{Q}_0P \le PQ_0P$. One can choose a $\delta > 0$ such that $A \ge \delta Q_1$ and $\tilde{Q}_1 \ge \delta PAP$. Hence $\tilde{Q}_1 \ge \delta^2 PQ_1P$ so that $(I-\tilde{Q}_1)PQ_1P(I-\tilde{Q}_1) = 0$, which implies that $Q_1P(I-\tilde{Q}_1) = 0$ so that $PQ_1P = \tilde{Q}_1PQ_1P\tilde{Q}_1 \le \tilde{Q}_1$ since $PQ_1P \le I$. Therefore, $P\tilde{Q}_1P \ge PQ_1P$ or $P\tilde{Q}_0P \le PQ_0P$.

(iv) $\Rightarrow$ (iii). Let $A, B \in B(\mathcal{H})^{sa}$ with $A \ge B > 0$. Letting $X := A^{-1/2}B^{1/2}$ we have $XX^* = A^{-1/2}BA^{-1/2} \le I$ and so $\|X\| \le I$. Since $B = X^*AX$, (iv) implies that

$$f(B) \le X^*f(A)X = B^{1/2}A^{-1/2}f(A)^{-1/2}B^{1/2}$$

and hence $A^{-1}f(A) = A^{-1/2}f(A)A^{-1/2} \ge B^{-1/2}f(B)B^{-1/2} = B^{-1}f(B)$. This means that $f(t)/t$ is operator monotone on $(0,\alpha)$.

(iii) $\Rightarrow$ (ii). First we prove that if g is a continuous operator monotone function on $[0,\alpha)$, then $tg(t)$ is operator convex on $[0,\alpha)$. Let $h(t) := tg(t)$ for $t \in [0,\alpha)$. To prove (vi) for h, we may assume that $A > 0$. In fact, one can take the limit of the inequality in (vi) for $A + \varepsilon I$ as $\varepsilon \searrow 0$. Since $A^{1/2}PA^{1/2} \le A$, we have $g(A^{1/2}PA^{1/2}) \le g(A)$. Multiplying $PA^{1/2}$ form the left and $A^{1/2}P$ for the right, we have

$$PA^{1/2}g(A^{1/2}PA^{1/2})A^{1/2}P \le PA^{1/2}g(A)A^{1/2}P.$$

Since $g(A^{1/2}PA^{1/2})A^{1/2}P = A^{1/2}Pg(PAP)$ by Lemma 2.5.1 (2), we have $h(PAP) \le Ph(A)P$, so h is operator convex on $[0,\alpha)$.

Assume that $f(t)/t$ is operator monotone on $(0,\alpha)$. By Theorem 2.4.1, $f(t)/t$ is continuous on $(0,\alpha)$. For each $\varepsilon > 0$, $f(t+\varepsilon)/(t+\varepsilon)$ is a continuous and operator monotone function on $[0,\alpha-\varepsilon)$. By what we proved just above, $\frac{t}{t+\varepsilon}f(t+\varepsilon)$ is operator convex on $[0,\alpha-\varepsilon)$. Hence f is operator convex on $(0,\alpha)$ by letting $\varepsilon \searrow 0$. This implies that (iii) $\Rightarrow$ (ii). Moreover, the convexity of f together with the non-decreasingness of $f(t)/t$ implies that $f(+0)$ exists and $f(+0) \le 0$. □

Theorem 2.5.3. *If $\alpha = \infty$ and $f(t) \le 0$ for all $t \in [0,\infty)$, then the conditions of Theorem 2.5.2 is also equivalent to* (vii) *$-f$ is operator monotone on* $[0,\infty)$.

Proof. Assume that $f \le 0$ on $[0,\infty)$. First we prove that (vii) is equivalent to that $-f$ is operator monotone on $(0,\infty)$ and $f(+0) \le f(0)$ (≤ 0). In fact, if (vii) holds, then it is immediate to see that $f(+0)$ exists and $f(+0) \le f(0)$. Conversely, assume that $-f$ is operator monotone on $(0,\infty)$ and $f(+0) \le f(0)$. Define f_0 on $[0,\infty)$ as in the proof of (ii) $\Rightarrow$ (i) above, so $-f_0$ is operator monotone on $[0,\infty)$. Let $A \ge B \ge 0$ in $B(\mathcal{H})$. Let Q_0 and $\tilde{Q}_0$ be the orthogonal projections onto the kernels of A and B, respectively. Then

$$f(A) = f_0(A) + \alpha Q_0, \qquad f(B) = f_0(B) + \alpha\tilde{Q}_0,$$

where $\alpha := f(0) - f(+0) \ge 0$. Since $A \ge B \ge 0$ yields $Q_0 \le \tilde{Q}_0$. With $f_0(A) \le f_0(B)$ this implies that $f(A) \le f(B)$. Thus it suffices to prove the equivalence between (i) and (vii) for the function f_0. Since f_0 is continuous on $[0,\infty)$ by Theorem 2.4.1, we assume in the rest that f is continuous on $[0,\infty)$.

(vii) $\Rightarrow$ (i). For A, X as in (iv) define $\mathbf{A}, \mathbf{U}$ as in the proof of (i) $\Rightarrow$ (iv) of Theorem 2.5.2, and set $R := (I-XX^*)^{1/2}A(I-XX^*)^{1/2}$ and $S := X^*A(I-XX^*)^{1/2}$. Moreover, define $\mathbf{B} := \begin{bmatrix} X^*AX + \varepsilon I & 0 \\ 0 & \beta I \end{bmatrix}$ for $\varepsilon, \beta > 0$. Since $\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} X^*AX & S \\ S^* & R \end{bmatrix}$, $\mathbf{B} - \mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} \varepsilon I & -S \\ -S^* & \beta I - R \end{bmatrix}$ and

$$\begin{aligned}\left\langle(\mathbf{B}-\mathbf{U}^*\mathbf{A}\mathbf{U})\begin{bmatrix}\xi\\ \eta\end{bmatrix},\begin{bmatrix}\xi\\ \eta\end{bmatrix}\right\rangle &= \left\langle\begin{bmatrix}\varepsilon\xi-S\eta\\ -S^*\xi+\beta\eta-R\eta\end{bmatrix},\begin{bmatrix}\xi\\ \eta\end{bmatrix}\right\rangle\\ &=\varepsilon\|\xi\|^2-2\operatorname{Re}\langle\xi,S\eta\rangle+\beta\|\eta\|^2-\langle R\eta,\eta\rangle\\ &\ge \varepsilon\|\xi\|^2-2\|S\|\,\|\xi\|\,\|\eta\|+(\beta-\|R\|)\|\eta\|^2\\ &=\left(\sqrt{\varepsilon}\|\xi\|-\frac{\|S\|}{\sqrt{\varepsilon}}\|\eta\|\right)^2+\left(\beta-\|R\|-\frac{\|S\|^2}{\varepsilon}\right)\|\eta\|^2\end{aligned}$$

for $\xi,\eta\in\mathbb{C}^n$, where $\operatorname{Re} z$ denotes the real part of $z\in\mathbb{C}$. Hence, for any $\varepsilon>0$, we have $\mathbf{B}-\mathbf{U}^*\mathbf{A}\mathbf{U}\ge 0$ if $\beta>0$ is sufficiently large. Then (vii) implies that

$$\begin{aligned}\begin{bmatrix}f(X^*AX+\varepsilon I) & 0\\ 0 & f(\beta)I\end{bmatrix} &= f(\mathbf{B})\le f(\mathbf{U}^*\mathbf{A}\mathbf{U})=\mathbf{U}^*\begin{bmatrix}f(A) & 0\\ 0 & f(0)I\end{bmatrix}\mathbf{U}\\ &\le \mathbf{U}^*\begin{bmatrix}f(A) & 0\\ 0 & 0\end{bmatrix}\mathbf{U}=\begin{bmatrix}X^*f(A)X & *\\ * & *\end{bmatrix}\end{aligned}$$

so that $f(X^*AX+\varepsilon I)\le X^*f(A)X$. Letting $\varepsilon\searrow 0$ yields $f(X^*AX)\le X^*f(A)X$. Hence (iv) is satisfied.

(i) $\Rightarrow$ (vii). Assume that $A\ge B\ge 0$ in $B(\mathcal{H})$. For each $\lambda\in(0,1)$, since $\lambda A=\lambda B+(1-\lambda)\lambda(1-\lambda)^{-1}(A-B)$,

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

Since $f(\lambda(1-\lambda)^{-1}(A-B))\le 0$ thanks to $f\le 0$, we have $f(\lambda A)\le\lambda f(B)$. Letting $\lambda\nearrow 1$ yields $f(A)\le f(B)$, which implies that $-f$ is operator monotone. □

Corollary 2.5.4. *If f is a function on $[0,\infty)$ such that $f\ge 0$, then f is operator monotone if and only if it is operator concave.*

Proof. This is the equivalence between (vii) and (i) above for $-f$. □

Lemma 2.5.5. *On $(0,\infty)$, the function t^{-1} is operator convex and $-t^{-1}$ is operator monotone. On $(-\infty,0)$, t^{-1} is operator concave and $-t^{-1}$ is operator monotone.*

Proof. Since $((1+t)/2)^{-1}\le(1+t^{-1})/2$ for $t>0$, for every $C>0$ in $\mathbb{M}_n$ we have

$$\left(\frac{I+C}{2}\right)^{-1}\le\frac{I+C^{-1}}{2}.$$

For every $A,B>0$ in $\mathbb{M}_n$, apply the above inequality to $C:=A^{-1/2}BA^{-1/2}$ to obtain

$$\begin{aligned}\left(\frac{A+B}{2}\right)^{-1} &= \left(\frac{A^{1/2}(I+A^{-1/2}BA^{-1/2})A^{1/2}}{2}\right)^{-1}\\ &=A^{-1/2}\left(\frac{I+A^{-1/2}BA^{-1/2}}{2}\right)^{-1}A^{-1/2}\\ &\le A^{-1/2}\,\frac{I+(A^{-1/2}BA^{-1/2})^{-1}}{2}A^{-1/2}=\frac{A^{-1}+B^{-1}}{2}.\end{aligned}$$

Hence t^{-1} is operator convex on $(0,\infty)$.

Next, assume that $A\ge B>0$ in $\mathbb{M}_n$. Since $B^{-1/2}AB^{-1/2}\ge I$, we have $B^{1/2}A^{-1}B^{1/2}=(B^{-1/2}AB^{-1/2})^{-1}\le I$ and hence $A^{-1}\le B^{-1}$. Hence $-t^{-1}$ is operator monotone on $(0,\infty)$. The assertions on $(-\infty,0)$ immediately follow from those on $(0,\infty)$ by taking account of the transformation $A\mapsto -A$. □

Corollary 2.5.6. *If f is a function on $(0,\infty)$ such that $f>0$, then* (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (iv) *hold concerning the following:*

(i) *f is operator monotone;*
(ii) *$t/f(t)$ is operator monotone;*
(iii) *f is operator concave;*
(iv) *$1/f(t)$ is operator convex.*

Proof. (i) $\Rightarrow$ (ii). For any $\varepsilon>0$, since $f(t+\varepsilon)$ is operator monotone on $[0,\infty)$ with $-f(t+\varepsilon)<0$, Theorem 2.5.3 implies that $-f(t+\varepsilon)/t$ is operator monotone on $(0,\infty)$. So Lemma 2.5.5 implies that $t/f(t+\varepsilon)=-(-f(t+\varepsilon)/t)^{-1}$ is operator monotone on $(0,\infty)$. Hence (ii) follows by letting $\varepsilon\searrow 0$.

(ii) $\Rightarrow$ (i). For any $\varepsilon>0$, since $(t+\varepsilon)/f(t+\varepsilon)$ is operator monotone on $[0,\infty)$ with $-(t+\varepsilon)/f(t+\varepsilon)<0$, Theorem 2.5.3 implies that $-(t+\varepsilon)/tf(t+\varepsilon)$ is operator monotone on $(0,\infty)$. So Lemma 2.5.5 implies that $tf(t+\varepsilon)/(t+\varepsilon)$ is operator monotone on $(0,\infty)$. Letting $\varepsilon\searrow 0$ gives (i).

(i) $\Leftrightarrow$ (iii). By Corollary 2.5.4 we see that

$$\begin{aligned}\text{(i)} &\iff f(t+\varepsilon) \text{ is operator monotone on } [0,\infty) \text{ for any } \varepsilon > 0\\ &\iff f(t+\varepsilon) \text{ is operator concave on } [0,\infty) \text{ for any } \varepsilon > 0\\ &\iff \text{(iii)}.\end{aligned}$$

(iii) $\Rightarrow$ (iv). Let $g(t) := 1/f(t)$ and assume that $A, B > 0$ in $\mathbb{M}_n$. Since (iii) implies that

$$f\left(\frac{A+B}{2}\right) \geq \frac{f(A)+f(B)}{2},$$

we have by Lemma 2.5.5

$$\begin{aligned}g\left(\frac{A+B}{2}\right) = f\left(\frac{A+B}{2}\right)^{-1} &\leq \left(\frac{f(A)+f(B)}{2}\right)^{-1}\\ &\leq \frac{f(A)^{-1}+f(B)^{-1}}{2} = \frac{g(A)+g(B)}{2}.\end{aligned}$$

Hence g is operator convex. $\square$

Note that (iv) $\Rightarrow$ (iii) is not valid in Corollary 2.5.6. For instance, when $1 \leq p \leq 2$, the functions t^{α} is operator convex on $(0,\infty)$ (see Example 2.5.9 (4)) but t^{-p} is not operator concave (even not concave in the usual sense) on $(0,\infty)$.

The following modification of Theorem 2.5.2 is also useful because the domain of f is a general interval and the condition $f(0) \leq 0$ in Theorem 2.5.2 is irrelevant.

Theorem 2.5.7. *Let f be a real-valued function on an interval J. Then the following conditions are equivalent, where finite-dimensional Hilbert spaces $\mathcal{H}, \mathcal{H}_j, \mathcal{K}$ are arbitrary and not fixed:*

(i) *f is operator convex;*

(ii) *for every $A \in B(\mathcal{H})^{sa}$ with $\sigma(A) \subset J$ and every isometry $X \in B(\mathcal{K},\mathcal{H})$,*

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

(iii) *for every $m \in \mathbb{N}$, every $A_j \in B(\mathcal{H}_j)^{sa}$ with $\sigma(A_j) \subset J$, $1 \leq j \leq m$, and every $X_j \in B(\mathcal{K},\mathcal{H}_j)$, $1 \leq j \leq m$, such that $\sum_{j=1}^m X_j^*X_j = I_{\mathcal{K}}$,*

$$f\left(\sum_{j=1}^m X_j^*A_jX_j\right) \leq \sum_{j=1}^m X_j^*f(A_j)X_j.$$

(iv) *for every $A, B \in B(\mathcal{H})^{sa}$ with $\sigma(A), \sigma(B) \subset J$ and every orthogonal projection $P \in B(\mathcal{H})$,*

$$f(PAP + (I-P)B(I-P)) \leq Pf(A)P + (I-P)f(B)(I-P).$$

Proof. (i) $\Rightarrow$ (ii). Let A, X be as in (ii), and choose any $B \in B(\mathcal{K})^{sa}$ with $\sigma(B) \subset J$. Since $X^*X = I_{\mathcal{K}}$ and hence $(XX^*)^2 = X(X^*X)X = XX^*$, it follows that $XX^* \in B(\mathcal{H})$ is an orthogonal projection. Define $Q := I_{\mathcal{H}} - XX^* \in B(\mathcal{H})$, an orthogonal projection, and $\mathbf{A}, \mathbf{U}, \mathbf{V} \in B(\mathcal{K} \oplus \mathcal{H})$ by

$$\mathbf{A} := \begin{bmatrix} B & 0 \\ 0 & A \end{bmatrix}, \quad \mathbf{U} := \begin{bmatrix} 0 & X^* \\ X & Q \end{bmatrix}, \quad \mathbf{V} := \begin{bmatrix} 0 & -X^* \\ -X & Q \end{bmatrix}.$$

We have

$$(QX)^*(QX) = X^*QX = X^*X - (X^*X)^2 = 0$$

so that $QX = 0$ and $X^*Q = 0$. Hence,

$$\mathbf{U}^*\mathbf{U} = \begin{bmatrix} X^*X & X^*Q \\ QX & XX^*+Q \end{bmatrix} = \begin{bmatrix} I_{\mathcal{K}} & 0 \\ 0 & I_{\mathcal{H}} \end{bmatrix}$$

so that $\mathbf{U}$ is a unitary and similarly for $\mathbf{V}$. Moreover, $\mathbf{A} \in B(\mathcal{K} \oplus \mathcal{H})^{sa}$ and $\sigma(\mathbf{A}) \subset J$. Since

$$\mathbf{U}^*\mathbf{A}\mathbf{U} = \begin{bmatrix} X^*AX & X^*AQ \\ QAX & XBX^*+QAQ \end{bmatrix}, \quad \mathbf{V}^*\mathbf{A}\mathbf{V} = \begin{bmatrix} X^*AX & -X^*AQ \\ -QAX & XBX^*+QAQ \end{bmatrix},$$

by (i) with Lemma 2.5.1 (1) we have

$$\begin{aligned}\begin{bmatrix} f(X^*AX) & 0 \\ 0 & f(XBX^*+QAQ)\end{bmatrix} &= f\left(\frac{\mathbf{U}^*\mathbf{A}\mathbf{U}+\mathbf{V}^*\mathbf{A}\mathbf{V}}{2}\right)\\ &\le \frac{f(\mathbf{U}^*\mathbf{A}\mathbf{U})+f(\mathbf{V}^*\mathbf{A}\mathbf{V})}{2}\\ &= \frac{1}{2}\mathbf{U}^*\begin{bmatrix} f(B) & 0 \\ 0 & f(A)\end{bmatrix}\mathbf{U}+\frac{1}{2}\mathbf{V}^*\begin{bmatrix} f(B) & 0 \\ 0 & f(A)\end{bmatrix}\mathbf{V}\\ &= \begin{bmatrix} X^*f(A)X & 0 \\ 0 & Xf(B)X^*+Qf(A)Q\end{bmatrix},\end{aligned}$$

implying that $f(X^*AX) \le X^*f(A)X$.

(ii) $\Rightarrow$ (iii). Let A_j, X_j for $1 \le j \le m$ be as in (iii). Define $\mathbf{A} \in B(\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_m)$ and $\mathbf{X} \in B(\mathcal{K}, \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_m)$ by

$$\mathbf{A} := \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_m\end{bmatrix}, \quad \text{i.e.,} \quad \mathbf{A}(x_1 \oplus \cdots \oplus x_m) = A_1x_1 \oplus \cdots \oplus A_mx_m \quad \text{for } x_j \in \mathcal{H}_j,$$

$$\mathbf{X} := \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m\end{bmatrix}, \quad \text{i.e.,} \quad \mathbf{X}y = X_1y \oplus \cdots \oplus X_my \quad \text{for } y \in \mathcal{K}.$$

Then $\mathbf{A} = \mathbf{A}^*$ and $\sigma(\mathbf{A}) \subset J$. Moreover, $\mathbf{X}$ is an isometry since $\mathbf{X}^*\mathbf{X} = \sum_{j=1}^m X_j^*X_j = I_{\mathcal{K}}$. Hence (ii) implies that

$$f\left(\sum_{j=1}^m X_j^*A_jX_j\right) = f(\mathbf{X}^*\mathbf{A}\mathbf{X}) \le \mathbf{X}^*f(\mathbf{A})\mathbf{X} = \sum_{j=1}^m X_j^*f(A_j)X_j.$$

(iii) $\Rightarrow$ (iv) is obvious.

(iv) $\Rightarrow$ (i). Let $A, B \in B(\mathcal{H})^{sa}$ with $\sigma(A), \sigma(B) \subset J$, and let $0 < \lambda < 1$. Define $\mathbf{A}, \mathbf{U}, \mathbf{P} \in B(\mathcal{H} \oplus \mathcal{H})$ in the same way as in the proof of (vi) $\Rightarrow$ (i) of Theorem 2.5.2. Since

$$\mathbf{P}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{P} + (I-\mathbf{P})\mathbf{U}^*\mathbf{A}\mathbf{U}(I-\mathbf{P}) = \begin{bmatrix} \lambda A + (1-\lambda)B & 0 \\ 0 & (1-\lambda)A + \lambda B\end{bmatrix},$$

(iv) implies that

$$\begin{aligned}&\begin{bmatrix} f(\lambda A + (1-\lambda)B) & 0 \\ 0 & f((1-\lambda)A + \lambda B)\end{bmatrix}\\ &\le \mathbf{P}f(\mathbf{U}^*\mathbf{A}\mathbf{U})\mathbf{P} + (I-\mathbf{P})f(\mathbf{U}^*\mathbf{A}\mathbf{U})(I-\mathbf{P})\\ &= \mathbf{P}\mathbf{U}^*\begin{bmatrix} f(A) & 0 \\ 0 & f(B)\end{bmatrix}\mathbf{U}\mathbf{P} + (I-\mathbf{P})\mathbf{U}^*\begin{bmatrix} f(A) & 0 \\ 0 & f(B)\end{bmatrix}\mathbf{U}(I-\mathbf{P})\\ &= \begin{bmatrix} \lambda f(A) + (1-\lambda)f(B) & 0 \\ 0 & (1-\lambda)f(A) + \lambda f(B)\end{bmatrix}\end{aligned}$$

so that $f(\lambda A + (1-\lambda)B) \le \lambda f(A) + (1-\lambda)f(B)$. □

Exercise 2.5.8. When f is a real-valued function on $J = [a, b]$, show that the conditions of Theorem 2.5.7 are also equivalent to

(v) f is operator convex on (a, b), $f(a+0) \le f(a)$ and $f(b-0) \le f(b)$, where $f(a+0) := \lim_{t\searrow a} f(t)$ and $f(b-0) := \lim_{t\nearrow b} f(t)$.

The following are basic examples of operator monotone and operator convex functions.

Example 2.5.9.

(1) When $\alpha \ge 0$, $\alpha t + \beta$ is operator monotone on $\mathbb{R}$.

(2) When $c \notin (\alpha, \beta)$, $(c-t)^{-1}$ is operator monotone on (α, β).

(3) When $0 \le p \le 1$, t^p is operator monotone and operator concave on $[0, \infty)$. Moreover,

$$\{p \in \mathbb{R} : t^p \text{ is operator monotone on } (0, \infty)\} = [0, 1].$$

(4) When $1 \le p \le 2$, t^p is operator convex on $[0, \infty)$. Moreover,

$$\{p \in \mathbb{R} : t^p \text{ is operator convex on } (0,\infty)\} = [-1,0] \cup [1,2].$$

(5) $f(t) := (t-1)/\log t$ on $[0,\infty)$ is operator monotone, where $f(0) = 0$ and $f(1) = 1$.
(6) $\log t$ on $(0,\infty)$ is operator monotone and operator concave.
(7) $t \log t$ on $[0,\infty)$ is operator convex, where $0 \log 0 = 0$.

Proof. (1) Obvious.

(2) Assume that $A, B \in \mathbb{M}_n^{sa}$, $\sigma(A), \sigma(B) \subset (\alpha, \beta)$ and $A \geq B$. If $c \leq \alpha$ then $A - cI \geq B - cI > 0$ and so $(A - cI)^{-1} \leq (B - cI)^{-1}$. Hence $(cI - A)^{-1} \geq (cI - B)^{-1}$. If $c \geq \beta$ then $0 < cI - A \leq cI - B$ and so $(cI - A)^{-1} \geq (cI - B)^{-1}$.

(3) The first assertion follows from Theorem 2.1.1 and Corollary 2.5.4. When $p < 0$, t^p on $(0,\infty)$ is not monotone increasing and so it is not operator monotone. When $p > 1$, $t/t^p = t^{1-p}$ on $(0,\infty)$ is not operator monotone, and hence by Corollary 2.5.6, t^p is not operator monotone.

(4) When $1 \leq p \leq 2$, $t^p/t = t^{p-1}$ on $(0,\infty)$ is operator monotone, and hence by Theorem 2.5.2, t^p is operator convex on $[0,\infty)$. Moreover, when $-1 \leq p \leq 0$, $t^p = 1/t^{-p}$ on $(0,\infty)$ is operator convex by Corollary 2.5.6. When $p > 0$, since the operator convexity of t^p on $(0,\infty)$ is equivalent to that on $[0,\infty)$ by continuity, Theorem 2.5.2 shows that t^p on $(0,\infty)$ is operator convex only if $t^p/t = t^{p-1}$ is operator monotone on $(0,\infty)$. Hence, when $p \in (0,1) \cup (2,\infty)$, t^p is not operator convex. That t^p is not operator convex when $p < -1$ is left for an exercise in Section 2.7 (Exercise 2.7.10).

(5) Notice that $f(t) = \int_0^1 t^p \, dp$ for all $t \geq 0$. This shows that $f(A) = \int_0^1 A^p \, dp \geq \int_0^1 B^p \, dp = f(B)$ if $A \geq B \geq 0$ in $\mathbb{M}_n$. Hence f is operator monotone on $[0,\infty)$.

(6) By (5), $t/\log(1+t)$ is operator monotone on $(0,\infty)$ and so $\log(1+t)$ is operator monotone and operator concave on $(0,\infty)$ by Corollary 2.5.6. Hence the result follows since $\log(\varepsilon + t) = \log \varepsilon + \log(1 + \varepsilon^{-1} t)$ is operator monotone and operator convex on $(0,\infty)$ for every $\varepsilon > 0$.

(7) Since $g(t) := t \log t$ is continuous on $[0,\infty)$ and $g(t)/t = \log t$ is operator monotone on $(0,\infty)$, g is operator convex by Theorem 2.5.2. $\square$

The following is Furuta's observation in [33], which provides a simple way to prove (6) and (7) above by only using the operator monotonicity (or the Löwner–Heinz inequality) and the operator concavity of the function t^p, $t > 0$, for $0 < p < 1$. Assume that $A \geq B > 0$ in $B(\mathcal{H})$. Since $A^p \geq B^p$ for all $p \in (0,1)$, we have

$$\log A = \lim_{p \searrow 0} \frac{1}{p}(A^p - I) \geq \lim_{p \searrow 0} \frac{1}{p}(B^p - I) = \log B.$$

For every $\lambda, p \in (0,1)$ we have $(\lambda A + (1-\lambda)B)^p \geq \lambda A^p + (1-\lambda)B^p$ so that

$$\frac{1}{p}\{(\lambda A + (1-\lambda)B)^p - I\} \geq \frac{\lambda}{p}(A^p - I) + \frac{1-\lambda}{p}(B^p - I),$$

$$\frac{1}{1-p}\{(\lambda A + (1-\lambda)B) - (\lambda A + (1-\lambda)B)^p\} \leq \frac{\lambda}{1-p}(A - A^p) + \frac{1-\lambda}{1-p}(B - B^p).$$

Taking the limits of the above as $p \searrow 0$ and $p \nearrow 1$, respectively, we have

$$\log(\lambda A + (1-\lambda)B) \geq \lambda \log A + (1-\lambda) \log B,$$

$$(\lambda A + (1-\lambda)B) \log(\lambda A + (1-\lambda)B) \leq \lambda A \log A + (1-\lambda) B \log B.$$

2.6 Pick functions

Let $\mathbb{C}^+$ denote the upper half-plane, i.e., $\mathbb{C}^+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$, where $\operatorname{Im} z$ is the imaginary part of z. A function $f : \mathbb{C}^+ \to \mathbb{C}$ is called a *Pick function* if f is analytic in $\mathbb{C}^+$ and the range $f(\mathbb{C}^+)$ is included in the closed half-plane $\{z \in \mathbb{C} : \operatorname{Im} z \geq 0\}$. The set of all Pick functions is denoted by $\mathcal{P}$. From the open mapping theorem in complex function theory (see [27, p. 99] for example) we note that if $f \in \mathcal{P}$ is not constant then the range $f(\mathbb{C}^+)$ is a domain (i.e., a connected open subset) of $\mathbb{C}$ and so $f(\mathbb{C}^+) \subset \mathbb{C}^+$. Obviously, $\mathcal{P}$ is a convex cone, and if $f, g \in \mathcal{P}$ with g non-constant, then $f \circ g \in \mathcal{P}$ as well. Typical examples of Pick functions are given in the following exercise.

Exercise 2.6.1. Verify the following:
(1) When $0 < p \leq 1$, the function $f(z) = z^p := r^p e^{ip\theta}$ (the principal branch of z^p) for $z = re^{i\theta}$ with $r > 0$ and $0 < \theta < \pi$ is in $\mathcal{P}$.
(2) $f(z) = \operatorname{Log} z := \log r + i\theta$ (the principal branch of $\log z$) for $z = re^{i\theta}$ is in $\mathcal{P}$.
(3) $f(z) = -1/z$ is in $\mathcal{P}$.
(4) $f(z) = \tan z := \sin z/\cos z$ is in $\mathcal{P}$, where $\cos z := (e^{iz} + e^{-iz})/2$ and $\sin z := (e^{iz} - e^{-iz})/2i$.

The next *Nevanlinna's theorem* provides the integral representation of Pick functions.

Theorem 2.6.2. *A function $f : \mathbb{C}^+ \to \mathbb{C}$ is in $\mathcal{P}$ if and only if there exist an $\alpha \in \mathbb{R}$, a $\beta \geq 0$ and a positive finite Borel measure ν on $\mathbb{R}$ such that*

$$f(z) = \alpha + \beta z + \int_{-\infty}^{\infty} \frac{1 + \lambda z}{\lambda - z}\, d\nu(\lambda), \qquad z \in \mathbb{C}^+. \tag{2.6.1}$$

The integral representation (2.6.1) is also written as

$$f(z) = \alpha + \beta z + \int_{-\infty}^{\infty} \left(\frac{1}{\lambda - z} - \frac{\lambda}{\lambda^2 + 1} \right) d\mu(\lambda), \qquad z \in \mathbb{C}^+, \tag{2.6.2}$$

where μ is a positive Borel measure on $\mathbb{R}$ given by $d\mu(\lambda) := (\lambda^2 + 1)\, d\nu(\lambda)$ and so

$$\int_{-\infty}^{\infty} \frac{1}{\lambda^2 + 1}\, d\mu(\lambda) < +\infty.$$

The proof of the "if" part is easy. Assume that f is defined on $\mathbb{C}^+$ as in (2.6.1). For each $z \in \mathbb{C}^+$, since

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} = \beta + \int_{\mathbb{R}} \frac{\lambda^2 + 1}{(\lambda - z)(\lambda - z - \Delta z)}\, d\nu(\lambda)$$

and

$$\sup\left\{ \left| \frac{\lambda^2 + 1}{(\lambda - z)(\lambda - z - \Delta z)} \right| : \lambda \in \mathbb{R},\ |\Delta z| < \frac{\operatorname{Im} z}{2} \right\} < +\infty,$$

it follows from the Lebesgue dominated convergence theorem that

$$\lim_{\Delta \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \beta + \int_{\mathbb{R}} \frac{\lambda^2 + 1}{(\lambda - z)^2}\, d\nu(\lambda).$$

Hence f is analytic in $\mathbb{C}^+$. Since

$$\operatorname{Im}\left(\frac{1 + \lambda z}{\lambda - z} \right) = \frac{(\lambda^2 + 1)\operatorname{Im} z}{|\lambda - z|^2}, \qquad z \in \mathbb{C}^+,$$

we have

$$\operatorname{Im} f(z) = \left(\beta + \int_{\mathbb{R}} \frac{\lambda^2 + 1}{|\lambda - z|^2}\, d\nu(\lambda) \right) \operatorname{Im} z \ge 0$$

for all $z \in \mathbb{C}^+$. Therefore, we have $f \in \mathcal{P}$. The equivalence between the two representations (2.6.1) and (2.6.2) is immediately seen from

$$\frac{1 + \lambda z}{\lambda - z} = (\lambda^2 + 1)\left(\frac{1}{\lambda - z} - \frac{\lambda}{\lambda^2 + 1} \right).$$

The "only if" is the significant part, whose proof based on the Poisson integral formula is exposed in Appendix A.4.

Moreover, we note that α, β and ν in Theorem 2.6.2 are uniquely determined by f. In fact, letting $z = i$ in (2.6.1) we have $\alpha = \operatorname{Re} f(i)$. Letting $z = iy$ with $y > 0$ we have

$$f(iy) = \alpha + i\beta y + \int_{-\infty}^{\infty} \frac{\lambda(1 - y^2) + iy(\lambda^2 + 1)}{\lambda^2 + y^2}\, d\nu(\lambda)$$

so that

$$\frac{\operatorname{Im} f(iy)}{y} = \beta + \int_{-\infty}^{\infty} \frac{\lambda^2 + 1}{\lambda^2 + y^2}\, d\nu(\lambda).$$

By the Lebesgue dominated convergence theorem this yields

$$\beta = \lim_{y \to \infty} \frac{\operatorname{Im} f(iy)}{y}.$$

Hence α and β are uniquely determined by f. By (2.6.2), for $z = x + iy$ we have

$$\operatorname{Im} f(x + iy) = \beta y + \int_{-\infty}^{\infty} \frac{y}{(x - \lambda)^2 + y^2}\, d\mu(\lambda), \qquad x \in \mathbb{R},\ y > 0. \tag{2.6.3}$$

Thus the uniqueness of μ (hence ν) is a consequence of the so-called *Stieltjes inversion formula*. For details omitted here, see [30, pp. 24–26] and [13, pp. 139–141].

For any open interval (a, b), $-\infty \le a < b \le \infty$, we denote by $\mathcal{P}(a, b)$ the set of all Pick functions which admit an analytic continuation across (a, b) by reflection into the lower half-plane $\mathbb{C}^- := \{z \in \mathbb{C} : \operatorname{Im} z < 0\}$. More precisely,

$\mathcal{P}(a,b)$ is the set of all $f \in \mathcal{P}$ with an analytic continuation (denoted by the same f) in $(\mathbb{C}\setminus\mathbb{R})\cup(a,b)$ so that $f(\overline{z}) = \overline{f(z)}$ for all $z \in \mathbb{C}^+$. Note that if $f \in \mathcal{P}(a,b)$ then $f(x) \in \mathbb{R}$ for all $x \in (a,b)$.

The next theorem is a specialization of Nevanlinna's theorem to functions in $\mathcal{P}(a,b)$.

Theorem 2.6.3. *A function $f : \mathbb{C}^+ \to \mathbb{C}$ is in $\mathcal{P}(a,b)$ if and only if f is represented as in* (2.6.1) *with $\alpha \in \mathbb{R}$, $\beta \geq 0$ and a positive finite Borel measure ν on $\mathbb{R}\setminus(a,b)$.*

Proof. Let $f \in \mathcal{P}$ be represented as in (2.6.1) with $\alpha \in \mathbb{R}$, $\beta \geq 0$ and a positive finite Borel measure ν on $\mathbb{R}$. It suffices to prove that $f \in \mathcal{P}(a,b)$ if and only if $\nu((a,b)) = 0$. First, assume that $\nu((a,b)) = 0$. The function f expressed by (2.6.1) is analytic in $\mathbb{C}^+ \cup \mathbb{C}^-$ so that $f(\overline{z}) = \overline{f(z)}$ for all $z \in \mathbb{C}^+$. For every $x \in (a,b)$, since

$$\sup\left\{\left|\frac{\lambda^2+1}{(\lambda-x)(\lambda-x-\Delta z)}\right| : \lambda \in \mathbb{R}\setminus(a,b),\ |\Delta z| < \frac{1}{2}\min\{x-a, b-x\}\right\} < +\infty,$$

the above proof of the "if" part of Theorem 2.6.2 by using the Lebesgue dominated convergence theorem can work for $z = x$ as well, and so f is differentiable (in the complex variable z) at $z = x$. Hence $f \in \mathcal{P}(a,b)$.

Conversely, assume that $f \in \mathcal{P}(a,b)$. It follows from (2.6.3) that

$$\int_{-\infty}^{\infty} \frac{1}{(x-\lambda)^2+y^2}\, d\mu(\lambda) = \frac{\operatorname{Im} f(x+iy)}{y} - \beta, \qquad x \in \mathbb{R},\ y > 0.$$

For any $x \in (a,b)$, since $f(x) \in \mathbb{R}$, we have

$$\frac{\operatorname{Im} f(x+iy)}{y} = \operatorname{Im}\frac{f(x+iy)-f(x)}{y} = \operatorname{Re}\frac{f(x+iy)-f(x)}{iy} \longrightarrow \operatorname{Re} f'(x) \quad \text{as } y \searrow 0,$$

and so the monotone convergence theorem yields

$$\int_{-\infty}^{\infty} \frac{1}{(x-\lambda)^2}\, d\mu(\lambda) = \operatorname{Re} f'(x), \qquad x \in (a,b).$$

Hence, for any closed interval $[c,d]$ included in (a,b), we have

$$R := \sup_{x\in[c,d]} \int_{-\infty}^{\infty} \frac{1}{(x-\lambda)^2}\, d\mu(\lambda) = \sup_{x\in[c,d]} \operatorname{Re} f'(x) < +\infty.$$

For each $m \in \mathbb{N}$ let $c_k := c + (k/m)(d-c)$ for $k = 0, 1, \ldots, m$. Then

$$\begin{aligned}\mu([c,d)) &= \sum_{k=1}^{m} \mu([c_{k-1},c_k)) \leq \sum_{k=1}^{m} \int_{[c_{k-1},c_k)} \frac{(c_k-c_{k-1})^2}{(c_k-\lambda)^2}\, d\mu(\lambda) \\ &\leq \sum_{k=1}^{m}\left(\frac{d-c}{m}\right)^2 \int_{-\infty}^{\infty} \frac{1}{(c_k-\lambda)^2}\, d\mu(\lambda) \leq \frac{(d-c)^2R}{m}.\end{aligned}$$

Letting $m \to \infty$ gives $\mu([c,d)) = 0$. This implies that $\mu((a,b)) = 0$ and so $\nu((a,b)) = 0$. $\square$

Now let $f \in \mathcal{P}(a,b)$. The above theorem says that $f(x)$ on (a,b) admits the integral representation

$$\begin{aligned}f(x) &= \alpha + \beta x + \int_{\mathbb{R}\setminus(a,b)} \frac{1+\lambda x}{\lambda - x}\, d\nu(\lambda) \\ &= \alpha + \beta x + \int_{\mathbb{R}\setminus(a,b)} (\lambda^2+1)\left(\frac{1}{\lambda-x} - \frac{\lambda}{\lambda^2+1}\right) d\nu(\lambda), \qquad x \in (a,b),\end{aligned}$$

where α, β and ν are as in the theorem. For any $n \in \mathbb{N}$ and $A, B \in \mathbb{M}_n^{sa}(a,b)$, if $A \geq B$ then $(\lambda I - A)^{-1} \geq (\lambda I - B)^{-1}$ for all $\lambda \in \mathbb{R}\setminus(a,b)$ (see Example 2.5.9 (2)) and hence we have

$$\begin{aligned}f(A) &= \alpha I + \beta A + \int_{\mathbb{R}\setminus(a,b)} (\lambda^2+1)\left((\lambda I - A)^{-1} - \frac{\lambda}{\lambda^2+1} I\right) d\nu(\lambda) \\ &\geq \alpha I + \beta B + \int_{\mathbb{R}\setminus(a,b)} (\lambda^2+1)\left((\lambda I - B)^{-1} - \frac{\lambda}{\lambda^2+1} I\right) d\nu(\lambda) = f(B).\end{aligned}$$

Therefore, f is operator monotone on (a,b). In the next section we will prove Löwner's theorem saying that the converse is also true so that f is operator monotone on (a,b) if and only if $f \in \mathcal{P}(a,b)$.

The following are examples of integral representations for typical Pick functions from Exercise 2.6.1.

Example 2.6.4. The principal branch $\operatorname{Log} z$ of the logarithm in Exercise 2.6.1 (2) is in $\mathcal{P}(0,\infty)$. Its integral representation in the form (2.6.2) is

$$\operatorname{Log} z = \int_{-\infty}^{0} \left(\frac{1}{\lambda - z} - \frac{\lambda}{\lambda^2 + 1} \right) d\lambda, \qquad z \in \mathbb{C}^+.$$

To show this, it suffices to verify the above expression for $z = x \in (0, \infty)$, that is,

$$\log x = \int_{0}^{\infty} \left(-\frac{1}{\lambda + x} + \frac{\lambda}{\lambda^2 + 1} \right) d\lambda, \qquad x \in (0, \infty),$$

which is immediate by a direct computation.

Example 2.6.5. When $0 < p < 1$, the principal branch of z^p in Example 2.6.1 (1) is in $\mathcal{P}(0, \infty)$, whose integral representation in the form (2.6.2) is

$$z^p = \cos \frac{p\pi}{2} + \frac{\sin p\pi}{\pi} \int_{-\infty}^{0} \left(\frac{1}{\lambda - z} - \frac{\lambda}{\lambda^2 + 1} \right) |\lambda|^p \, d\lambda, \qquad z \in \mathbb{C}^+.$$

For this it suffices to verify that

$$x^p = \cos \frac{p\pi}{2} + \frac{\sin p\pi}{\pi} \int_{0}^{\infty} \left(-\frac{1}{\lambda + x} + \frac{\lambda}{\lambda^2 + 1} \right) \lambda^p \, d\lambda, \qquad x \in (0, \infty), \tag{2.6.4}$$

which can be shown as in the following exercise.

Exercise 2.6.6. When $0 < p < 1$, show (2.6.4) as follows:

(a) Consider the integration of the function

$$\frac{z^{p-1}}{1+z} := \frac{r^{p-1} e^{i(p-1)\theta}}{1 + re^{i\theta}}, \qquad z = re^{i\theta}, \ 0 < \theta < 2\pi,$$

that is analytic in the cut plane $\mathbb{C} \setminus [0, \infty)$ except -1, along the contour

$$z = \begin{cases} re^{i\theta} & (\varepsilon \le r \le R, \, \theta = +0), \\ Re^{i\theta} & (0 < \theta < 2\pi), \\ re^{i\theta} & (R \ge r \ge \varepsilon, \, \theta = 2\pi - 0), \\ \varepsilon e^{i\theta} & (2\pi > \theta > 0), \end{cases}$$

where $0 < \varepsilon < 1 < R$. Apply the residue theorem (see [27, p. 112]) and let $\varepsilon \searrow 0$ and $R \nearrow \infty$ to show that

$$\int_{0}^{\infty} \frac{t^{p-1}}{1+t} \, dt = \frac{\pi}{\sin p\pi}. \tag{2.6.5}$$

(b) For each $x > 0$, substitute λ/x for t in (2.6.5) to obtain

$$x^p = \frac{\sin p\pi}{\pi} \int_{0}^{\infty} \frac{x \lambda^{p-1}}{\lambda + x} \, d\lambda, \qquad x \in (0, \infty).$$

(c) Since

$$\frac{x}{\lambda + x} = \frac{1}{\lambda^2 + 1} + \left(\frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x} \right) \lambda,$$

it follows that

$$x^p = \frac{\sin p\pi}{\pi} \int_{0}^{\infty} \frac{\lambda^{p-1}}{\lambda^2 + 1} \, d\lambda + \frac{\sin p\pi}{\pi} \int_{0}^{\infty} \left(\frac{\lambda}{\lambda^2 + 1} - \frac{1}{\lambda + x} \right) \lambda^p \, d\lambda, \qquad x \in (0, \infty).$$

Substitute λ^2 for t in (2.6.5) with p replaced by $p/2$ to obtain

$$\int_{0}^{\infty} \frac{\lambda^{p-1}}{\lambda^2 + 1} \, d\lambda = \frac{\pi}{2 \sin \frac{p\pi}{2}}.$$

Hence (2.6.4) follows.

2.7 Löwner's theorem

The main aim of this section is to prove the primary result in Löwner's theory saying that an operator monotone function on (a, b) belongs to $\mathcal{P}(a, b)$. Apart from Löwner's original proof, three different proofs are known so far, which are by Bendat and Sherman [12] based on the Hamburger moment problem, by Korányi [51] (also found in [3]) based on the the spectral theorem of self-adjoint operators, and by Hansen and Pedersen [35] based on the Krein–

Milman theorem. In all of them, the integral representation of operator monotone functions was obtained to prove Löwner's theorem. The proof below is based on [35].

Operator monotone (or operator convex) functions on an finite open interval (a,b) are transformed into those on a symmetric interval $(-1,1)$ via the affine function $x \in (-1,1) \mapsto \frac{b-a}{2}x + \frac{b+a}{2}$. So it is essential to analyze operator monotone (or operator convex) functions on $(-1,1)$. Theorem 2.4.1 says that every operator monotone function f on $(-1,1)$ is C^1 on $(-1,1)$ and $f'(0) > 0$ unless f is constant. Taking $(f - f(0))/f'(0)$ we may assume that $f(0) = 0$ and $f'(0) = 1$. So let $\mathcal{K}$ denote the set of all operator monotone functions on $(-1,1)$ such that $f(0) = 0$ and $f'(0) = 1$.

Lemma 2.7.1. *Let f be an operator monotone function on $(-1,1)$. Then*

(1) *For every $\alpha \in [-1,1]$, $(x+\alpha)f(x)$ is operator convex on $(-1,1)$.*
(2) *For every $\alpha \in [-1,1]$, $(1+\frac{\alpha}{x})f(x)$ is operator monotone on $(-1,1)$.*
(3) *f is twice differentiable at 0 and*

$$\frac{f''(0)}{2} = \lim_{x\to 0} \frac{f(x) - f'(0)x}{x^2}.$$

Proof. (1) For any $\varepsilon \in (0,1)$, since $f(x-1+\varepsilon)$ is operator monotone on $[0,2-\varepsilon)$, it follows from Theorem 2.5.2 that $xf(x-1+\varepsilon)$ is operator convex on $[0,2-\varepsilon)$ and so $(x+1-\varepsilon)f(x)$ is operator convex on $(-1+\varepsilon,1)$. By letting $\varepsilon \searrow 0$, $(x+1)f(x)$ is operator convex on $(-1,1)$. Applying this to the operator monotone function $-f(-x)$ implies that $-(x+1)f(-x)$ is operator convex on $(-1,1)$. Hence so is $(x-1)f(x)$ by changing x to $-x$. Moreover, for every $\alpha \in [-1,1]$, since $(x+\alpha)f(x) = \frac{1+\alpha}{2}(x+1)f(x) + \frac{1-\alpha}{2}(x-1)f(x)$, $(x+\alpha)f(x)$ is operator convex on $(-1,1)$.

(2) For every $\alpha \in [-1,1]$, set $g(x) := (x+\alpha)f(x)$. By (1) and Corollary 2.4.6, $g^{[1]}(0,x) = g(x)/x = (1+\frac{\alpha}{x})f(x)$ is operator monotone on $(-1,1)$.

(3) Although Theorem 2.4.7 implies that f is actually C^∞ on $(-1,1)$, we give a proof that is tailor-made for the situation of our exposition. By (2) and Theorem 2.4.1, $(1+\frac{1}{x})f(x)$ as well as $f(x)$ is C^1 on $(-1,1)$ so that the function h on $(-1,1)$ defined by $h(x) := f(x)/x$ for $x \neq 0$ and $h(0) := f'(0)$ is C^1. This implies that

$$h'(x) = \frac{f'(x)x - f(x)}{x^2} \longrightarrow h'(0) \quad \text{as } x \to 0.$$

Therefore,

$$f'(x)x = f(x) + h'(0)x^2 + o(|x|^2)$$

so that

$$f'(x) = h(x) + h'(0)x + o(|x|) = h(0) + 2h'(0)x + o(|x|) \quad \text{as } x \to 0,$$

which shows that f is twice differentiable at 0 with $f''(0) = 2h'(0)$. Hence

$$\frac{f''(0)}{2} = h'(0) = \lim_{x\to 0}\frac{h(x) - h(0)}{x} = \lim_{x\to 0}\frac{f(x) - f'(0)x}{x^2}. \qquad \square$$

Lemma 2.7.2. *If $f \in \mathcal{K}$ then*

$$f(x) \le \frac{x}{1-x} \quad \text{for } 0 \le x < 1,$$
$$f(x) \ge \frac{x}{1+x} \quad \text{for } -1 < x \le 0,$$
$$|f''(0)| \le 2.$$

Proof. For every $x \in (-1,1)$, Theorem 2.4.3 implies that

$$\begin{bmatrix} f^{[1]}(x,x) & f^{[1]}(x,0) \\ f^{[1]}(x,0) & f^{[1]}(0,0) \end{bmatrix} = \begin{bmatrix} f'(x) & f(x)/x \\ f(x)/x & 1 \end{bmatrix} \ge 0,$$

and hence

$$\frac{f(x)^2}{x^2} \le f'(x). \tag{2.7.1}$$

By Lemma 2.7.1 (1),

$$\frac{d}{dx}(x \pm 1)f(x) = f(x) + (x \pm 1)f'(x)$$

is increasing on $(-1,1)$. Since $f(0) \pm f'(0) = \pm 1$, we have

$$f(x) + (x-1)f'(x) \geq -1 \quad \text{for} \quad 0 < x < 1, \tag{2.7.2}$$
$$f(x) + (x+1)f'(x) \leq 1 \quad \text{for} \ -1 < x < 0, \tag{2.7.3}$$

By (2.7.1) and (2.7.2) we have

$$f(x) + 1 \geq \frac{(1-x)f(x)^2}{x^2}.$$

If $f(x) > \frac{x}{1-x}$ for some $x \in (0,1)$, then

$$f(x) + 1 > \frac{(1-x)f(x)}{x^2} \cdot \frac{x}{1-x} = \frac{f(x)}{x}$$

so that $f(x) < \frac{x}{1-x}$, a contradiction. Hence $f(x) \leq \frac{x}{1-x}$ for all $x \in [0,1)$. A similar argument using (2.7.1) and (2.7.3) yields that $f(x) \geq \frac{x}{1+x}$ for all $x \in (-1,0]$.

Moreover, by Lemma 2.7.1 (3) and the two inequalities just proved,

$$\frac{f''(0)}{2} \leq \lim_{x\searrow 0} \frac{\frac{x}{1-x} - x}{x^2} = \lim_{x\searrow 0} \frac{1}{1-x} = 1$$

and

$$\frac{f''(0)}{2} \geq \lim_{x\nearrow 0} \frac{\frac{x}{1+x} - x}{x^2} = \lim_{x\searrow 0} \frac{-1}{1+x} = -1$$

so that $|f''(0)| \leq 2$. □

Lemma 2.7.3. *The set $\mathcal{K}$ is convex and compact if it is considered as a subset of the topological vector space consisting of real functions on $(-1,1)$ with the locally convex topology of pointwise convergence.*

Proof. It is obvious that $\mathcal{K}$ is convex. Since $\{f(x) : f \in \mathcal{K}\}$ is bounded for each $x \in (-1,1)$ thanks to Lemma 2.7.2, it follows that $\mathcal{K}$ is relatively compact. To prove that $\mathcal{K}$ is closed, let $\{f_i\}$ be a net in $\mathcal{K}$ converging to a function f on $(-1,1)$. Then it is clear that f is operator monotone on $(-1,1)$ and $f(0) = 0$. By Lemma 2.7.1 (2), $(1 + \frac{1}{x})f_i(x)$ is operator monotone on $(-1,1)$ for every i. Since $\lim_{x\to 0}(1 + \frac{1}{x})f_i(x) = f_i'(0) = 1$, we thus have

$$\left(1 - \frac{1}{x}\right)f_i(-x) \leq 1 \leq \left(1 + \frac{1}{x}\right)f_i(x), \qquad x \in (0,1).$$

Therefore,

$$\left(1 - \frac{1}{x}\right)f(-x) \leq 1 \leq \left(1 + \frac{1}{x}\right)f(x), \qquad x \in (0,1).$$

Since f is C^1 on $(-1,1)$ by Theorem 2.4.1, the above inequalities yield $f'(0) = 1$. □

Lemma 2.7.4. *The extreme points of $\mathcal{K}$ have the form*

$$f(x) = \frac{x}{1-\lambda x}, \quad \textit{where} \quad \lambda = \frac{f''(0)}{2}.$$

Proof. Let f be an extreme point of $\mathcal{K}$. For each $\alpha \in (-1,1)$ define

$$g_\alpha(x) := \left(1 + \frac{\alpha}{x}\right)f(x) - \alpha, \qquad x \in (-1,1).$$

By Lemma 2.7.1 (2), g_α is operator monotone on $(-1,1)$. Notice

$$g_\alpha(0) = f(0) + \alpha f'(0) - \alpha = 0$$

and

$$g_\alpha'(0) = \lim_{x\to 0} \frac{(1 + \frac{\alpha}{x})f(x) - \alpha}{x} = f'(0) + \alpha \lim_{x\to 0} \frac{f(x) - f'(0)x}{x^2} = 1 + \frac{1}{2}\alpha f''(0)$$

by Lemma 2.7.1 (3). Since $1 + \frac{1}{2}\alpha f''(0) > 0$ by Lemma 2.7.2, the function

$$h_\alpha(x) := \frac{(1 + \frac{\alpha}{x})f(x) - \alpha}{1 + \frac{1}{2}\alpha f''(0)}$$

is in $\mathcal{K}$. Since

$$f = \frac{1}{2}\left(1 + \frac{1}{2}\alpha f''(0)\right)h_\alpha + \frac{1}{2}\left(1 - \frac{1}{2}\alpha f''(0)\right)h_{-\alpha},$$

the extremality of f implies that $f = h_\alpha$ so that

$$\left(1 + \frac{1}{2}\alpha f''(0)\right)f(x) = \left(1 + \frac{\alpha}{x}\right)f(x) - \alpha$$

for all $\alpha \in (-1, 1)$. This immediately implies that $f(x) = x/(1 - \frac{1}{2}f''(0)x)$. □

Theorem 2.7.5. *Let f be a non-constant operator monotone function on $(-1, 1)$, Then there exists a unique probability Borel measure μ on $[-1, 1]$ such that*

$$f(x) = f(0) + f'(0)\int_{-1}^{1} \frac{x}{1 - \lambda x}\, d\mu(\lambda), \qquad x \in (-1, 1). \tag{2.7.4}$$

Proof. Since $f'(0) > 0$ thanks to Theorem 2.4.1, it is enough to assume that $f \in \mathcal{K}$ by considering $(f - f(0))/f'(0)$. Let $\phi_\lambda(x) := x/(1 - \lambda x)$ for $\lambda \in [-1, 1]$. By Lemmas 2.7.3 and 2.7.4, the Krein–Milman theorem says that $\mathcal{K}$ is the closed convex hull of $\{\phi_\lambda : \lambda \in [-1, 1]\}$. Hence there exists a net $\{f_i\}$ in the convex hull of $\{\phi_\lambda : \lambda \in [-1, 1]\}$ such that $f_i(x) \to f(x)$ for all $x \in (-1, 1)$. Each f_i is written as $f_i(x) = \int_{-1}^{1} \phi_\lambda(x)\, d\mu_i(\lambda)$ with a probability measure μ_i on $[-1, 1]$ with finite support. Note that the set $\mathcal{M}_1([-1, 1])$ of probability Borel measures on $[-1, 1]$ is compact in the weak* topology when considered as a subset of the dual Banach space of $C([-1, 1])$. Taking a subnet we may assume that μ_i converges in the weak* topology to some $\mu \in \mathcal{M}_1([-1, 1])$. For each $x \in (-1, 1)$, since $\phi_\lambda(x)$ is continuous in $\lambda \in [-1, 1]$, we have

$$f(x) = \lim_i f_i(x) = \lim_i \int_{-1}^{1} \phi_\lambda(x)\, d\mu_i(\lambda) = \int_{-1}^{1} \phi_\lambda(x)\, d\mu(\lambda).$$

To prove the uniqueness of the representing measure μ, let $\mu_1, \mu_2 \in \mathcal{M}([-1, 1])$ be such that

$$f(x) = \int_{-1}^{1} \phi_\lambda(x)\, d\mu_1(\lambda) = \int_{-1}^{1} \phi_\lambda(x)\, d\mu_2(\lambda), \qquad x \in (-1, 1).$$

Since $\phi_\lambda(x) = \sum_{k=0}^{\infty} x^{k+1}\lambda^k$ is uniformly convergent in $\lambda \in [-1, 1]$ for any $x \in (-1, 1)$ fixed, it follows that

$$\sum_{k=0}^{\infty} x^{k+1} \int_{-1}^{1} \lambda^k\, d\mu_1(\lambda) = \sum_{k=0}^{\infty} x^{k+1} \int_{-1}^{1} \lambda^k\, d\mu_2(\lambda), \qquad x \in (-1, 1).$$

Hence $\int_{-1}^{1} \lambda^k\, d\mu_1(\lambda) = \int_{-1}^{1} \lambda^k\, d\mu_2(\lambda)$ for all $k = 0, 1, 2, \ldots$, which implies that $\mu_1 = \mu_2$. □

The integral representation of the above theorem is an example of Choquet's theorem while we proved it in a direct way. The uniqueness of the representing measure μ shows that $\{\phi_\lambda : \lambda \in [-1, 1]\}$ is actually the set of extreme points of $\mathcal{K}$. Since the pointwise convergence topology on $\{\phi_\lambda : \lambda \in [-1, 1]\}$ agrees with the usual topology on $[-1, 1]$, we see that $\mathcal{K}$ is a so-called Bauer simplex (see [67]).

Theorem 2.7.6. *Let f be a non-linear operator convex function on $(-1, 1)$. Then there exists a unique probability Borel measure μ on $[-1, 1]$ such that*

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}\int_{-1}^{1} \frac{x^2}{1 - \lambda x}\, d\mu(\lambda), \qquad x \in (-1, 1).$$

Proof. It is enough to assume that $f(0) = f'(0) = 0$ by considering $f(x) - f(0) - f'(0)x$. By Corollary 2.4.6, $g(x) := f^{[1]}(0, x) = f(x)/x$ is a non-constant operator monotone function on $(-1, 1)$. Hence by Theorem 2.7.5 there exists a probability Borel measure μ on $[-1, 1]$ such that

$$g(x) = g'(0)\int_{-1}^{1} \frac{x}{1 - \lambda x}\, d\mu(\lambda), \qquad x \in (-1, 1).$$

Since $g'(0) = f''(0)/2$ is easily seen from Theorem 2.4.2, we have

$$f(x) = \frac{f''(0)}{2}\int_{-1}^{1} \frac{x^2}{1 - \lambda x}\, d\mu(\lambda), \qquad x \in (-1, 1).$$

Moreover, the uniqueness of μ follows from that of the representing measure for g. □

Finally, we establish the equivalence between the operator monotone functions on (a, b) and the Pick functions in $\mathcal{P}(a, b)$ in the following way.

Theorem 2.7.7. *Let $-\infty \le a < b \le \infty$ and f be a real-valued function on (a,b). Then f is operator monotone on (a,b) if and only if $f \in \mathcal{P}(a,b)$.*

Proof. The "if" part was shown after Theorem 2.6.3 in the preceding section. To prove the "only if", it is enough to assume that (a,b) is a finite open interval. In fact, if the assertion holds in this case, then for every finite interval (c,d) included in (a,b), $f|_{(c,d)}$ is operator monotone and so $f \in \mathcal{P}(c,d)$. Hence $f \in \mathcal{P}(a,b)$ follows by letting $c \searrow a$ and $d \nearrow b$. Moreover, when (a,b) is a finite interval, f is transformed into an operator monotone function on $(-1,1)$ and $\mathcal{P}(a,b)$ is transformed into $\mathcal{P}(-1,1)$ via the affine function mentioned in the beginning of this section. So it suffices to prove the "only if" part when $(a,b)=(-1,1)$. If f is a non-constant operator monotone function on $(-1,1)$, then by using the integral representation (2.7.4) of Theorem 2.7.5 one can define an analytic continuation of f across $(-1,1)$ by

$$f(z) = f(0) + f'(0)\int_{-1}^{1} \frac{z}{1-\lambda z}\, d\mu(\lambda), \qquad z \in \mathbb{C}^+ \cup \mathbb{C}^-.$$

Since

$$\operatorname{Im} f(z) = f'(0)\int_{-1}^{1} \frac{\operatorname{Im} z}{|1-\lambda z|^2}\, d\mu(\lambda),$$

it follows that f maps $\mathbb{C}^+$ into itself. Hence $f \in \mathcal{P}(-1,1)$. □

Since

$$\frac{x}{1-\lambda x} = \frac{\lambda + x}{1-\lambda x}\cdot\frac{1}{\lambda^2+1} - \frac{\lambda}{\lambda^2+1},$$

one can substitute u^{-1} for $\lambda \neq 0$ in the integral of (2.7.4) so that (2.7.4) is rewritten as

$$\begin{aligned} f(x) &= f(0) - f'(0)\int_{-1}^{1} \frac{\lambda}{\lambda^2+1}\, d\mu(\lambda) + f'(0)\mu(\{0\})x \\ &\quad + f'(0)\int_{[-1,0)\cup(0,1]} \frac{\lambda + x}{1-\lambda x}\cdot\frac{1}{\lambda^2+1}\, d\mu(\lambda) \\ &= \alpha + \beta x + f'(0)\int_{\mathbb{R}\setminus(-1,1)} \frac{1+ux}{u-x}\cdot\frac{u^2}{1+u^2}\, d\mu(u^{-1}) \\ &= \alpha + \beta x + \int_{\mathbb{R}\setminus(-1,1)} \frac{1+ux}{u-x}\, d\nu(u), \end{aligned}$$

where

$$\alpha := f(0) - f'(0)\int_{-1}^{1} \frac{\lambda}{\lambda^2+1}\, d\mu(\lambda), \quad \beta := f'(0)\mu(\{0\}), \quad d\nu(u) := \frac{f'(0)u^2}{1+u^2}\, d\mu(u^{-1}).$$

In this way, the integral representation (2.7.4) can be transformed into the form (2.6.1) of Nevanlinna's theorem. This may be an alternative proof of Theorem 2.7.7.

The next corollary improves the last statement of Corollary 2.4.6.

Corollary 2.7.8. *Let $-\infty \le a < b \le \infty$ and f be a real-valued function on (a,b). Then the following conditions are equivalent:*

(i) *f is operator convex;*
(ii) *f is C^1 and $f^{[1]}(s,\cdot)$ is operator monotone on (a,b) for every $s \in (a,b)$;*
(iii) *$f^{[1]}(s,\cdot)$ is operator monotone on (a,b) for some $s \in (a,b)$ (with continuation of value at s).*

Consequently, if g is a real-valued function on (a,b) and $f(x) := (x-s)g(x)$ for any $s \in (a,b)$, then f is operator convex on (a,b) if and only if g is operator monotone on (a,b).

Proof. By the same argument as in the proof of Theorem 2.7.7, we may assume that $(a,b)=(-1,1)$. (ii) $\Rightarrow$ (iii) is trivial. Although (i) $\Rightarrow$ (ii) is included in Corollary 2.4.6, we prove it below based on the integral representation of Theorem 2.7.6.

(i) $\Rightarrow$ (ii). By Theorem 2.7.6, f admits a representation

$$f(x) = \alpha + \beta x + \int_{-1}^{1} \frac{x^2}{1-\lambda x}\, d\nu(\lambda), \qquad x \in (-1,1),$$

where $\alpha, \beta \in \mathbb{R}$ and ν is a positive finite Borel measure on $[-1,1]$. For any $s \in (-1,1)$ we write

$$f^{[1]}(s,x) = \frac{f(x)-f(s)}{x-s} = \beta + \int_{-1}^{1} \frac{1}{x-s}\left(\frac{x^2}{1-\lambda x} - \frac{s^2}{1-\lambda s}\right) d\nu(\lambda)$$
$$= \beta + \int_{-1}^{1} \frac{(1-\lambda s)x + s}{(1-\lambda s)(1-\lambda x)}\, d\nu(\lambda)$$
$$= \beta + \int_{-1}^{1} \frac{x + \frac{s}{1-\lambda s}}{1-\lambda x}\, d\nu(\lambda).$$

Hence it suffices to show that $(x + \frac{s}{1-\lambda s})/(1-\lambda x)$ is operator monotone on $(-1,1)$ for every $\lambda \in [-1,1]$. This is clear when $\lambda = 0$. When $\lambda \neq 0$, we have

$$\frac{x + \frac{s}{1-\lambda s}}{1-\lambda x} = -\frac{1}{\lambda} + \frac{1}{\lambda^2(1-\lambda s)} \cdot \frac{1}{\lambda^{-1} - x},$$

which is operator monotone on $(-1,1)$ thanks to Example 2.5.9 (2) and $1/\lambda^2(1-\lambda s) > 0$.

(iii) $\Rightarrow$ (i). By Theorem 2.7.5, $f^{[1]}(s,\cdot)$ admits a representation

$$f^{[1]}(s,x) = \alpha + \int_{-1}^{1} \frac{x}{1-\lambda x}\, d\nu(x), \qquad x \in (-1,1),$$

where $\alpha \in \mathbb{R}$ and ν is a positive finite Borel measure on $[-1,1]$. This implies that

$$f(x) = f(s) + \alpha(x-s) + \int_{-1}^{1} \frac{x(x-s)}{1-\lambda x}\, d\nu(\lambda).$$

Hence it suffices to show that $(x^2 - sx)/(1-\lambda x)$ is operator convex on $(-1,1)$ for every $\lambda \in [-1,1]$. This is clear when $\lambda = 0$. When $\lambda \neq 0$,

$$\lambda^2\, \frac{x^2 - sx}{1-\lambda x} = -\lambda x + \lambda s - 1 + \frac{1-\lambda s}{1-\lambda x}.$$

From Example 2.5.9 (4) it is immediate to see that $(1-\lambda x)^{-1}$ is operator convex on $(-1,1)$. Hence thanks to $1-\lambda s > 0$, $(x^2 - sx)/(1-\lambda x)$ is operator convex on $(-1,1)$.

The latter statement is just a rewriting of the proved equivalence. □

Corollary 2.7.9. *Any operator monotone function on the whole real line is a linear function $f(x) = \alpha + \beta x$ with $\alpha \in \mathbb{R}$ and $\beta \geq 0$. Any operator convex function on the whole real line is a quadratic function $f(x) = \alpha + \beta x + \gamma x^2$ with $\alpha, \beta \in \mathbb{R}$ and $\gamma \geq 0$.*

Proof. Assume that f is operator monotone on $(-\infty,\infty)$. Since $f \in \mathcal{P}(-\infty,\infty)$ thanks to Theorem 2.7.7, Theorem 2.6.3 implies that $f(z) = \alpha + \beta z$ with $\alpha \in \mathbb{R}$ and $\beta \geq 0$. Hence $f(x) = \alpha + \beta x$. Next assume that f is operator convex on $(-\infty,\infty)$. By Corollary 2.4.6, $(f(x) - f(0))/x$ is operator monotone on $(-\infty,\infty)$ and hence $(f(x)-f(0))/x = \beta + \gamma x$ with $\beta \in \mathbb{R}$ and $\gamma \geq 0$. Therefore, $f(x) = \alpha + \beta x + \gamma x^2$ with $\alpha = f(0)$. □

Exercise 2.7.10. When $p < -1$, show that the function f on $(0,\infty)$ defined by $f(x) := (x^p - 1)/(x-1)$, $x \neq 1$, and $f(1) := p$ cannot be analytically continued across $(0,\infty)$ to $\mathbb{C}^+$ in such a way that $f(\mathbb{C}^+) \subset \mathbb{C}^+$. For this, take account of the fact that for small $r > 0$ the argument of $f(z) = (z^p - 1)/(z-1)$, $z = re^{i\theta}$, nearly behaves as $-z^p$. By Corollary 2.4.6 and Theorem 2.7.7, this proves that x^p is not operator convex on $(0,\infty)$ when $p < -1$, settling the remaining part of Example 2.5.9 (4).

Finally, we transform the integral expression (2.7.4) for operator monotone functions on $(-1,1)$ into the following expression on $[0,\infty)$, which will play an important role in the next chapter.

Theorem 2.7.11. *Let f be a continuous and nonnegative function on $[0,\infty)$. Then f is operator monotone if and only if there exists a positive finite Borel measure m on $[0,\infty]$ such that*

$$f(t) = \int_{[0,\infty]} \frac{t(1+\lambda)}{t+\lambda}\, dm(\lambda), \qquad t \in [0,\infty),$$

where $t(1+\lambda)/(t+\lambda)$ is 1 if $\lambda = 0$ and t if $\lambda = \infty$. In this case, the measure m is unique, and if $a := m(\{0\})$ and $b := m(\{\infty\})$ then

$$f(t) = a + bt + \int_{(0,\infty)} \frac{t(1+\lambda)}{t+\lambda}\, dm(\lambda), \qquad t \in [0,\infty). \tag{2.7.5}$$

Also, $a = f(0)$ and $b = \lim_{t\to\infty} f(t)/t$.

Moreover, a continuous real-function function f on $[0,\infty)$ is operator monotone if and only if there exist a $b \geq 0$ and a positive finite Borel measure m on $(0,\infty)$ such that

$$f(t) = f(0) + bt + \int_{(0,\infty)} \frac{t(1+\lambda)}{t+\lambda}\, dm(\lambda), \qquad t \in [0,\infty).$$

Proof. The "if" part is immediately seen since, for every $\lambda \in [0,\infty)$,

$$\frac{t(1+\lambda)}{t+\lambda} = 1 + \lambda - \frac{\lambda(1+\lambda)}{t+\lambda}$$

is operator monotone on $[0,\infty)$. Conversely, assume that $f \geq 0$ is continuous and operator monotone on $[0,\infty)$. Transform $f(t)$ on $(0,\infty)$ to an operator monotone function $g(x) := f(\psi(x))$ on $(-1,1)$ by

$$t = \psi(x) := \frac{1+x}{1-x} = -1 + \frac{2}{1-x} : (-1,1) \to (0,\infty).$$

Theorem 2.7.5 implies that there exists a probability Borel measure μ on $[-1,1]$ such that

$$g(x) = g(0) + g'(0) \int_{[-1,1]} \frac{x}{1-\lambda x}\, d\mu(\lambda), \qquad x \in (-1,1).$$

Since $g(-1) = \lim_{x\to -1+0} g(x) = f(0) \geq 0$, we notice that

$$\int_{[-1,1]} \frac{1}{1+\lambda}\, d\mu(\lambda) < +\infty,$$

in particular $\mu(\{-1\}) = 0$, and hence

$$g(x) - g(-1) = g'(0) \int_{(-1,1]} \frac{1+x}{(1-\lambda x)(1+\lambda)}\, d\mu(\lambda).$$

Transforming this to the expression of $f(t)$ by $x = \psi^{-1}(t)$ and $\lambda = \psi^{-1}(\zeta)$ and introducing the measure m on $(0,\infty]$ by

$$m := \tilde{\mu} \circ \psi^{-1}, \quad \text{where} \quad d\tilde{\mu}(\lambda) := \frac{g'(0)}{1+\lambda}\, d\mu(\lambda),$$

we have

$$f(t) - f(0) = \int_{(0,\infty]} \frac{t(1+\zeta)}{t+\zeta}\, dm(\zeta), \qquad t \in [0,\infty).$$

Adding the mass $f(0)\delta_0$ to m we have

$$f(t) = \int_{[0,\infty]} \frac{t(1+\zeta)}{t+\zeta}\, dm(\zeta), \qquad t \in [0,\infty).$$

The uniqueness of the measure m follows from that of μ in Theorem 2.7.5, and the remaining assertions are easily verified. Finally, the last assertion is immediately seen by applying the above case to $f - f(0)$. □

For example, from the integral expression

$$\log t = \int_0^\infty \left(\frac{1}{1+\lambda} - \frac{1}{t+\lambda} \right) d\lambda,$$

the operator monotone function $\log(1+t)$ on $[0,\infty)$ has the expression

$$\log(1+t) = \int_1^\infty \frac{t}{\lambda(t+\lambda)}\, d\lambda$$

and the representing measure in Theorem 2.7.11 is $\chi_{[1,\infty)} \frac{1}{\lambda(1+\lambda)}\, d\lambda$. For $0 < p < 1$ the operator monotone function t^p on $[0,\infty)$ has the expression

$$t^p = \frac{\sin p\pi}{\pi} \int_0^\infty \frac{t\lambda^{p-1}}{t+\lambda}\, d\lambda$$

(see Exercise 2.6.6 (b)) and the representing measure is $\frac{\sin p\pi}{\pi} \cdot \frac{\lambda^{p-1}}{1+\lambda}\, d\lambda$.

2.8 Bhatia and Sano's characterization of operator convex functions

Concerning matrix/operator convex functions, after showing Kraus' characterization in terms of the second divided difference matrices in Section 2.4, we present characterizations due to Hansen and Pedersen for operator convex functions on $(0,\infty)$ in Section 2.5. In this section we present different characterizations for those functions in terms of the first divided difference matrices, which were recently obtained by Bhatia and Sano [20]. We begin with

Definition 2.8.1. We write $\mathbb{C}_0^n$ for the subspace (of codimension 1) of $\mathbb{C}^n$ consisting of $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{C}^n$ such that $\sum_{i=1}^n \xi_n = 0$. An $n \times n$ Hermitian matrix A is said to be *conditionally positive definite* (c.p.d. for short) if $\langle x, Ax \rangle \geq 0$ for all $x \in \mathbb{C}_0^n$, and also *conditionally negative definite* (c.n.d. for short) if $\langle x, Ax \rangle \leq 0$ for all $x \in \mathbb{C}_0^n$. The $n \times n$ matrix of all entries equal to 1 is denoted by J_n. Obviously, J_n is a positive semidefinite matrix (of rank 1), and $x \in \mathbb{C}^n$ belongs to $\mathbb{C}_0^n$ if and only if $J_n x = 0$.

Lemma 2.8.2. *If $A = [a_{ij}]_{i,j=1}^n \in \mathbb{M}_n^{sa}$ is c.p.d., then the $(n-1) \times (n-1)$ matrix whose (i,j)-entry is*

$$a_{ij} - a_{in} - a_{nj} + a_{nn}, \qquad i, j = 1, \ldots, n-1,$$

is positive semidefinite.

Proof. Write $B := [a_{ij} - a_{in} - a_{nj} + a_{nn}]_{i,j=1}^{n-1}$. For every $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{n-1} \end{bmatrix} \in \mathbb{C}^{n-1}$ let $\tilde{x} := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{n-1} \\ \xi_n \end{bmatrix} \in \mathbb{C}_0^n$ with $\xi_n := -\sum_{i=1}^{n-1} \xi_i$. Then

$$\begin{aligned} \langle x, Bx \rangle &= \sum_{i,j=1}^{n-1} (a_{ij} - a_{in} - a_{nj} + a_{nn}) \overline{\xi_i} \xi_j \\ &= \sum_{i,j=1}^{n-1} a_{ij} \overline{\xi_i} \xi_j - \sum_{i=1}^{n-1} a_{in} \overline{\xi_i} \left(\sum_{j=1}^{n-1} \xi_j \right) - \sum_{j=1}^{n-1} a_{nj} \left(\sum_{i=1}^{n-1} \overline{\xi_i} \right) \xi_j + a_{nn} \left(\sum_{i=1}^{n-1} \overline{\xi_i} \right) \left(\sum_{j=1}^{n-1} \xi_j \right) \\ &= \sum_{i,j=1}^{n} a_{ij} \overline{\xi_i} \xi_j = \langle \tilde{x}, A \tilde{x} \rangle \geq 0 \end{aligned}$$

so that B is positive semidefinite. $\square$

The next theorem shows Bhatia and Sano's characterizations in [20] for operator convex functions on $(0, \infty)$ with some improvements. A similar improvement was also obtained in [76] by a different method.

Theorem 2.8.3. *Let f be a real C^1-function on $(0, \infty)$, and $g(t) := t f(t)$ for $t \in (0, \infty)$. Then the following conditions are equivalent:*

(i) *f is operator convex;*
(ii) *$\liminf_{t \to \infty} f(t)/t > -\infty$ and $[f^{[1]}(t_i, t_j)]_{i,j=1}^n$ is c.n.d. for all $n \in \mathbb{N}$ and all $t_1, \ldots, t_n \in (0, \infty)$;*
(iii) *$\limsup_{t \searrow 0} g(t) \geq 0$ and $[g^{[1]}(t_i, t_j)]_{i,j=1}^n$ is c.p.d. for all $n \in \mathbb{N}$ and all $t_1, \ldots, t_n \in (0, \infty)$.*

Proof. (i) $\Rightarrow$ (ii). First, the condition $\liminf_{t \to \infty} f(t)/t > -\infty$ is trivially satisfied as long as f is a convex function on $(0, \infty)$. For any $\varepsilon > 0$ define

$$h_\varepsilon(t) := f(t + \varepsilon) - f(\varepsilon) - f'(\varepsilon) t, \qquad t \geq 0,$$

which is operator convex and nonnegative on $[0, \infty)$ with $h_\varepsilon(+0) = h_\varepsilon(0) = 0$ and $h'_\varepsilon(0) = 0$. Then Theorem 2.5.2 implies that $h_\varepsilon(t)/t$ is operator monotone on $(0, \infty)$ with $\lim_{t \searrow 0} h_\varepsilon(t)/t = h'_\varepsilon(0) = 0$. Hence by Theorem 2.7.11, the function $h_\varepsilon(t)/t$ is represented as

$$\frac{h_\varepsilon(t)}{t} = ct + \int_{(0,\infty)} \frac{t(1+\lambda)}{t+\lambda} \, dm(\lambda),$$

where $c \geq 0$ and m is a positive finite measure on $(0, \infty)$. Therefore, we have

$$f_\varepsilon(t) := f(t + \varepsilon) = a + bt + ct^2 + \int_{(0,\infty)} \frac{t^2(1+\lambda)}{t+\lambda} \, dm(\lambda),$$

where $a := f(\varepsilon)$ and $b := f'(\varepsilon)$. Letting $\phi_\lambda(t) := t^2/(t+\lambda)$ for $\lambda \in (0, \infty)$ and $t \in [0, \infty)$, one can write

$$f_\varepsilon^{[1]}(s, t) = b + c(s+t) + \int_{(0,\infty)} \phi_\lambda^{[1]}(s, t)(1+\lambda) \, dm(\lambda), \qquad s, t \in [0, \infty).$$

Since

$$\phi_\lambda(s,t) = \frac{1}{s-t}\left(\frac{s^2}{s+\lambda} - \frac{t^2}{t+\lambda}\right) = \frac{st+\lambda(s+t)}{(s+\lambda)(t+\lambda)}$$
$$= 1 - \frac{\lambda^2}{(s+\lambda)(t+\lambda)}, \qquad s,t \in [0,\infty),$$

we have, for every $t_1,\dots,t_n \in (0,\infty)$,

$$[f_\varepsilon^{[1]}(t_i,t_j)]_{i,j=1}^n = bJ_n + c(DJ_n + J_nD) + \int_{(0,\infty)} (J_n - \lambda^2 D_\lambda J_n D_\lambda)(1+\lambda)\,dm(\lambda),$$

where

$$D := \mathrm{Diag}(t_1,\dots,t_n), \qquad D_\lambda := \mathrm{Diag}\left(\frac{1}{t_1+\lambda},\dots,\frac{1}{t_n+\lambda}\right).$$

For every $x \in \mathbb{C}_0^n$, since $J_nx = 0$ and $D_\lambda J_n D_\lambda \geq 0$, we obtain

$$\langle x, [f_\varepsilon^{[1]}(t_i,t_j)]x\rangle = -\int_{(0,\infty)} \lambda^2\langle x, D_\lambda J_n D_\lambda x\rangle(1+\lambda)\,dm(\lambda) \leq 0,$$

which shows that $[f_\varepsilon^{[1]}(t_i,t_j)]$ is c.n.d. Since $f_\varepsilon^{[1]}(t_i,t_j) = f^{[1]}(t_i+\varepsilon, t_j+\varepsilon)$ and $\varepsilon > 0$ is arbitrary, (ii) holds.

(ii) $\Rightarrow$ (i). For any $\varepsilon > 0$, since

$$\lim_{t\searrow 0} \frac{f(t+\varepsilon)-f(\varepsilon)}{t} = f'(\varepsilon)$$

and

$$\liminf_{t\to\infty} \frac{f(t+\varepsilon)-f(\varepsilon)}{t} > -\infty$$

thanks to the condition $\liminf_{t\to\infty} f(t)/t > -\infty$, one can see that

$$\inf_{t\in(0,\infty)} \frac{f(t+\varepsilon)-f(\varepsilon)}{t} > -\infty.$$

So choose a $\gamma_\varepsilon \in \mathbb{R}$ smaller than the above infimum, and define

$$f_\varepsilon(t) := f(t+\varepsilon), \qquad h_\varepsilon(t) := f(t+\varepsilon) - f(\varepsilon) - \gamma_\varepsilon t, \qquad t \in [0,\infty),$$

so that $h_\varepsilon(t) > 0$ for all $t \in (0,\infty)$. For every $n \in \mathbb{N}$ and every $t_1,\dots,t_n \in [0,\infty)$, since

$$[h_\varepsilon^{[1]}(t_i,t_j)]_{i,j=1}^n = [f_\varepsilon^{[1]}(t_i,t_j)]_{i,j=1}^n - \gamma_\varepsilon J_n = [f^{[1]}(t_i+\varepsilon, t_j+\varepsilon)]_{i,j=1}^n - \gamma_\varepsilon J_n,$$

it follows that $[h_\varepsilon^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. Letting $t_1,\dots,t_{n-1} > 0$ and $t_n = 0$, we see by Lemma 2.8.2 that

$$-\left[h_\varepsilon^{[1]}(t_i,t_j) - \frac{h_\varepsilon(t_i)}{t_i} - \frac{h_\varepsilon(t_j)}{t_j} + h_\varepsilon'(0)\right]_{i,j=1}^{n-1}$$

is positive semidefinite. One can compute the above (i,j)-entry as follows:

$$\frac{h_\varepsilon(t_i)-h_\varepsilon(t_j)}{t_i-t_j} - \frac{h_\varepsilon(t_i)}{t_i} - \frac{h_\varepsilon(t_j)}{t_j} + f'(\varepsilon) - \gamma_\varepsilon$$
$$= \frac{t_j^2 h_\varepsilon(t_i) - t_i^2 h_\varepsilon(t_j)}{t_i(t_i-t_j)t_j} + f'(\varepsilon) - \gamma_\varepsilon$$
$$= -\frac{h_\varepsilon(t_i)}{t_i}\cdot\frac{\frac{t_i^2}{h_\varepsilon(t_i)} - \frac{t_j^2}{h_\varepsilon(t_j)}}{t_i-t_j}\cdot\frac{h_\varepsilon(t_j)}{t_j} + f'(\varepsilon) - \gamma_\varepsilon,$$

noting that $h_\varepsilon(t) > 0$ for $t > 0$. Hence it follows that

$$\left[\left(\frac{t^2}{h_\varepsilon(t)}\right)^{[1]}(t_i,t_j)\right]_{i,j=1}^{n-1} - (f'(\varepsilon)-\gamma_\varepsilon)DJ_{n-1}D$$

is positive semidefinite, where

$$D := \mathrm{Diag}\left(\frac{t_1}{h_\varepsilon(t_1)},\dots,\frac{t_n}{h_\varepsilon(t_n)}\right).$$

Since $f'(\varepsilon) - \gamma_\varepsilon > 0$ by the choice of γ_ε,

$$\left[\left(\frac{t^2}{h_\varepsilon(t)}\right)^{[1]}(t_i,t_j)\right]_{i,j=1}^{n-1}$$

is positive semidefinite. This implies by Theorem 2.4.3 that $t^2/h_\varepsilon(t)$ is operator monotone on $(0,\infty)$. Hence by Corollary 2.5.6, $h_\varepsilon(t)/t = t/(t^2/h_\varepsilon(t))$ is operator monotone on $(0,\infty)$ again. Finally, since $h_\varepsilon(0) = 0$, Theorem 2.5.2 implies that $h_\varepsilon(t)$ is operator convex on $[0,\infty)$. Hence so is $f(t+\varepsilon) = h_\varepsilon(t) + f(\varepsilon) + \gamma_\varepsilon t$. Since $\varepsilon > 0$ is arbitrary, f is operator convex on $(0,\infty)$.

(i) $\Rightarrow$ (iii). The condition $\limsup_{t\searrow 0} g(t) \ge 0$ is obvious since $f(+0) > -\infty$ as long as f is a convex function on $(0,\infty)$. With the same $h_\varepsilon(t)$ as in the proof of (i) $\Rightarrow$ (ii), we have

$$g_\varepsilon(t) := tf(t+\varepsilon) = at + bt^2 + ct^3 + \int_{(0,\infty)} \frac{t^3(1+\lambda)}{t+\lambda}\, dm(\lambda),$$

where a, b, $c \ge 0$, and m are as in the proof of (i) $\Rightarrow$ (ii). Letting $\psi_\lambda(t) := t^3/(t+\lambda)$, one can write

$$g_\varepsilon^{[1]}(s,t) = a + b(s+t) + c(s^2+st+t^2) + \int_{(0,\infty)} \psi_\lambda^{[1]}(s,t)(1+\lambda)\, dm(\lambda), \qquad s,t \in [0,\infty).$$

Since

$$\begin{aligned}\psi_\lambda^{[1]}(s,t) &= \frac{1}{s-t}\left(\frac{s^3}{s+\lambda} - \frac{t^3}{t+\lambda}\right) = \frac{st(s+t) + \lambda(s^2+st+t^2)}{(s+\lambda)(t+\lambda)} \\ &= s + t - \lambda + \frac{\lambda^3}{(s+\lambda)(t+\lambda)}, \qquad s,t \in [0,\infty),\end{aligned}$$

we have, for every $t_1,\dots,t_n \in (0,\infty)$,

$$\begin{aligned}[g_\varepsilon^{[1]}(t_i,t_j)]_{i,j=1}^n &= aJ_n + b(DJ_n + J_nD) + c(D^2J_n + DJ_nD + J_nD^2) \\ &\quad + \int_{(0,\infty)} (DJ_n + J_nD - \lambda J_n + \lambda^3 D_\lambda J_n D_\lambda)(1+\lambda)\, dm(\lambda),\end{aligned}$$

where D and D_λ are as in the proof of (i) $\Rightarrow$ (ii). For every $x \in \mathbb{C}_0^n$, since $J_nx = 0$, $DJ_nD \ge 0$ and $D_\lambda J_n D_\lambda \ge 0$, we have

$$\langle x, [g_\varepsilon^{[1]}(t_i,t_j)]x\rangle = c\langle x, DJ_nDx\rangle + \int_{(0,\infty)} \lambda^3\langle x, D_\lambda J_n D_\lambda x\rangle(1+\lambda)\, dm(\lambda) \ge 0$$

so that $[g_\varepsilon^{[1]}(t_i,t_j)]$ is c.p.d. Since $\lim_{\varepsilon\searrow 0} g_\varepsilon(t) = g(t)$ and $\lim_{\varepsilon\searrow 0} g_\varepsilon'(t) = g'(t)$ for all $t \in (0,\infty)$, it follows that $[g^{[1]}(t_i,t_j)]$ is c.p.d. and (iii) holds.

(iii) $\Rightarrow$ (i). Thanks to the condition $\limsup_{t\searrow 0} g(t) \ge 0$, one can choose a sequence $\varepsilon_k \searrow 0$ such that $g(\varepsilon_k) > 0$ for all k when $\limsup_{t\searrow 0} g(t) > 0$, or else $\lim_{k\to\infty} g(\varepsilon_k) = 0$ when $\limsup_{t\searrow 0} g(t) = 0$. Define

$$h_k(t) := g(t+\varepsilon_k) - g(\varepsilon_k) - g'(\varepsilon_k)t, \qquad t \in [0,\infty).$$

For every $n \in \mathbb{N}$ and every $t_1,\dots,t_n \in [0,\infty)$, since

$$[h_k^{[1]}(t_i,t_j)]_{i,j=1}^n = [g^{[1]}(t_i+\varepsilon_k, t_j+\varepsilon_k)]_{i,j=1}^n - g'(\varepsilon_k)J_n,$$

it follows that $[h_k^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. Now let $t_1,\dots,t_{n-1} > 0$ and $t_n = 0$. Since $h_k'(0) = 0$, we see by Lemma 2.8.2 that

$$\left[h_k^{[1]}(t_i,t_j) - \frac{h_k(t_i)}{t_i} - \frac{h_k(t_j)}{t_j}\right]_{i,j=1}^{n-1}$$

is positive semidefinite. Since the above (i,j)-entry is equal to

$$\frac{h_k(t_i) - h_k(t_j)}{t_i - t_j} - \frac{h_k(t_i)}{t_i} - \frac{h_k(t_j)}{t_j} = t_i \cdot \frac{\frac{h_k(t_i)}{t_i^2} - \frac{h_k(t_j)}{t_j^2}}{t_i - t_j} \cdot t_j,$$

it follows that

$$\left[\left(\frac{h_k(t)}{t^2}\right)^{[1]}(t_i,t_j)\right]_{i,j=1}^{n-1}$$

is positive semidefinite. Therefore, $h_k(t)/t^2$ is operator monotone on $(0,\infty)$. Furthermore, since $\lim_{t\searrow 0} h_k(t)/t = h_k'(0) = 0$, Theorem 2.5.2 implies that $h_k(t)/t$ is operator convex on $(0,\infty)$. Noting that

$$\frac{h_k(t)}{t} = \frac{(t+\varepsilon_k)f(t+\varepsilon_k)}{t} - \frac{g(\varepsilon_k)}{t} - g'(\varepsilon_k),$$

we see that

$$f_k(t) := \frac{(t+\varepsilon_k)f(t+\varepsilon_k)}{t} - \frac{g(\varepsilon_k)}{t}$$

is operator convex on $(0,\infty)$. When $g(\varepsilon_k) > 0$ for all k, $g(\varepsilon_k)/t$ is operator convex and so

$$\frac{(t+\varepsilon_k)f(t+\varepsilon_k)}{t} = f_k(t) + \frac{g(\varepsilon_k)}{t}$$

is operator convex on $(0,\infty)$. Since $\lim_{k\to\infty}(t+\varepsilon_k)f(t+\varepsilon_k)/t = f(t)$ for all $t > 0$, f is operator convex on $(0,\infty)$. When $\lim_{k\to\infty} g(\varepsilon_k) = 0$, $\lim_{k\to\infty} f_k(t) = f(t)$ for all $t > 0$ so that f is operator convex on $(0,\infty)$ as well. □

Remark 2.8.4. The conditions $\limsup_{t\to\infty} f(t)/t > -\infty$ in (ii) and $\limsup_{t\searrow 0} g(t) \ge 0$ in (iii) are essential in Theorem 2.8.3, as seen from the discussions below. When $1 \le \alpha \le 2$, the function t^α is operator convex on $(0,\infty)$. Hence Theorem 2.8.3 implies that $[(t^{\alpha+1})^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. and so $[(-t^{\alpha+1})^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. for all $t_1,\dots,t_n \in (0,\infty)$. But $-t^{\alpha+1}$ is not operator convex on $(0,\infty)$. Note that $\lim_{t\to\infty}(-t^{\alpha+1}/t) = -\infty$.

Next, when $-1 \le \alpha \le 0$, the function t^α is operator convex on $(0,\infty)$. Hence Theorem 2.8.3 implies that $[(t^{\alpha})^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. and so $[(-t^{\alpha})^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. for all $t_1,\dots,t_n \in (0,\infty)$. But $-t^{\alpha-1}$ is not operator convex on $(0,\infty)$. Note that $\lim_{t\searrow 0} t(-t^{\alpha-1}) \le -1$.

The next corollary is Theorem 2.8.3 in a special situation.

Corollary 2.8.5. *Let f be a real C^1-function form $(0,\infty)$ into itself. Then the following conditions are equivalent:*

(i) *f is operator convex;*

(ii) *all divided difference matrices of f are c.n.d., i.e., $[f^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. for all $n \in \mathbb{N}$ and all $t_1,\dots,t_n \in (0,\infty)$;*

(iii) *all divided difference matrices of $tf(t)$ are c.p.d.*

Moreover, if the above conditions hold, then all divided difference matrices of $t/f(t)$ and of $f(t)/t^2$ are c.n.d.

Proof. Since the additional conditions in (ii) and (iii) of Theorem 2.8.3 are trivially satisfied in this special situation, Theorem 2.8.3 shows that (i)–(iii) are equivalent. When f is operator convex, then so are $f/f(t)$ and $f(t)/t^2$ by Corollary 2.5.6. Hence the last assertion follows. □

For instance, all divided difference matrices of the power function t^r on $(0,\infty)$ are c.n.d. for $-1 \le r \le 0$, positive semidefinite for $0 \le r \le 1$, c.n.d. for $1 \le r \le 2$, and c.p.d. for $2 \le r \le 3$. Since t^r is not operator convex on $(0,\infty)$ for $r > 2$, Corollary 2.8.5 shows that t^r for any $r > 3$ has a non-c.p.d. divided difference matrix and also a non-c.n.d. one. See [20] for more properties of divided difference matrices of t^r.

Remark 2.8.6. As is easily verified by translation $t \mapsto t+\alpha$ and by reversing $t \mapsto -t$, Theorem 2.8.3 similarly holds also when f is a function on (α,∞) or $(-\infty,\beta)$. In the latter case, the roles of c.n.d. and c.p.d. are exchanged. However, the theorem is not true when f is a function on a finite interval (α,β). For instance, according to Theorem 2.7.6, the functions

$$g_\lambda(t) := \frac{t^2}{1-\lambda t}, \quad \text{where } \lambda \in [-1,1], \tag{2.8.1}$$

are operator convex on $(-1,1)$. The function g_λ has c.n.d. divided difference matrices for $\lambda \in [-1,0)$ while g_λ does c.p.d. divided difference matrices for $\lambda \in (0,1]$, as shown in [21]. But it was also shown in [21] that the function $tg_\lambda(t)$ has c.p.d. divided difference matrices for all $\lambda \in [-1,1]$ so that, for every operator convex function f on $(-1,1)$, the function $tf(t)$ has c.p.d. divided difference matrices, which was indeed formerly proved by Horn [45] by a different method. The proofs of these facts are left for the next exercise.

Exercise 2.8.7. Let g_λ, $\lambda \in [-1,1]$, be the functions on $[-1,1]$ defined by (2.8.1). Prove the following:

(1) If $t_1,\dots,t_n \in (-1,1)$, then $[g_\lambda^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. for all $\lambda \in [-1,0)$ and is c.p.d. for all $\lambda \in (0,1]$.

(2) If $h_\lambda(t) := tg_\lambda(t)$ and $t_1,\dots,t_n \in (-1,1)$, then $[h_\lambda^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. for all $\lambda \in [-1,1]$.

Exercise 2.8.8. For every $m \in \mathbb{N}$ and every $t_1,\dots,t_n \in (0,\infty)$, show that

$$\langle x, [(t^m)^{[1]}(t_i,t_j)]x\rangle = -\langle x, [(t^{2-m})^{[1]}(t_i^{-1},t_j^{-1})]x\rangle, \qquad x \in \mathbb{C}_0^n.$$

This implies that $[(t^m)^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. (resp., c.n.d.) if and only if $[(t^{2-m})^{[1]}(t_i^{-1},t_j^{-1})]_{i,j=1}^n$ is c.n.d. (resp., c.p.d.).

More recently in [44] we considered the following conditions for a C^1 function f on $(0,\infty)$ and for each fixed integer $n \ge 1$:

(i)$_n$ f is matrix convex of order n on $(0,\infty)$;

(ii)$_n$ $\liminf_{t\to\infty} f(t)/t > -\infty$ and $[f^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.n.d. for all $t_1,\dots,t_n \in (0,\infty)$;

(iii)$_n$ $\limsup_{t\searrow 0} g(t) \ge 0$ and $[g^{[1]}(t_i,t_j)]_{i,j=1}^n$ is c.p.d. for all $t_1,\dots,t_n \in (0,\infty)$, where $g(t) := tf(t)$ for $t \in (0,\infty)$.

We further improved the above proof of Theorem 2.8.3 without use of integral representation of operator convex functions and proved the implications

$$(\mathrm{i})_{2n+1} \Longrightarrow (\mathrm{ii})_n, \qquad (\mathrm{ii})_{4n+1} \Longrightarrow (\mathrm{i})_n, \qquad (\mathrm{i})_{n+1} \Longrightarrow (\mathrm{iii})_n, \qquad (\mathrm{iii})_{2n+1} \Longrightarrow (\mathrm{i})_n.$$

In this way, it turned out that the results in [20] (also [75]) are refined to those for each matrix order.

3. Operator Means

3.1 Operator means and parallel sum

This chapter is a brief survey on operator means. An axiomatic approach for operator means was investigated by Kubo and Ando [58]. To introduce operator means in an axiomatic way, it is convenient to treat positive operators on a fixed separable and infinite-dimensional Hilbert space instead of $n \times n$ positive matrices for all separate n. So, throughout this chapter, we fix such an infinite-dimensional Hilbert space $\mathcal{H}$, and let $B(\mathcal{H})$ be the set of all bounded linear operators on $\mathcal{H}$. Let $B(\mathcal{H})^{sa}$ denote the set of all self-adjoint $A \in B(\mathcal{H})$ and $B(\mathcal{H})^+$ the set of all positive $A \in B(\mathcal{H})$.

First, we review the functional calculus and the spectral decomposition for self-adjoint operators on $\mathcal{H}$, which are the infinite-dimensional extensions of those for matrices explained in Section 1.4. For $A \in B(\mathcal{H})$ let $\sigma(A)$ be the *spectrum* of A, i.e., the set of $\lambda \in \mathbb{C}$ such that $\lambda I - A$ is not invertible in $B(\mathcal{H})$, which is a non-empty compact subset of $\mathbb{C}$ with $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \le \|A\|\}$, $\|A\|$ being the *operator norm* of A. Hence $r(A) \le \|A\|$, where $r(A) := \max\{|\lambda| : \lambda \in \sigma(A)\}$, the *spectral radius* of A. Moreover, note that $r(A) = \lim_{n\to\infty} \|A^n\|^{1/n}$ and that $r(A) = \|A\|$ if A is normal (i.e., $A^*A = AA^*$). Let $\mathbb{P}$ denote the linear space of all polynomials with complex coefficients. For $p(t) = \sum_{k=0}^N \alpha_k t^k$ in $\mathbb{P}$ we put $p(A) := \sum_{k=0}^N \alpha_k A^k$ as usual. When $A \in B(\mathcal{H})^{sa}$ (hence $\sigma(A) \subset [-\|A\|, \|A\|]$), we have

$$\sigma(p(A)) = p(\sigma(A)) = \{p(\lambda) : \lambda \in \sigma(A)\}$$

and hence

$$\|p(A)\| = \|p\|_{\sigma(A)} := \max\{|p(\lambda)| : \lambda \in \sigma(A)\}.$$

This means that $p \in \mathbb{P} \mapsto p(A) \in B(\mathcal{H})$ is an isometry with respect to the norm $\|p\|_{\sigma(A)}$ on $\mathbb{P}$ and the operator norm. This isometry can uniquely extend to $C(\sigma(A))$, the Banach space of continuous complex functions on $\sigma(A)$ with sup-norm, since $\mathbb{P}$ is dense in $C(\sigma(A))$. This extended isometry is written as $f \in C(\sigma(A)) \mapsto f(A) \in B(\mathcal{H})$ and $f(A)$ is called the *functional calculus* of A by f. We have $\sigma(f(A)) = f(\sigma(A))$, the *spectral mapping theorem*. When $A \ge 0$ (i.e., $A \in B(\mathcal{H})^+$) and $f(t) = t^r$ on $[0, \infty)$ with $r \ge 0$, we write A^r for $f(A)$. In particular, $A^0 = I$, the identity operator, by convention.

For $A, A_n \in B(\mathcal{H})$, $n \in \mathbb{N}$, it is said that A_n converges to A in the *strong operator topology* (or simply $A_n \to A$ strongly) if $\|(A_n - A)x\| \to 0$ for all $x \in \mathcal{H}$. Of course, the operator norm convergence $\|A_n - A\| \to 0$ implies the convergence in the strong operator topology. When $\mathcal{H}$ is finite dimensional, both convergences are equivalent.

Now let $A \in B(\mathcal{H})^{sa}$. For each $x, y \in \mathcal{H}$ define $\varphi_{x,y}(f) := \langle x, f(A)y\rangle$, $f \in C(\sigma(A))$, which is a bounded linear functional on $C(\sigma(A))$. Hence by the *Riesz–Markov theorem*, there is a unique complex Borel measure $\mu_{x,y}$ on $\sigma(A)$ such that

$$\langle x, f(A)y\rangle = \int_{\sigma(A)} f\, d\mu_{x,y}, \qquad f \in C(\sigma(A)).$$

For each Borel subset S of $\sigma(A)$, it follows that $\mu_{x,y}(S)$ is a sesqui-linear form on $\mathcal{H}$, i.e., $\mu_{x,y}(S)$ is conjugate-linear in x and linear in y, which is bounded as $|\mu_{x,y}(S)| \le \|x\|\,\|y\|$ for all $x, y \in \mathcal{H}$. Hence by the *Riesz representation theorem*, there is an $E(S) \in B(\mathcal{H})$ such that $\mu_{x,y}(S) = \langle x, E(S)y\rangle$ for all $x, y \in \mathcal{H}$. Since $\varphi_{x,x} \ge 0$ on $C(\sigma(A))$, $\mu_{x,x}$ is a positive measure for any $x \in \mathcal{H}$, so $E(S) \in B(\mathcal{H})^+$ for all Borel sets $S \subset \sigma(A)$. Here one can show that $E(S)$ is an orthogonal projection for every Borel set S and $E(\cdot)$ is σ-additive in the sense that

$$E(S) = \sum_{k=1}^{\infty} E(S_k) = \lim_{N\to\infty} \sum_{k=1}^{N} E(S_k)$$

in the strong operator topology if S_k, $k \in \mathbb{N}$, are mutually disjoint Borel subsets of $\sigma(A)$ with $S = \bigcup_{k=1}^\infty S_k$. Moreover, $E(\sigma(A)) = I$ since $\mu_{x,y}(\sigma(A)) = \langle x, y\rangle$ for $x, y \in \mathcal{H}$. In this way, one obtains a *spectral measure* $E(\cdot)$ on $\sigma(A)$ such that

$$\langle x, Ay\rangle = \int_{\sigma(A)} t\, d\mu_{x,y}(t) = \int_{\sigma(A)} t\, d\langle x, E(t)y\rangle, \qquad x, y \in \mathcal{H},$$

that is,

$$A = \int_{\sigma(A)} t\, dE(t), \tag{3.1.1}$$

which is called the *spectral decomposition* of A. One can also define a *resolution of identity* $\{E_t\}_{-\infty<t<\infty}$ by $E_t := E(\sigma(A) \cap (-\infty, t])$, which is a non-increasing and right-continuous one-parameter family of orthogonal projections with $E_t = 0$ for $t < -\|A\|$ and $E_t = I$ for $t \leq \|A\|$. Then the representation (3.1.1) is also written as

$$A = \int_{-\|A\|}^{\|A\|} t\, dE_t.$$

When $\mathcal{H}$ is finite dimensional and $\alpha_1, \ldots, \alpha_m$ are different eigenvalues of $A \in B(\mathcal{H})^{sa}$ with P_j the orthogonal projection onto the eigenspace $\ker(A - \alpha_j I)$ for $1 \leq j \leq m$, the spectral decomposition of A reduces to (1.4.2) in Section 1.4.

Exercise 3.1.1. Let f be a continuous complex function on an interval $[\alpha, \beta]$. Let $A, A_n \in B(\mathcal{H})^{sa}$, and assume that $A_n \to A$ in the strong operator topology and that $\sigma(A_n) \subset [\alpha, \beta]$ for all $n \in \mathbb{N}$. Show that $\sigma(A) \subset [\alpha, \beta]$ and $f(A_n) \to f(A)$ in the strong operator topology.

After the above short review on functional calculus and spectral decomposition, we now introduce the notion of operator means in the following:

Definition 3.1.2. A binary operation $\sigma : B(\mathcal{H})^+ \times B(\mathcal{H})^+ \to B(\mathcal{H})^+$ is called an *operator connection* if it satisfies the following conditions (i)–(iii) for $A, B, C, D \in B(\mathcal{H})^+$:

(i) $A \leq C$ and $B \leq D$ imply $A\,\sigma\, B \leq C\,\sigma\, D$ (joint monotonicity),
(ii) $C(A\,\sigma\, B)C \leq (CAC)\,\sigma\,(CBC)$ (transformer inequality),
(iii) $A_n, B_n \in B(\mathcal{H})^+$, $A_n \downarrow A$, and $B_n \downarrow B$ imply $A_n\,\sigma\, B_n \downarrow A\,\sigma\, B$ (upper semicontinuity), where $A_n \downarrow A$ means that $A_1 \geq A_2 \geq \ldots$ and $A_n \to A$ in the strong operator topology.

An operator connection σ is called an *operator mean* if

(iv) $I\,\sigma\, I = I$.

In the rest of this chapter, we always assume that A, B, C, D are elements of $B(\mathcal{H})^+$.

Proposition 3.1.3. *Assume that σ is an operator connection. If C is invertible, then*

$$C(A\,\sigma\, B)C = (CAC)\,\sigma\,(CBC). \tag{3.1.2}$$

For every $\alpha \geq 0$,

$$\alpha(A\,\sigma\, B) = (\alpha A)\,\sigma\,(\alpha B) \quad (\textit{positive homogeneity}). \tag{3.1.3}$$

Proof. From property (ii) above,

$$C^{-1}\{(CAC)\,\sigma\,(CBC)\}C^{-1} \leq A\,\sigma\, B$$

so that

$$(CAC)\,\sigma\,(CBC) \leq C(A\,\sigma\, B)C.$$

This and (ii) imply (3.1.2). When $\alpha > 0$, letting $C := \alpha^{1/2} I$ in (3.1.2) implies (3.1.3). When $\alpha = 0$, let $0 < \alpha_n \searrow 0$. Then $(\alpha_n)\,\sigma\,(\alpha_n I) \downarrow 0\,\sigma\, 0$ by (iii) above while $(\alpha_n I)\,\sigma\,(\alpha_n I) = \alpha_n(I\,\sigma\, I) \downarrow 0$. Hence $0 = 0\,\sigma\, 0$, which is (3.1.3) for $\alpha = 0$. □

Lemma 3.1.4. *For invertible $A, B \in B(\mathcal{H})^+$ define $A : B \in B(\mathcal{H})^+$ by*

$$A : B := (A^{-1} + B^{-1})^{-1}. \tag{3.1.4}$$

Then

(1) *Let $A, B, C, D \in B(\mathcal{H})^+$ be invertible. If $A \leq C$ and $B \leq D$, then $A : B \leq C : D$.*
(2) *Let $A, B, A_n, B_n \in B(\mathcal{H})^+$ for $n \geq 1$. If A, B are invertible, $A_n \downarrow A$ and $B_n \downarrow B$, then $A_n : B_n \downarrow A : B$.*
(3) *Let $A, B, A_n, B_n \in B(\mathcal{H})^+$ for $n \geq 1$. If A_n, B_n are invertible for $n \geq 1$, $A_n \downarrow A$ and $B_n \downarrow B$, then the limit $\lim_n A_n : B_n$ in the strong operator topology exists, and the limit is independent of the choices of A_n, B_n.*

Proof. (1) Since $A \leq C$ and $B \leq D$, $A^{-1} \geq C^{-1}$ and $B^{-1} \geq D^{-1}$ so that $A^{-1} + B^{-1} \geq C^{-1} + D^{-1}$. Hence $(A^{-1} + B^{-1})^{-1} \leq (C^{-1} + D^{-1})^{-1}$.

(2) Assume that A, B are invertible, $A_n \downarrow A$ and $B_n \downarrow B$. Then $A_1^{-1} \leq A_2^{-1} \leq \ldots$, $B_1^{-1} \leq B_2^{-1} \leq \ldots$, $A_n^{-1} \leq A^{-1}$ and $B_n^{-1} \leq B^{-1}$. Notice that

$$\begin{aligned}\langle x, (A^{-1} - A_n^{-1})x\rangle &= \langle x, A_n^{-1}(A_n - A)A^{-1}x\rangle = \langle A_n^{-1}x, (A_n - A)A^{-1}x\rangle \\ &\leq \|A_n^{-1}x\|\,\|(A_n - A)A^{-1}x\| \longrightarrow 0\end{aligned}$$

for every $x \in \mathcal{H}$. This implies that $A_n^{-1} \uparrow A^{-1}$ (i.e., A_n^{-1} increasingly converges to A in the strong operator topology). (In fact, this is also seen from Exercise 3.1.1.) Similarly $B_n^{-1} \uparrow B^{-1}$. Hence $A_n^{-1} + B_n^{-1} \uparrow A^{-1} + B^{-1}$. An argument similar to the above shows that $A_n : B_n \downarrow A : B$.

(3) For general $A, B \in B(\mathcal{H})^+$, let A_n, B_n be invertible with $A_n \downarrow A$ and $B_n \downarrow B$. It follows from (1) that $A_1 : B_1 \geq A_2 : B_2 \geq \ldots$ and so $\lim_n A_n : B_n$ in the strong operator topology exists. For any other invertible A'_n, B'_n with $A_n \downarrow A$ and $B_n \downarrow B$, since $A_n \leq A_n + A'_m - A$ and $B_n \leq B_n + B'_m - B$ for every $n, m \in \mathbb{N}$, we have by (1)

$$A_n : B_n \leq (A_n + A'_m - A) : (B_n + B'_m - B). \tag{3.1.5}$$

Since $A_n + A'_m - A \downarrow A'_m$ and $B_n + B'_m - B \downarrow B$ as $n \to \infty$, it follows from (2) that $(A_n + A'_m - A) : (B_n + B'_m - B) \downarrow A'_m : B'_m$ as $n \to \infty$. Passing to the limit of (3.1.5) as $n \to \infty$ we have $\lim_n A_n : B_n \leq A'_m : B'_m$. Hence letting $m \to \infty$ gives $\lim_n A_n : B_n \leq \lim_n A'_n : B'_n$. By symmetry we have $\lim_n A_n : B_n = \lim_n A'_n : B'_n$. $\square$

Thanks to Lemma 3.1.4 (3), one can extend (3.1.4) to general $A, B \in B(\mathcal{H})^+$ as follows:

$$A : B := \lim_{\varepsilon \searrow 0} (A + \varepsilon I) : (B + \varepsilon I) \quad \text{(in the strong operator topology)}. \tag{3.1.6}$$

This $A : B$ is called the *parallel sum* of A, B. The next variational expression of $A : B$ is useful.

Lemma 3.1.5. *For every $x \in \mathcal{H}$,*

$$\langle x, (A : B)x \rangle = \inf\{\langle y, Ay \rangle + \langle z, Bz \rangle : y, z \in \mathcal{H},\ y + z = x\}.$$

Proof. When A, B are invertible,

$$A : B = \{B^{-1}(A + B)A^{-1}\}^{-1} = \{(A + B) - B\}(A + B)^{-1}B = B - B(A + B)^{-1}B.$$

For every $x, y \in \mathcal{H}$ we have

$$\begin{aligned}
&\langle y, Ay \rangle + \langle x - y, B(x - y) \rangle - \langle x, (A : B)x \rangle \\
&\quad = \langle x, Bx \rangle + \langle y, (A + B)y \rangle - 2\,\mathrm{Re}\langle y, Bx \rangle - \langle x, (A : B)x \rangle \\
&\quad = \langle x, B(A + B)^{-1}Bx \rangle + \langle y, (A + B)y \rangle - 2\,\mathrm{Re}\langle y, Bx \rangle \\
&\quad = \|(A + B)^{-1/2}Bx\|^2 + \|(A + B)^{1/2}y\|^2 - 2\,\mathrm{Re}\langle (A + B)^{1/2}y, (A + B)^{-1/2}Bx \rangle \\
&\quad \geq 0.
\end{aligned}$$

In particular, the above is equal to 0 if $y = (A + B)^{-1}Bx$. Hence the assertion is shown when A, B are invertible. For general A, B,

$$\begin{aligned}
\langle x, (A : B)x \rangle &= \inf_{\varepsilon > 0} \langle x, \{(A + \varepsilon I) : (B + \varepsilon I)\}x \rangle \\
&= \inf_{\varepsilon > 0} \inf_{y} \{\langle y, (A + \varepsilon I)y \rangle + \langle x - y, (B + \varepsilon I)(x - y) \rangle\} \\
&= \inf_{y} \{\langle y, Ay \rangle + \langle x - y, B(x - y) \rangle\}.
\end{aligned}$$ $\square$

Corollary 3.1.6. *The parallel sum $A : B$ is an operator connection and the following hold:*

(1) *For every $S \in B(\mathcal{H})$, $S^*(A : B)S \leq (S^*AS) : (S^*BS)$.*

(2) *$(A : B) + (C : D) \leq (A + C) : (B + D)$.*

Proof. Let us first show (1) and (2).

(1) When $y + z = x$, Lemma 3.1.5 implies that

$$\begin{aligned}
\langle x, S^*(A : B)Sx \rangle = \langle Sx, (A : B)Sx \rangle &\leq \langle Sy, ASy \rangle + \langle Sz, BSz \rangle \\
&= \langle y, S^*ASy \rangle + \langle z, S^*BSz \rangle.
\end{aligned}$$

Hence $S^*(A : B)S \leq (S^*AS) : (S^*BS)$ by Lemma 3.1.5 again.

(2) When $y + z = x$,

$$\begin{aligned}
\langle x, \{(A : B) + (C : D)\}x \rangle &\leq \langle y, Ay \rangle + \langle z, Bz \rangle + \langle y, Cy \rangle + \langle z, Dz \rangle \\
&= \langle y, (A + C)y \rangle + \langle z, (B + D)z \rangle.
\end{aligned}$$

Hence $(A : B) + (C : D) \leq (A + C) : (B + D)$.

Next, we show that $A : B$ is an operator connection. (i) of Definition 3.1.2 is obvious from Lemma 3.1.4 (1) and Definition (3.1.6). (ii) is contained in (1). To show (iii), let $A_n \downarrow A$ and $B_n \downarrow B$. Since $A : B \leq A_n : B_n$ by (i), we have $A : B \leq \lim_n A_n : B_n$. For any $\varepsilon > 0$, since $A_n : B_n \leq (A_n + \varepsilon I) : (B_n + \varepsilon I)$, Lemma 3.1.4 (2) implies that $\lim_n A_n : B_n \leq (A + \varepsilon I) : (B + \varepsilon I)$. Hence $\lim_n A_n : B_n \leq A : B$ so that $A_n : B_n \downarrow A : B$. (It is also easy to show (i) and (iii) from Lemma 3.1.5.) $\square$

3.2 Kubo and Ando's theorem

The next fundamental theorem of Kubo and Ando says that there is a one-to-one correspondence between operator connections and operator monotone functions on $[0, \infty)$.

Theorem 3.2.1. *For each operator connection σ there exists a unique operator monotone function $f \geq 0$ on* $[0,\infty)$ *such that*

$$f(t)I = I\,\sigma\,(tI), \qquad t \geq 0. \tag{3.2.1}$$

Furthermore, the following properties are satisfied:

(1) *The map $\sigma \mapsto f$ is an affine order-isomorphism between the operator connections and the nonnegative operator monotone functions on* $[0,\infty)$. *Here, the order-isomorphism means that when $\sigma_i \mapsto f_i$ for* $i = 1,2$, *$A\,\sigma_1\,B \leq A\,\sigma_2\,B$ for all $A, B \in B(\mathcal{H})^+$ if and only if $f_1(t) \leq f_2(t)$ for all $t \geq 0$.*

(2) *If A is invertible, then*

$$A\,\sigma\,B = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}. \tag{3.2.2}$$

(3) *σ is an operator mean if and only if* $f(1) = 1$. *In this case, $A\,\sigma\,A = A$ for all A.*

Proof. Let σ be an operator connection. First we show that if a projection $P \in B(\mathcal{H})$ commutes with A and B, then P commutes $A\,\sigma\,B$ and

$$\{(AP)\,\sigma\,(BP)\}P = (A\,\sigma\,B)P. \tag{3.2.3}$$

Since $PAP = AP \leq A$ and $PBP = BP \leq B$, it follows from (ii) and (i) of Definition 3.1.2 that

$$P(A\,\sigma\,B)P \leq (PAP)\,\sigma\,(PBP) = (AP)\,\sigma\,(BP) \leq A\,\sigma\,B. \tag{3.2.4}$$

Hence $\{A\,\sigma\,B - P(A\,\sigma\,B)P\}^{1/2}$ exists so that

$$|\{A\,\sigma\,B - P(A\,\sigma\,B)P\}^{1/2}P|^2 = P\{A\,\sigma\,B - P(A\,\sigma\,B)P\}P = 0.$$

Therefore, $\{A\,\sigma\,B - P(A\,\sigma\,B)P\}^{1/2}P = 0$ and so $(A\,\sigma\,B)P = P(A\,\sigma\,B)P$. This implies that P commutes with $A\,\sigma\,B$. Similarly, P commutes with $(AP)\,\sigma\,(BP)$ as well, and (3.2.3) follows from (3.2.4). Hence we see that there is a function $f \geq 0$ on $[0,\infty)$ satisfying (3.2.1). The uniqueness of such function f is obvious, and it follows from (iii) of Definition 3.1.2 that f is right-continuous for $t \geq 0$. Since $t^{-1}f(t)I = (t^{-1}I)\,\sigma\,I$ for $t > 0$ thanks to (3.1.3), if follows from (iii) of Definition 3.1.2 again that $t^{-1}f(t)$ is left-continuous for $t > 0$ and so is $f(t)$. Hence f is continuous on $[0,\infty)$.

To show the operator monotonicity of f, let us prove that

$$f(A) = I\,\sigma\,A. \tag{3.2.5}$$

Let $A = \sum_{i=1}^m \alpha_i P_i$, where $\alpha_i > 0$ and P_i are projections with $\sum_{i=1}^m P_i = I$. Since each P_i commute with A, using (3.2.3) twice we have

$$\begin{aligned} I\,\sigma\,A &= \sum_{i=1}^m (I\,\sigma\,A)P_i = \sum_{i=1}^m \{P_i\,\sigma\,(AP_i)\}P_i = \sum_{i=1}^m \{P_i\,\sigma\,(\alpha_i P_i)\}P_i \\ &= \sum_{i=1}^m \{I\,\sigma\,(\alpha_i I)\}P_i = \sum_{i=1}^m f(\alpha_i)P_i = f(A). \end{aligned}$$

For general $A \in B(\mathcal{H})^+$ choose a sequence $\{A_n\}$ in $B(\mathcal{H})^+$ of the above form such that $A_n \downarrow A$. By (iii) of Definition 3.1.2 and Exercise 3.1.1 we have

$$I\,\sigma\,A = \lim_{n\to\infty} I\,\sigma\,A_n = \lim_{n\to\infty} f(A_n) = f(A)$$

in the strong operator topology, and so (3.2.5) is shown. Hence, if $A, B \in B(\mathcal{H})^+$ and $A \leq B$, then

$$f(A) = I\,\sigma\,A \leq I\,\sigma\,B = f(B),$$

showing that f is operator monotone. In the rest we prove (1)–(3).

(1) It suffices to show that $\sigma \mapsto f$ is surjective onto the set of nonnegative operator monotone functions on $[0,\infty)$, since the remaining assertions are obvious from (3.2.1) and (3.2.5). So let $f \geq 0$ be operator monotone on $[0,\infty)$. By Theorem 2.7.11 we have $a, b \geq 0$ and a finite positive measure m on $(0,\infty)$ so that

$$f(t) = a + bt + \int_{(0,\infty)} \frac{t(1+\lambda)}{t+\lambda}\,dm(\lambda).$$

Define a binary operation σ on $B(\mathcal{H})^+$ by

$$A\,\sigma\,B := aA + bB + \int_{(0,\infty)} \frac{1+\lambda}{\lambda}\{(\lambda A) : B\}\,dm(\lambda).$$

In fact, since

$$(\lambda A) : B \le (\lambda \|A\| I) : (\|B\| I) = \frac{\|A\|\,\|B\|\lambda}{\|A\|\lambda + \|B\|} I$$

so that

$$\frac{1+\lambda}{\lambda}\,\|(\lambda A) : B\| \le \frac{\|A\|\,\|B\|(1+\lambda)}{\|A\|\lambda + \|B\|},$$

it follows that $\frac{\lambda}{1+\lambda}\,\{(\lambda A) : B\}$ is uniformly bounded for $\lambda > 0$. Hence $A\,\sigma\,B$ is well defined as an element of $B(\mathcal{H})^+$. Now it is easy to see from Corollary 3.1.6 that $A\,\sigma\,B$ is an operator connection. For instance, we show (iii) of Definition 3.1.2: if $A_n \downarrow A$ and $B_n \downarrow B$, then $(\lambda A_n) : B_n \downarrow (\lambda A) : B$ for all $\lambda \in (0,\infty)$ and so by the Lebesgue convergence theorem we have

$$\begin{aligned}\lim_{n\to\infty} \langle x, (A_n\,\sigma\,B_n)x\rangle &= \lim_{n\to\infty}\left[a\langle x, A_n x\rangle + b\langle x, B_n x\rangle + \int_{(0,\infty)} \langle x, \{(\lambda A_n) : B_n\}x\rangle\, dm(\lambda)\right] \\ &= a\langle x, Ax\rangle + b\langle x, Bx\rangle + \int_{(0,\infty)} \langle x, \{(\lambda A) : B\}x\rangle\, dm(\lambda) \\ &= \langle x, (A\,\sigma\,B)x\rangle.\end{aligned}$$

Hence $A_n\,\sigma\,B_n \downarrow A\,\sigma\,B$. For this operator connection σ, $I\,\sigma\,(tI) = f(t)I$ holds for all $t \ge 0$ so that f is the operator monotone function corresponding to σ.

(2) When A is invertible, it follows from (3.1.2) and (3.2.5) that

$$A\,\sigma\,B = A^{1/2}(I\,\sigma\,A^{-1/2}BA^{-1/2})A^{1/2} = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}.$$

(3) The first assertion is immediate to see. When $f(1) = 1$, it follows from (3.2.2) that $A\,\sigma\,A = A$ for invertible A. By continuity this holds for all $A \in B(\mathcal{H})^+$. □

Let σ be an operator connection and f be the operator monotone function on $[0,\infty)$ corresponding to σ as described in the above theorem. When A is invertible, $A\,\sigma\,B$ is written as in (3.2.2). For general $A, B \in B(\mathcal{H})^+$, by (iii) of Definition 3.1.2 one can define as

$$A\,\sigma\,B = \lim_{\varepsilon\searrow 0} A_\varepsilon\,\sigma\,B_\varepsilon = \lim_{\varepsilon\searrow 0} A_\varepsilon^{1/2} f(A_\varepsilon^{-1/2}B_\varepsilon A_\varepsilon^{-1/2})A_\varepsilon^{1/2} \quad \text{(in the strong operator topology)},$$

where $A_\varepsilon := A + \varepsilon I$ and $B_\varepsilon := B + \varepsilon I$. We call f the *representing function* of σ. For scalars $s, t \ge 0$ note that $(sI)\,\sigma\,(tI)$ is a scalar multiple of I thanks to (3.1.3) and (3.2.1), so we write $s\,\sigma\,t$ for this scalar. In fact, $s\,\sigma\,t = sf(t/s)$ for $s > 0$.

The next proposition is seen from Theorem 2.7.11 and the proof of Theorem 3.2.1.

Proposition 3.2.2. *For every operator connection σ, there exists a unique positive finite Borel measure m on* $[0,\infty]$ *such that*

$$A\,\sigma\,B = aA + bB + \int_{(0,\infty)} \frac{1+\lambda}{\lambda}\,\{(\lambda A) : B\}\, dm(\lambda), \qquad A, B \in B(\mathcal{H})^+, \tag{3.2.6}$$

where $a := m(\{0\})$ and $b := m(\{\infty\})$. The map $\sigma \mapsto m$ is a bijective affine correspondence between the operator connections and the positive finite Borel measures on $[0,\infty]$.

Due to the integral expression (3.2.6), one can derive properties of general operator connections by checking them for only parallel sum. For instance, the following corollary is obvious from Corollary 3.1.6 and (3.2.6).

Corollary 3.2.3. *For every operator connection σ the following hold:*

(1) *For every $S \in B(\mathcal{H})$, $S^*(A\,\sigma\,B)S \le (S^*AS)\,\sigma\,(S^*BS)$ (transformer inequality) and equality holds if S is invertible.*

(2) *$(A\,\sigma\,B) + (C\,\sigma\,D) \le (A + C)\,\sigma\,(B + D)$ (concavity).*

It is quite instructive to consider operator connections form the point of view of electrical circuits. An impedance of an n-port resistive network is represented by an $n \times n$ positive matrix A. The equation $v = Ax$ holds for n-dimensional vectors of current x and voltage v, and the electrical power is given by $\langle x, Ax\rangle$. For two impedances A and B, their series and parallel connections are given by the sum $A + B$ and the parallel sum $A : B$, respectively. Lemma 3.1.4 means Maxwell's principle that current runs through a parallel connection so as to minimize the electrical power. A general operator connection represent a formation of making a new impedance from two given impedances A, B. The integral expression (3.2.6) shows that such a formation can be realized as a weighted series connection of (infinite) weighted parallel connections. In this way, the theory of operator connections can be regarded as a mathematical theory of electrical circuits. Indeed, the notion of parallel sum for positive operators was introduced from the viewpoint of electrical circuits in [2].

3.3 Examples and properties of operator means

The following are typical examples of operator means.

Example 3.3.1.

(1) *Arithmetic mean*: $A \triangledown B := \frac{1}{2}(A+B)$, whose representing function is $(1+t)/2$.

(2) *Harmonic mean*: $A \,!\, B := 2(A : B)$, whose representing function is $2t/(1+t)$.

(3) *Geometric mean*: $A \# B := \lim_{\varepsilon \searrow 0} A_\varepsilon^{1/2}(A_\varepsilon^{-1/2} B_\varepsilon A_\varepsilon^{-1/2})^{1/2} A_\varepsilon^{1/2}$, whose representing function is $t^{1/2}$. The notion of geometric mean was first introduced in [69] and developed in [3, 58].

(4) For $0 \le \alpha \le 1$ let $\#_\alpha$ denote the α-*power mean*, which is the operator mean corresponding to the operator monotone function t^α. Namely, for each $A, B \in B(\mathcal{H})^+$ with A invertible, $A \#_\alpha B$ is defined by
$$A \#_\alpha B := A^{1/2}(A^{-1/2} B A^{-1/2})^\alpha A^{1/2}.$$
Here we recall the convention $B^0 = I$ for any $B \in B(\mathcal{H})^+$. Note that $A \#_0 B = A$, $A \#_1 B = B$, and $A \#_{1/2} B = A \# B$.

(5) The operator mean corresponding to the operator monotone function $(t-1)/\log t$ (see Example 2.5.9 (5)) is called the *logarithmic mean* and denoted by $A \,\lambda\, B$.

Among the operator means $\triangledown$, !, #, and λ, the following orders hold:

Proposition 3.3.2. *For every $A, B \in B(\mathcal{H})^+$,*
$$A \,!\, B \le A \# B \le A \,\lambda\, B \le A \triangledown B.$$

Proof. Thanks to Theorem 3.2.1 (1) it suffices to show that
$$\frac{2t}{1+t} \le t^{1/2} \le \frac{t-1}{\log t} \le \frac{1+t}{2}, \qquad t > 0. \tag{3.3.1}$$
But these can be shown by an elementary calculus, and so the details are left for an exercise. □

Exercise 3.3.3. Show (3.3.1).

The following are the variational expressions for the geometric and the harmonic means in terms of 2×2 operator matrices.

Proposition 3.3.4.

(1) $A \# B = \max\left\{X \in B(\mathcal{H})^+ : \begin{bmatrix} A & X \\ X & B \end{bmatrix} \ge 0\right\}$.

(2) $A \,!\, B = \max\left\{X \in B(\mathcal{H})^+ : \begin{bmatrix} 2A & 0 \\ 0 & 2B \end{bmatrix} \ge \begin{bmatrix} X & X \\ X & X \end{bmatrix}\right\}$.

Proof. (1) When A, B are invertible, since
$$\begin{bmatrix} I & A^{-1/2}XB^{-1/2} \\ B^{-1/2}XA^{-1/2} & I \end{bmatrix} = \begin{bmatrix} A^{-1/2} & 0 \\ 0 & B^{-1/2} \end{bmatrix} \begin{bmatrix} A & X \\ X & B \end{bmatrix} \begin{bmatrix} A^{-1/2} & 0 \\ 0 & B^{-1/2} \end{bmatrix},$$
we notice that $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \ge 0$ if and only if $\begin{bmatrix} I & A^{-1/2}XB^{-1/2} \\ B^{-1/2}XA^{-1/2} & I \end{bmatrix} \ge 0$. By Lemma 1.7.2 this is equivalent to $\|A^{-1/2}XB^{-1/2}\| \le 1$, that is, $B^{-1/2}XA^{-1}XB^{-1/2} \le I$ or $XA^{-1}X \le B$. If $X = A \# B$, then
$$\begin{aligned} XA^{-1}X &= A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}A^{-1}A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \\ &= A^{1/2}(A^{-1/2}BA^{-1/2})A^{1/2} = B. \end{aligned}$$
Also, if $XA^{-1}X \le B$, then $A^{-1/2}XA^{-1}XA^{-1/2} \le A^{-1/2}BA^{-1/2}$ so that $A^{-1/2}XA^{-1/2} \le (A^{-1/2}BA^{-1/2})^{1/2}$, implying that $X \le A \# B$. Hence $A \# B$ is the largest $X \in B(\mathcal{H})^+$ satisfying $XA^{-1}X \le B$.

For general A, B, let $A_\varepsilon := A + \varepsilon I$ and $B_\varepsilon := B + \varepsilon I$ for $\varepsilon > 0$. Since $\begin{bmatrix} A_\varepsilon & A_\varepsilon \# B_\varepsilon \\ A_\varepsilon \# B_\varepsilon & B_\varepsilon \end{bmatrix} \ge 0$, we have $\begin{bmatrix} A & A \# B \\ A \# B & B \end{bmatrix} \ge 0$ by letting $\varepsilon \searrow 0$. If $X \in B(\mathcal{H})^+$ and $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \ge 0$, then $\begin{bmatrix} A_\varepsilon & X \\ X & B_\varepsilon \end{bmatrix} \ge 0$ so that $X \le A_\varepsilon \# B_\varepsilon$ and letting $\varepsilon \searrow 0$ gives $X \le A \# B$. Therefore, we obtain the conclusion.

(2) This expression is a reformulation of that in Lemma 3.1.5, whose proof is left for an exercise below. □

Exercise 3.3.5. Prove (2) of Proposition 3.3.4.

Proposition 3.3.6. *Let σ be an operator mean and a, b be as in Proposition 3.2.2. Then for every orthogonal projections P, Q,*

$$P\,\sigma\,Q = a(P - P\wedge Q) + b(Q - P\wedge Q) + P\wedge Q,$$

where $P\wedge Q$ is the orthogonal projection onto $\operatorname{ran} P \cap \operatorname{ran} Q$.

Proof. For any $A, B \in B(\mathcal{H})^+$, we notice by Lemma 3.1.5 that $\ker A \cup \ker B \subset \ker(A : B)$. Hence

$$\overline{\operatorname{ran}}\,(A : B) = (\ker(A : B))^\perp \subset (\ker A \cup \ker B)^\perp = \overline{\operatorname{ran}}\,A \cap \overline{\operatorname{ran}}\,B,$$

where $\overline{\operatorname{ran}}\,A$ denotes the closure of $\operatorname{ran} A$. Hence we have $\operatorname{ran}(\lambda P) : Q) \subset \operatorname{ran}(P\wedge Q)$ for all $\lambda > 0$. Since $P\wedge Q$ commutes with λP, Q, using (3.2.3) twice we have

$$\begin{aligned}(\lambda P) : Q &= \{(\lambda P) : Q\}(P\wedge Q) = \{(\lambda P(P\wedge Q)) : (Q(P\wedge Q))\}(P\wedge Q)\\ &= \{(\lambda(P\wedge Q)) : (P\wedge Q)\}(P\wedge Q) = \{(\lambda I) : I\}(P\wedge Q) = \frac{\lambda}{1+\lambda}(P\wedge Q).\end{aligned}$$

Therefore, Proposition 3.2.2 implies that

$$\begin{aligned}P\,\sigma\,Q &= aP + bQ + \left\{\int_{(0,\infty)} dm(\lambda)\right\}(P\wedge Q)\\ &= aP + bQ + (1 - a - b)(P\wedge Q)\\ &= a(P - P\wedge Q) + b(Q - P\wedge Q) + P\wedge Q.\end{aligned}$$ □

Corollary 3.3.7. *For every orthogonal projections P, Q,*

$$P\,!\,Q = P\,\#\,Q = P\,\lambda\,Q = P\wedge Q.$$

Proof. When f is the representing function of an operator mean σ, recall (Theorem 2.7.11) that $a = f(0)$ and $b = \lim_{t\to\infty} f(t)/t$. We have $a = b = 0$ for !, #, and λ. □

In the rest of this section we discuss certain operations on operator means in connection with the corresponding operations on operator monotone functions.

Let σ be an operator mean with the corresponding function f. Note that $f > 0$ on $(0,\infty)$. In fact, suppose that $f(\alpha) = 0$ for some $\alpha > 0$. Then $f(t) = 0$ for all $t \in [0,\alpha]$, and the concavity of f implies that $f \equiv 0$, contradicting $f(1) = 1$. Hence Corollary 2.5.6 implies that $t/f(t)$ is operator monotone on $(0,\infty)$. Furthermore, it is easy to see that $f(t^{-1})^{-1}$ is operator monotone on $(0,\infty)$. Hence $tf(t^{-1})$ is also operator monotone on $(0,\infty)$ by Corollary 2.5.6 once again. By fixing the value at 0 as the limit as $t \searrow 0$, the functions $t/f(t)$, $f(t^{-1})^{-1}$, and $tf(t^{-1})$ are operator monotone functions on $[0,\infty)$. Now the following definitions are meaningful.

Definition 3.3.8. Let σ be an operator mean and f be the corresponding function.
(1) The operator mean with the representing function $tf(t^{-1})$ is called the *transpose* of σ and denoted by σ'. If $\sigma = \sigma'$, σ is said to be *symmetric*.
(2) The operator mean with the representing function $f(t^{-1})^{-1}$ is called the *adjoint* of σ and denoted by σ^*.
(3) The operator mean with the representing function $t/f(t)$ is called the *dual* of σ and denoted by $\sigma^\perp$.

The assertions in the following proposition are immediately verified from the above definitions.

Proposition 3.3.9. *Let σ be an operator mean and f be the corresponding function.*
(1) $A\,\sigma'\,B = B\,\sigma A$.
(2) *σ is symmetric if and only if $f(t) = tf(t^{-1})$ for all $t > 0$.*
(3) *When A, B are invertible, $A\,\sigma^*\,B = (A^{-1}\,\sigma\,B^{-1})^{-1}$.*
(4) *$(\sigma')' = \sigma$, $(\sigma^*)^* = \sigma$, and $(\sigma^\perp)^\perp = \sigma$.*
(5) *$\sigma^\perp = (\sigma')^* = (\sigma^*)'$, $\sigma' = (\sigma^*)^\perp = (\sigma^\perp)^*$, and $\sigma^* = (\sigma')^\perp = (\sigma^\perp)'$.*

Exercise 3.3.10. Prove Proposition 3.3.9.

For operator connections (resp., operator means) σ_1, σ_2, and σ_3, the operation defined by

$$(A, B) \mapsto (A\,\sigma_1\,B)\,\sigma_3\,(A\,\sigma_2\,B)$$

becomes an operator connection (resp., operator mean) again. If $A\,\sigma_1\,B \le A\,\sigma_2\,B$ holds for all A, B, then we write $\sigma_1 \le \sigma_2$.

Proposition 3.3.11. *If σ is a symmetric operator mean, then $! \le \sigma \le \triangledown$. That is, $\triangledown$ is maximal and $!$ is minimal among the symmetric operator means.*

Proof. Let f be the representing function of σ. From Theorem 3.2.1 (1) it suffices to show that

$$\frac{2t}{1+t} \le f(t) \le \frac{1+t}{2}, \qquad t \ge 0.$$

Recall (see Theorem 2.4.1) that f is differentiable on $(0,\infty)$. Since $f(t) = tf(t^{-1})$ for $t > 0$ by Proposition 3.3.9 (2), it follows that $f'(1) = f(1) - f'(1) = 1 - f'(1)$ and so $f'(1) = 1/2$. The concavity of f implies that $f(t) \le (1+t)/2$. Since σ^* is also symmetric by Proposition 3.3.9 (5), the same argument applied to $f(t^{-1})^{-1}$ implies that $f(t) \ge 2t/(1+t)$. □

Furthermore, let σ be an arbitrary operator mean with the corresponding function f, and let $\alpha := f'(1)$. From the concavity of f we have $0 \le \alpha \le 1$ and

$$\frac{t}{(1-\alpha)t+\alpha} \le f(t) \le (1-\alpha) + \alpha t, \qquad t \ge 0, \tag{3.3.2}$$

as in the proof of the above proposition. Hence $A\,\sigma\,B$ is between the weighted arithmetic mean $(1-\alpha)A + \alpha B$ and the weighted harmonic mean $\lim_{\varepsilon\searrow 0}((1-\alpha)A_\varepsilon^{-1} + \alpha B_\varepsilon^{-1})^{-1}$.

Proposition 3.3.12. *For every operator mean σ the following hold:*
(1) $(A\,\sigma\,B) + (B\,\sigma\,A) \le A + B$.
(2) $(A\,\sigma\,B) : (B\,\sigma\,A) \ge A : B$.
(3) $(A\,\sigma\,B)\,\#\,(A\,\sigma^\perp\,B) = A\,\#\,B$.
(4) $(A\,\sigma\,B) + (A\,\sigma^\perp\,B) \le A + B$.
(5) $(A\,\sigma\,B) : (A\,\sigma^\perp\,B) \ge A : B$.

Proof. Since $(A,B) \mapsto \frac{1}{2}\{(A\,\sigma\,B) + (B\,\sigma\,A)\}$ and $(A,B) \mapsto 2\{(A\,\sigma\,B) : (B\,\sigma\,A)\}$ are symmetric operator means, Proposition 3.3.11 implies that

$$\frac{1}{2}\{(A\,\sigma\,B) + (B\,\sigma\,A) \le A \triangledown B, \qquad 2\{(A\,\sigma\,B) : (B\,\sigma\,A)\} \ge A\,!\,B.$$

Hence (1) and (2) hold.

Let f be the representing function of σ. The left-hand sides of (3)–(5) are operator connections, whose representing functions are

$$\begin{aligned}(1\,\sigma\,t)\,\#\,(1\,\sigma^\perp\,t) &= \left\{f(t)\cdot\frac{t}{f(t)}\right\}^{1/2} = t^{1/2},\\ (1\,\sigma\,t) + (1\,\sigma^\perp\,t) &= f(t) + \frac{1}{f(t)} = \frac{t+f(t)^2}{f(t)},\\ (1\,\sigma\,t) : (1\,\sigma^\perp\,t) &= \left\{\frac{1}{f(t)} + \frac{f(t)}{t}\right\}^{-1} = \frac{tf(t)}{t+f(t)^2}.\end{aligned}$$

Hence (3) holds. For (4) it suffices to show that

$$\frac{t+f(t)^2}{f(t)} \le 1+t, \tag{3.3.3}$$

which is equivalent to $\{f(t)-1\}\{t-f(t)\} \ge 0$. This is seen because from (3.3.2) we have $t \le f(t) \le 1$ if $0 \le t \le 1$ and $1 \le f(t) \le t$ if $t \ge 1$. For (5) it suffices to show that

$$\frac{tf(t)}{t+f(t)^2} \ge \frac{t}{1+t},$$

which is equivalent to (3.3.3). □

The following is a counterpart of Corollary 3.2.3 (2).

Proposition 3.3.13. *For every operator connection σ,*

$$(A\,\sigma\,B) : (C\,\sigma\,D) \ge (A : C)\,\sigma\,(B : D).$$

Proof. Since the result is clear in the case $\sigma = 0$, we assume that $\sigma \ne 0$. Then we may assume that σ is an operator mean. From upper semicontinuity, it is enough to show the result when A, B, C, D are invertible. Then the inequality in question is written as

$$(A^{-1}\,\sigma^*\,B^{-1} + C^{-1}\,\sigma^*\,D^{-1})^{-1} \ge \{(A^{-1} + C^{-1})\,\sigma^*\,(B^{-1} + D^{-1})\}^{-1}.$$

This is seen from Corollary 3.2.3 (2). □

4. Spectral Variation and Majorization

4.1 Majorization for vectors

Let us start with the majorization for real vectors, which was introduced by Hardy, Littlewood, and Pólya. For two vectors $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$ in $\mathbb{R}^n$, the *weak majorization* $a \prec_w b$ means that

$$\sum_{i=1}^k a_{[i]} \le \sum_{i=1}^k b_{[i]}, \qquad 1 \le k \le n, \tag{4.1.1}$$

where $(a_{[1]}, \dots, a_{[n]})$ is the *decreasing rearrangement* of a, i.e., $a_{[1]} \ge \cdots \ge a_{[n]}$ are the components of a in decreasing order. The *majorization* $a \prec b$ means that $a \prec_w b$ and the equality holds for $k = n$ in (4.1.1). The characterizations of majorization and weak majorization in the following propositions are fundamental.

Proposition 4.1.1. *The following conditions for $a, b \in \mathbb{R}^n$ are equivalent:*
(i) $a \prec b$;
(ii) $\sum_{i=1}^n |a_i - r| \le \sum_{i=1}^n |b_i - r|$ *for all* $r \in \mathbb{R}$;
(iii) $\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n f(b_i)$ *for any convex function* f *on an interval containing all* a_i, b_i;
(iv) *a is a convex combination of coordinate permutations of b;*
(v) $a = Db$ *for some doubly stochastic* $n \times n$ *matrix* D, *i.e.,* $D = [d_{ij}]_{i,j=1}^n$ *with* $d_{ij} \ge 0$, $\sum_{j=1}^n d_{ij} = 1$ *for* $1 \le i \le n$, *and* $\sum_{i=1}^n d_{ij} = 1$ *for* $1 \le j \le n$.

Proof. (i) $\Rightarrow$ (iv). We show that there exist a finite number of matrices $D_1, \dots, D_N$ of the form $\lambda I + (1-\lambda)\Pi$, where $0 \le \lambda \le 1$ and Π is a permutation matrix interchanging two coordinates only, such that $a = D_N \cdots D_1 b$. Then (iv) follows because $D_N \cdots D_1$ becomes a convex combination of permutation matrices. We may assume that $a_1 \ge \cdots \ge a_n$ and $b_1 \ge \cdots \ge b_n$. Suppose $a \ne b$ and choose the largest j such that $a_j < b_j$. Then there exists a k with $k > j$ such that $a_k > b_k$. Choose the smallest such k. Let $\lambda_1 := 1 - \min\{b_j - a_j, a_k - b_k\}/(b_j - b_k)$ and Π_1 be the permutation matrix interchanging the jth and kth coordinates. Then $0 < \lambda_1 < 1$ since $b_j > a_j \ge a_k > b_k$. Define $D_1 := \lambda_1 I + (1-\lambda_1)\Pi_1$ and $b^{(1)} := D_1 b$. Now it is easy to check that $a \prec b^{(1)} \prec b$ and $b_1^{(1)} \ge \cdots \ge b_n^{(1)}$. Moreover the jth or kth coordinates of a and $b^{(1)}$ are equal. When $a \ne b^{(1)}$, we can apply the above argument to a and $b^{(1)}$. Repeating this finite times we reach the conclusion.

(iv) $\Rightarrow$ (v) is trivial from the fact that any convex combination of permutation matrices is doubly stochastic.

(v) $\Rightarrow$ (ii). For every $r \in \mathbb{R}$ we have

$$\sum_{i=1}^n |a_i - r| = \sum_{i=1}^n \left| \sum_{j=1}^n d_{ij}(b_j - r) \right| \le \sum_{i,j=1}^n d_{ij}|b_j - r| = \sum_{j=1}^n |b_j - r|.$$

(ii) $\Rightarrow$ (i). Taking large r and small r in the inequality of (ii) we have $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$. Noting that $|x| + x = 2x_+$ for $x \in \mathbb{R}$, where $x_+ = \max\{x, 0\}$, we have

$$\sum_{i=1}^n (a_i - r)_+ \le \sum_{i=1}^n (b_i - r)_+, \qquad r \in \mathbb{R}. \tag{4.1.2}$$

Now prove that (4.1.2) implies that $a \prec_w b$. When $b_{[k]} \ge r \ge b_{[k+1]}$, $\sum_{i=1}^k a_{[i]} \le \sum_{i-1}^k b_{[i]}$ follows since

$$\sum_{i=1}^n (a_i - r)_+ \ge \sum_{i=1}^k (a_{[i]} - r)_+ \ge \sum_{i=1}^k a_{[i]} - kr, \qquad \sum_{i=1}^n (b_i - r)_+ = \sum_{i=1}^k b_{[i]} - kr.$$

(iv) $\Rightarrow$ (iii). Suppose that $a_i = \sum_{k=1}^N \lambda_k b_{\pi_k(i)}$, $1 \le i \le n$, where $\lambda_k > 0$, $\sum_{k=1}^N \lambda_k = 1$, and π_k are permutations on $\{1, \dots, n\}$. Then the convexity of f implies that

$$\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n \sum_{k=1}^N \lambda_k f(b_{\pi_k(i)}) = \sum_{i=1}^n f(b_i).$$

(iii) $\Rightarrow$ (ii) is trivial since $f(x) = |x - r|$ is convex. $\square$

Note that (v) $\Rightarrow$ (iv) is seen directly from the well-known theorem of Birkhoff [22] saying that any doubly stochastic matrix is a convex combination of permutation matrices.

Exercise 4.1.2. Let Δ_n denote the set of all probability vectors in $\mathbb{R}^n$, i.e., $\Delta_n := \{p = (p_1, \dots, p_n) : p_i \ge 0, \sum_{i=1}^n p_i = 1\}$. Prove that

$$(1/n, 1/n, \dots, 1/n) \prec p \prec (1, 0, \dots, 0), \qquad p \in \Delta_n.$$

The *Shannon entropy* of $p \in \Delta_n$ is $H(p) := -\sum_{i=1}^n p_i \log p_i$. Show that $H(q) \le H(p) \le \log n$ for all $p \prec q$ in Δ_n and that, for $p \in \Delta_n$, $H(p) = \log n$ if and only if $p = (1/n, \dots, 1/n)$.

Proposition 4.1.3. *The following conditions* (i)–(iv) *for $a, b \in \mathbb{R}^n$ are equivalent:*

(i) $a \prec_w b$;

(ii) *there exists a $c \in \mathbb{R}^n$ such that $a \le c \prec b$, where $a \le c$ means that $a_i \le c_i$, $1 \le i \le n$;*

(iii) $\sum_{i=1}^n (a_i - r)_+ \le \sum_{i=1}^n (b_i - r)_+$ *for all $r \in \mathbb{R}$;*

(iv) $\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n f(b_i)$ *for any non-decreasing convex function f on an interval containing all a_i, b_i.*

Moreover, if $a, b \ge 0$, then the above conditions are equivalent to the following:

(v) *$a = Sb$ for some doubly substochastic $n \times n$ matrix S, i.e., $S = [s_{ij}]_{i,j=1}^n$ with $s_{ij} \ge 0$, $\sum_{j=1}^n s_{ij} \le 1$ for $1 \le i \le n$, and $\sum_{i=1}^n s_{ij} \le 1$ for $1 \le j \le n$.*

Proof. (i) $\Rightarrow$ (ii). By induction on n. We may assume that $a_1 \ge \cdots \ge a_n$ and $b_1 \ge \cdots \ge b_n$. Let $\alpha := \min_{1 \le k \le n}(\sum_{i=1}^k b_i - \sum_{i=1}^k a_i)$ and define $\tilde{a} := (a_1 + \alpha, a_2, \ldots, a_n)$. Then $a \le \tilde{a} \prec_w b$ and $\sum_{i=1}^k \tilde{a}_i = \sum_{i=1}^k b_i$ for some $1 \le k \le n$. When $k = n$, $a \le \tilde{a} \prec b$. When $k < n$, we have $(\tilde{a}_1, \ldots, \tilde{a}_k) \prec (b_1, \ldots, b_k)$ and $(\tilde{a}_{k+1}, \ldots, \tilde{a}_n) \prec_w (b_{k+1}, \ldots, b_n)$. Hence the induction assumption implies that $(\tilde{a}_{k+1}, \ldots, \tilde{a}_n) \le (c_{k+1}, \ldots, c_n) \prec (b_{k+1}, \ldots, b_n)$ for some $(c_{k+1}, \ldots, c_n) \in \mathbb{R}^{n-k}$. Then $a \le (\tilde{a}_1, \ldots, \tilde{a}_k, c_{k+1}, \ldots, c_n) \prec b$ is immediate from $\tilde{a}_k \ge b_k \ge b_{k+1} \ge c_{k+1}$.

(ii) $\Rightarrow$ (iv). Let $a \le c \prec b$. If f is non-decreasing and convex on an interval $[\alpha, \beta]$ containing a_i, b_i, then $c_i \in [\alpha, \beta]$ and

$$\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n f(c_i) \le \sum_{i=1}^n f(b_i)$$

by Proposition 4.1.1.

(iv) $\Rightarrow$ (iii) is trivial and (iii) $\Rightarrow$ (i) was already shown in the proof (ii) $\Rightarrow$ (i) of Proposition 4.1.1.

Now assume $a, b \ge 0$ and prove that (ii) $\Leftrightarrow$ (v). If $a \le c \prec b$, then we have, by Proposition 4.1.1, $c = Db$ for some doubly stochastic matrix D and $a_i = \alpha_i c_i$ for some $0 \le \alpha_i \le 1$. So $a = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n)Db$ and $\mathrm{Diag}(\alpha_1, \ldots, \alpha_n)D$ is a doubly substochastic matrix. Conversely if $a = Sb$ for a doubly substochastic matrix S, then a doubly stochastic matrix D exists so that $S \le D$ entrywise, whose proof is left for the exercise below, and hence $a \le Db \prec b$. □

Proposition 4.1.4. *Let $a, b \in \mathbb{R}^n$.*

(1) *If $a \prec b$ and f is a convex function on an interval containing all a_i, b_i, then $f(a) \prec_w f(b)$, where $f(a) := (f(a_1), \ldots, f(a_n))$.*

(2) *If $a \prec_w b$ and f is a non-decreasing convex function on an interval containing all a_i, b_i, then $f(a) \prec_w f(b)$.*

Proof. (1) If f is a convex function, then so is $(f(x) - r)_+$ for any $r \in \mathbb{R}$. Hence the result follows from (i) $\Rightarrow$ (iii) of Proposition 4.1.1 and (iii) $\Rightarrow$ (i) of Proposition 4.1.3.

(2) If f is a non-decreasing convex function, then so is $(f(x) - r)_+$ for any $r \in \mathbb{R}$. Hence the result follows from (i) $\Rightarrow$ (iv) and (iii) $\Rightarrow$ (i) of Proposition 4.3. □

Exercise 4.1.5. For any doubly substochastic matrix $S = [s_{ij}]_{i,j=1}^n$, show that there exists a doubly stochastic matrix $D = [d_{ij}]_{i,j=1}^n$ such that $s_{ij} \le d_{ij}$ for all $i, j = 1, \ldots, n$.

Let $a, b \in \mathbb{R}^n$ and $a, b \ge 0$. We define the *weak log-majorization* $a \prec_{w(\log)} b$ when

$$\prod_{i=1}^k a_{[i]} \le \prod_{i=1}^k b_{[i]}, \qquad 1 \le k \le n, \tag{4.1.3}$$

and the *log-majorization* $a \prec_{(\log)} b$ when $a \prec_{w(\log)} b$ and equality holds for $k = n$ in (4.1.3). It is obvious that if a and b are strictly positive, then $a \prec_{(\log)} b$ (resp., $a \prec_{w(\log)} b$) if and only if $\log a \prec \log b$ (resp., $\log a \prec_w \log b$), where $\log a := (\log a_1, \ldots, \log a_n)$.

Proposition 4.1.6. *Let $a, b \in \mathbb{R}^n$ with $a, b \ge 0$, and assume that $a \prec_{w(\log)} b$. If f is a continuous non-decreasing function on $[0, \infty)$ such that $f(e^x)$ is convex, then $f(a) \prec_w f(b)$. In particular, $a \prec_{w(\log)} b$ implies $a \prec_w b$.*

Proof. First assume that $a, b \in \mathbb{R}^n$ are strictly positive and $a \prec_{w(\log)} b$, so that $\log a \prec_w \log b$. Thanks to the assumption on f, the function $(f(e^x) - r)_+$ is non-decreasing and convex for any $r \in \mathbb{R}$. Hence we have $f(a) \prec_w f(b)$ by (i) $\Rightarrow$ (iv) and (iii) $\Rightarrow$ (i) of Proposition 4.1.3. When $a, b \ge 0$ and $a \prec_{w(\log)} b$, we can choose $a^{(m)}, b^{(m)} > 0$ such that $a^{(m)} \prec_{w(\log)} b^{(m)}$, $a^{(m)} \to a$, and $b^{(m)} \to b$. Since $f(a^{(m)}) \prec_w f(b^{(m)})$ and f is continuous, we obtain $f(a) \prec_w f(b)$. □

4.2 Singular values of matrices

Let $\mathcal{H}$ is an n-dimensional Hilbert space and $A \in B(\mathcal{H})$. Let $s(A) = (s_1(A), \ldots, s_n(A))$ denote the vector of the *singular values* of A in decreasing order, i.e., $s_1(A) \ge \cdots \ge s_n(A)$ are the eigenvalues of $|A| = (A^*A)^{1/2}$ with counting multiplicities. When A is self-adjoint, the vector of the eigenvalues of A in decreasing order is denoted by $\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A))$. Of course, $s(A) = \lambda(A)$ if $A \ge 0$.

For every $A \in \mathbb{M}_n$, combining the polar decomposition of A (Theorem 1.4.7) and the diagonalization of $|A|$ (Theorem 1.4.6), one has the expression

$$A = U\mathrm{Diag}(s_1(A), \dots, s_n(A))V \tag{4.2.1}$$

with unitary matrices $U, V \in \mathbb{M}_n$, which is called the *singular value decomposition* of A.

The basic properties of $s(A)$ are summarized as follows:

Proposition 4.2.1. *Let $A, B, X, Y \in B(\mathcal{H})$ and $k, m \in \{1, \dots, n\}$.*

(1) $s_1(A) = \|A\|$, *the operator norm of A.*

(2) $s_k(\alpha A) = |\alpha| s_k(A)$ *for $\alpha \in \mathbb{C}$.*

(3) $s_k(A) = s_k(A^*)$.

(4) *Mini-max expression:*

$$s_k(A) = \inf\{\|A(I-P)\| : P \text{ is a projection, } \operatorname{rank} P = k-1\}, \tag{4.2.2}$$

where $\|X\|$ is the operator norm of X and $\operatorname{rank} X := \dim(\operatorname{ran} X)$ for $X \in B(\mathcal{H})$. If $A \geq 0$ then

$$s_k(A) = \inf\left\{\max_{x \in \mathcal{M}^\perp,\, \|x\|=1} \langle x, Ax\rangle : \mathcal{M} \text{ is a subspace of } \mathcal{H}, \dim \mathcal{M} = k-1\right\}. \tag{4.2.3}$$

Furthermore, the inf *in* (4.2.2) *and* (4.2.3) *can be replaced by* min.

(5) *Approximation number expression:*

$$s_k(A) = \inf\{\|A - X\| : X \in B(\mathcal{H}),\ \operatorname{rank} X < k\}. \tag{4.2.4}$$

(6) *If $0 \leq A \leq B$ then $s_k(A) \leq s_k(B)$.*

(7) $s_k(XAY) \leq \|X\|\|Y\| s_k(A)$.

(8) $s_{k+m-1}(A+B) \leq s_k(A) + s_m(B)$ *if $k+m-1 \leq n$.*

(9) $s_{k+m-1}(AB) \leq s_k(A) s_m(B)$ *if $k+m-1 \leq n$.*

(10) $|s_k(A) - s_k(B)| \leq \|A - B\|$.

(11) $s_k(f(A)) = f(s_k(A))$ *if $A \geq 0$ and f is a non-decreasing function on $[0, \infty)$ with $f(0) \geq 0$.*

Proof. Let $A = U|A|$ be the polar decomposition of A (Theorem 1.4.7) and we write the Schmidt decomposition of $|A|$ as

$$|A| = \sum_{i=1}^n s_i(A) |u_i\rangle\langle u_i|$$

(see (1.4.1)), where U is a unitary and $\{u_1, \dots, u_n\}$ is an orthonormal basis of $\mathcal{H}$.

(1) follows since $s_1(A) = \|\,|A|\,\| = \|A\|$. (2) is clear from $|\alpha A| = |\alpha|\,|A|$. Also, (3) immediately follows since the Schmidt decomposition of $|A^*|$ is given as

$$|A^*| = U|A|U^* = \sum_{i=1}^n s_i(A) |Uu_i\rangle\langle Uu_i|.$$

(4) Let α_k be the right-hand side of (4.2.2). For $1 \leq k \leq n$ define $P_k := \sum_{i=1}^k |u_i\rangle\langle u_i|$, which is a projection of rank k. We have

$$\alpha_k \leq \|A(I - P_{k-1})\| = \left\|\sum_{i=k}^n s_i(A)|u_i\rangle\langle u_i|\right\| = s_k(A).$$

Conversely, for any $\varepsilon > 0$ choose a projection P with $\operatorname{rank} P = k-1$ such that $\|A(I-P)\| < \alpha_k + \varepsilon$. Then there exists a $y \in \mathcal{H}$ with $\|y\| = 1$ such that $P_k y = y$ but $Py = 0$. Since $y = \sum_{i=1}^k \langle u_i, y\rangle u_i$, we have

$$\begin{aligned}\alpha_k + \varepsilon &> \|\,|A|(I-P)y\| = \|\,|A|y\| = \left\|\sum_{i=1}^k \langle u_i, y\rangle s_i(A) u_i\right\| \\ &= \left\{\sum_{i=1}^k |\langle u_i, y\rangle|^2 s_i(A)^2\right\}^{1/2} \geq s_k(A).\end{aligned}$$

Hence $s_k(A) = \alpha_k$ and the inf in (4.2.2) is attained by $P = P_{k-1}$.

When $A \geq 0$, we have

$$s_k(A) = s_k(A^{1/2})^2 = \min\{\|A^{1/2}(I-P)\|^2 : P \text{ is a projection, } \operatorname{rank} P = k-1\}.$$

Since $\|A^{1/2}(I-P)\|^2 = \max_{x \in \mathcal{M}^\perp,\, \|x\|=1} \langle x, Ax\rangle$ with $\mathcal{M} := \operatorname{ran} P$, the latter expression follows.

(5) Let β_k be the right-hand side of (4.2.4). Let $X := AP_{k-1}$, where P_{k-1} is as in the above proof of (4). Then we have $\operatorname{rank} X \leq \operatorname{rank} P_{k-1} = k-1$ so that $\beta_k \leq \|A(I - P_{k-1})\| = s_k(A)$. Conversely, assume that $X \in B(\mathcal{H})$ has rank $< k$. Since $\operatorname{rank} X = \operatorname{rank} |X| = \operatorname{rank} X^*$ by Theorem 1.4.7, the projection P onto $\operatorname{ran} X^*$ has rank $< k$. Then $X(I-P) = 0$ and by (4.2.2) we have

$$s_k(A) \le \|A(I-P)\| = \|(A-X)(I-P)\| \le \|A-X\|,$$

implying that $s_k(A) \le \beta_k$.

(6) is an immediate consequence of (4.2.3). It is immediate from (4.2.2) that $s_k(XA) \le \|X\| s_k(A)$. Also $s_k(AY) = s_k(Y^*A^*) \le \|Y\| s_k(A)$ by (3). Hence (7) holds.

Next we show (8)–(10). By (4.2.4), for every $\varepsilon > 0$, there exist $X, Y \in B(\mathcal{H})$ with $\operatorname{rank} X < k$, $\operatorname{rank} Y < m$ such that $\|A - X\| < s_k(A) + \varepsilon$ and $\|B - Y\| < s_m(B) + \varepsilon$. Since $\operatorname{rank}(X + Y) \le \operatorname{rank} X + \operatorname{rank} Y < k + m - 1$, we have

$$s_{k+m-1}(A+B) \le \|(A+B)-(X+Y)\| < s_k(A) + s_m(B) + 2\varepsilon,$$

implying (8). For $Z := XB + (A-X)Y$ we have

$$\operatorname{rank} Z \le \operatorname{rank} X + \operatorname{rank} Y < k + m - 1,$$
$$\|AB - Z\| = \|(A-X)(B-Y)\| < (s_k(A)+\varepsilon)(s_m(B)+\varepsilon).$$

These imply (9). Letting $m = 1$ and replacing B by $B - A$ in (8) we have

$$s_k(B) \le s_k(A) + \|B - A\|,$$

which shows (10).

(11) When $A \ge 0$ has the Schmidt decomposition $A = \sum_{i=1}^n s_i(A)|u_i\rangle\langle u_i|$, we have $f(A) = \sum_{i=1}^n f(s_i(A))|u_i\rangle\langle u_i|$. Since $f(s_1(A)) \ge \cdots \ge f(s_n(A)) \ge 0$, $s_k(f(A)) = f(s_k(A))$ follows. □

The following exercise is the extension of the above (4.2.3) and (6) to self-adjoint $A, B \in B(\mathcal{H})$.

Exercise 4.2.2. When $A \in B(\mathcal{H})$ is self-adjoint, prove the mini-max expression

$$\lambda_k(A) = \min\left\{ \max_{x \in \mathcal{M}^\perp,\, \|x\|=1} \langle x, Ax\rangle : \mathcal{M} \text{ is a subspace of } \mathcal{H},\ \dim \mathcal{M} = k-1 \right\}$$

for $1 \le k \le n$. Hence, if $A, B \in B(\mathcal{H})$ are self-adjoint and $A \le B$, then $\lambda_k(A) \le \lambda_k(B)$ for $1 \le k \le n$.

Exercise 4.2.3. When $A \in B(\mathcal{H})$ is self-adjoint, prove the expression

$$\sum_{i=1}^k \lambda_i(A) = \max\{\operatorname{Tr} AP : P \text{ is a projection, } \operatorname{rank} P = k\}, \qquad 1 \le k \le n.$$

Exercise 4.2.4. Let $\mathfrak{S}(\mathcal{H})$ denote the set of all states on $B(\mathcal{H})$. For each $\omega \in \mathfrak{S}(\mathcal{H})$ let $D_\omega \in B(\mathcal{H})$ be the density operator for ω (see Exercise 1.5.4). For $\omega, \varphi \in \mathfrak{S}(\mathcal{H})$ we write $\omega \prec \varphi$ if $\lambda(D_\omega) \prec \lambda(D_\varphi)$. Prove

(1) $\frac{1}{n}\operatorname{Tr}$ (the tracial state) $\prec \omega \prec \rho$ for all $\omega \in \mathfrak{S}(\mathcal{H})$, where ρ is any pure state.

(2) $\omega \prec \varphi$ if and only if ω belongs to the convex hull of $\{\varphi(U \cdot U^*) : U \in B(\mathcal{H}) \text{ is a unitary}\}$. In this case, it is often said that ω is *more mixed than* φ.

4.3 The Lidskii–Wielandt and the Gelfand–Naimark theorems

The following majorization results are the celebrated *Lidskii–Wielandt theorem* for the eigenvalues of Hermitian matrices as well as for the singular values of general matrices. The first complete proof was obtained by Wielandt [79], where Wielandt's mini-max representation was proved by induction. The proof is contained in [5] and [13]; in fact, [13] contains two more different proofs of the Lidskii–Wielandt theorem, one of which was given in [41] (also found in [37]). All of those proofs are rather involved but a surprisingly elementary and short proof was finally obtained by Li and Mathias [60] as will be given below.

Theorem 4.3.1. *For every Hermitian $n \times n$ matrices A and B,*

$$\lambda(A) - \lambda(B) \prec \lambda(A - B),$$

or equivalently

$$(\lambda_i(A) + \lambda_{n-i+1}(B)) \prec \lambda(A + B).$$

Proof. What we need to prove is that for any choice of $1 \le i_1 < i_2 < \cdots < i_k \le n$ we have

$$\sum_{j=1}^k \{\lambda_{i_j}(A) - \lambda_{i_j}(B)\} \le \sum_{j=1}^k \lambda_j(A - B). \tag{4.3.1}$$

Choose the Schmidt decomposition of $A - B$ as

$$A - B = \sum_{i=1}^{n} \lambda_i(A - B)|u_i\rangle\langle u_i|$$

with an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathbb{C}^n$. We may assume without loss of generality that $\lambda_k(A - B) = 0$. In fact, we may replace B by $B + \lambda_k(A - B)I$, which reduces both sides of (4.3.1) by $k\lambda_k(A - B)$. In this situation, the Jordan decomposition $A - B = (A - B)_+ - (A - B)_-$ is given as

$$(A - B)_+ = \sum_{i=1}^{k} \lambda_i(A - B)|u_i\rangle\langle u_i|, \qquad (A - B)_- = -\sum_{i=k+1}^{n} \lambda_i(A - B)|u_i\rangle\langle u_i|.$$

Since $A = B + (A - B)_+ - (A - B)_- \le B + (A - B)_+$, it follows from Exercise 4.2.2 that

$$\lambda_i(A) \le \lambda_i(B + (A - B)_+), \qquad 1 \le i \le n.$$

Since $B \le B + (A - B)_+$, we also have

$$\lambda_i(B) \le \lambda_i(B + (A - B)_+), \qquad 1 \le i \le n.$$

Hence

$$\begin{aligned}
\sum_{j=1}^{k} \{\lambda_{i_j}(A) - \lambda_{i_j}(B)\} &\le \sum_{j=1}^{k} \{\lambda_{i_j}(B + (A - B)_+) - \lambda_{i_j}(B)\} \\
&\le \sum_{i=1}^{n} \{\lambda_i(B + (A - B)_+) - \lambda_i(B)\} \\
&= \mathrm{Tr}(B + (A - B)_+) - \mathrm{Tr}\, B \quad \text{(by Proposition 1.5.3)} \\
&= \mathrm{Tr}(A - B)_+ = \sum_{j=1}^{k} \lambda_j(A - B),
\end{aligned}$$

proving (4.3.1). Moreover, $\sum_{i=1}^{n} \{\lambda_i(A) - \lambda_i(B)\} = \mathrm{Tr}(A - B) = \sum_{i=1}^{n} \lambda_i(A - B)$.

Replacing B by $-B$ in (4.3.1) gives the latter expression since $\lambda_i(B) = -\lambda_{n-i+1}(-B)$ for $1 \le i \le n$. □

Theorem 4.3.2. *For every $n \times n$ matrices A and B,*

$$|s(A) - s(B)| \prec_w s(A - B),$$

that is,

$$\sum_{j=1}^{k} |s_{i_j}(A) - s_{i_j}(B)| \le \sum_{j=1}^{k} s_j(A - B)$$

for any choice of $1 \le i_1 < i_2 < \cdots < i_k \le n$.

Proof. For every $A, B \in \mathbb{M}_n$ define $\mathbf{A}, \mathbf{B} \in \mathbb{M}_{2n}^{sa}$ by

$$\mathbf{A} := \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}, \qquad \mathbf{B} := \begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix}.$$

Since $\mathbf{A}^*\mathbf{A} = \begin{bmatrix} A^*A & 0 \\ 0 & AA^* \end{bmatrix}$ and hence $|\mathbf{A}| = \begin{bmatrix} |A| & 0 \\ 0 & |A^*| \end{bmatrix}$, it follows from Proposition 4.2.1 (3) that

$$s(\mathbf{A}) = (s_1(A), s_1(A), s_2(A), s_2(A), \ldots, s_n(A), s_n(A)).$$

On the other hand, since

$$\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \mathbf{A} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = -\mathbf{A},$$

we have $\lambda_i(\mathbf{A}) = \lambda_i(-\mathbf{A}) = -\lambda_{2n-i+1}(\mathbf{A})$ for $1 \le i \le 2n$. Hence one can write

$$\lambda(\mathbf{A}) = (\lambda_1, \ldots, \lambda_n, -\lambda_n, \ldots, -\lambda_1),$$

where $\lambda_1 \ge \cdots \ge \lambda_n \ge 0$. Since

$$s(\mathbf{A}) = \lambda(|\mathbf{A}|) = (\lambda_1, \lambda_1, \lambda_2, \lambda_2, \ldots, \lambda_n, \lambda_n),$$

we have $\lambda_i = s_i(A)$ for $1 \le i \le n$ and hence

$$\lambda(\mathbf{A}) = (s_1(A), \ldots, s_n(A), -s_n(A), \ldots, -s_1(A)).$$

Similarly,

$$\lambda(\mathbf{B}) = (s_1(B), \dots, s_n(B), -s_n(B), \dots, -s_1(B)),$$
$$\lambda(\mathbf{A}-\mathbf{B}) = (s_1(A-B), \dots, s_n(A-B), -s_n(A-B), \dots, -s_1(A-B)).$$

Theorem 4.3.1 implies that

$$\lambda(\mathbf{A}) - \lambda(\mathbf{B}) \prec \lambda(\mathbf{A}-\mathbf{B}).$$

Now we note that the components of $\lambda(\mathbf{A}) - \lambda(\mathbf{B})$ are

$$|s_1(A) - s_1(B)|, \dots, |s_n(A) - s_n(B)|, -|s_1(A) - s_1(B)|, \dots, -|s_n(A) - s_n(B)|.$$

Therefore, for any choice of $1 \le i_1 < i_2 < \cdots < i_k \le n$ with $1 \le k \le n$, we have

$$\sum_{j=1}^{k} |s_{i_j}(A) - s_{i_j}(B)| \le \sum_{j=1}^{k} \lambda_j(\mathbf{A}-\mathbf{B}) = \sum_{j=1}^{k} s_j(A-B). \qquad \square$$

The following results due to Ky Fan are consequences of the above theorems, which are weaker versions of the Lidskii–Wielandt theorem.

Corollary 4.3.3.
(1) *For every $n \times n$ Hermitian matrices A and B,*

$$\lambda(A+B) \prec \lambda(A) + \lambda(B).$$

(2) *For every $n \times n$ matrices A and B,*

$$s(A+B) \prec_w s(A) + s(B).$$

Proof. (1) Apply Theorem 4.3.1 to $A+B$ and B. Then

$$\sum_{i=1}^{k} \{\lambda_i(A+B) - \lambda_i(B)\} \le \sum_{i=1}^{k} \lambda_i(A)$$

so that

$$\sum_{i=1}^{k} \lambda_i(A+B) \le \sum_{i=1}^{k} \{\lambda_i(A) + \lambda_i(B)\}.$$

Moreover, $\sum_{i=1}^{n} \lambda_i(A+B) = \mathrm{Tr}(A+B) = \sum_{i=1}^{n} \{\lambda_i(A) + \lambda_i(B)\}$.

(2) Similarly, by Theorem 4.3.2,

$$\sum_{i=1}^{k} |s_i(A+B) - s_i(B)| \le \sum_{i=1}^{k} s_i(A)$$

so that

$$\sum_{i=1}^{k} s_i(A+B) \le \sum_{i=1}^{k} \{s_i(A) + s_i(B)\}. \qquad \square$$

Another important majorization for singular values of matrices is the *Gelfand–Naimark theorem* as follows.

Theorem 4.3.4. *For every $n \times n$ matrices A and B,*

$$(s_i(A)s_{n-i+1}(B)) \prec_{(\log)} s(AB), \tag{4.3.2}$$

or equivalently

$$\prod_{j=1}^{k} s_{i_j}(AB) \le \prod_{j=1}^{k} \{s_j(A)s_{i_j}(B)\} \tag{4.3.3}$$

for any choice of $1 \le i_1 < i_2 < \cdots < i_k \le n$.

Proof. First assume that A and B are invertible. Let $A = U\mathrm{Diag}(s_1, \dots, s_n)V$ be the singular value decomposition (see (4.2.1)) with the singular values $s_1 \ge \cdots \ge s_n > 0$ of A and unitaries U, V. Write $D := \mathrm{Diag}(s_1, \dots, s_n)$. Then $s(AB) = s(UDVB) = s(DVB)$ and $s(B) = s(VB)$, so we may replace A, B by D, VB, respectively. Hence we may assume that $A = D = \mathrm{Diag}(s_1, \dots, s_n)$. Moreover, to prove (4.3.3), it suffices to assume that $s_k = 1$. In fact, when A is replaced by $s_k^{-1}A$, both sides of (4.3.3) are multiplied by same s_k^{-k}. Define $\tilde{A} := \mathrm{Diag}(s_1, \dots, s_k, 1, \dots, 1)$; then $\tilde{A}^2 \ge A^2$ and $\tilde{A}^2 \ge I$. We notice that, for every $i = 1, \dots, n$,

$$\begin{aligned} s_i(AB) = s_i((B^*A^2B)^{1/2}) &= s_i(B^*A^2B)^{1/2} \quad \text{(by Proposition 4.2.1 (11))} \\ &\le s_i(B^*\tilde{A}^2B)^{1/2} \quad \text{(by Proposition 4.2.1 (6))} \\ &= s_i(\tilde{A}B) \end{aligned}$$

and

$$s_i(\tilde{A}B) = s_i(B^*\tilde{A}^2B)^{1/2} \ge s_i(B^*B)^{1/2} = s_i(B).$$

Therefore, for any choice of $1 \le i_1 < \cdots < i_k \le n$, we have

$$\begin{aligned} \prod_{j=1}^k \frac{s_{i_j}(AB)}{s_{i_j}(B)} \le \prod_{j=1}^k \frac{s_{i_j}(\tilde{A}B)}{s_{i_j}(B)} \le \prod_{i=1}^n \frac{s_i(\tilde{A}B)}{s_i(B)} &= \frac{\det|\tilde{A}B|}{\det|B|} \\ &= \frac{\sqrt{\det(B^*\tilde{A}^2B)}}{\sqrt{\det(B^*B)}} = \frac{\det\tilde{A}\cdot|\det B|}{|\det B|} = \det\tilde{A} = \prod_{j=1}^k s_j(A), \end{aligned}$$

proving (4.3.3). By replacing A and B by AB and B^{-1}, respectively, (4.3.3) is rephrased as

$$\prod_{j=1}^k s_{i_j}(A) \le \prod_{j=1}^k \{s_j(AB)s_{i_j}(B^{-1})\}.$$

Since $s_i(B^{-1}) = s_{n-i+1}(B)^{-1}$ for $1 \le i \le n$ as readily verified, the above inequality means that

$$\prod_{j=1}^k \{s_{i_j}(A)s_{n-i_j+1}(B)\} \le \prod_{j=1}^k s_j(AB).$$

Hence (4.3.3) implies (4.3.2) and vice versa (as long as A, B are invertible).

For general $A, B \in \mathbb{M}_n$ choose a sequence of complex numbers $\alpha_l \in \mathbb{C}\setminus(\sigma(A)\cup\sigma(B))$ such that $\alpha_l \to 0$. Since $A_l := A - \alpha_l I$ and $B_l := B - \alpha_l I$ are invertible, (4.3.2) and (4.3.3) hold for those. By Proposition 4.2.2 (10), $s_i(A_l) \to s_i(A)$, $s_i(B_l) \to s_i(B)$ and $s_i(A_lB_l) \to s_i(AB)$ as $l \to \infty$ for $1 \le i \le n$. Hence (4.3.2) and (4.3.3) hold for general A, B. □

An immediate corollary of this theorem is the majorization result due to Horn.

Corollary 4.3.5. *For every matrices A and B,*

$$s(AB) \prec_{(\log)} s(A)s(B),$$

where $s(A)s(B) = (s_i(A)s_i(B))$.

Proof. A special case of (4.3.3) is

$$\prod_{i=1}^k s_i(AB) \le \prod_{i=1}^k \{s_i(A)s_i(B)\}$$

for every $k = 1, \ldots, n$. Moreover,

$$\prod_{i=1}^n s_i(AB) = \det|AB| = \det|A|\cdot\det|B| = \prod_{i=1}^n \{s_i(A)s_i(B)\}. \qquad \square$$

Exercise 4.3.6. Show that another formula equivalent to (4.3.2) and (4.3.3) is

$$\prod_{j=1}^k \{s_{n+1-j}(A)s_{i_j}(B)\} \le \prod_{j=1}^k s_{i_j}(AB)$$

for any choice of $1 \le i_1 < \cdots < i_k \le n$.

The most comprehensive literature on majorization theory for vectors and matrices is Marshall and Olkin's monograph [63]. Andō's two survey articles [5, 6] are the best sources on majorizations for the eigenvalues and the singular values of matrices. The contents of this chapter are mostly based on [37].

We end this section with a brief remark on the famous Horn conjecture that was affirmatively solved just before 2000. The conjecture is related to three real vectors $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$, and $c = (c_1, \ldots, c_n)$. If there are two $n \times n$ Hermitian matrices A and B such that $a = \lambda(A)$, $b = \lambda(B)$, and $c = \lambda(A+B)$, that is, a, b, c are the eigenvalues of $A, B, A+B$, then the three vectors obey many inequalities of the type

$$\sum_{k\in K} c_k \le \sum_{i\in I} a_i + \sum_{j\in J} b_j$$

for certain triples (I, J, K) of subsets of $\{1, \dots, n\}$, including those coming from the Lidskii–Wielandt theorem, together with the obvious equality

$$\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i.$$

Horn [47] proposed the procedure how to produce such triples (I, J, K) and conjectured that all the inequalities obtained in that way is sufficient to characterize a, b, c that are the eigenvalues of Hermitian matrices $A, B, A+B$. This long-standing Horn conjecture was solved by two papers put together, one by Klyachko [49] and the other by Knuston and Tao [50]. More information on this interesting conjecture is found in Fulton [32] and Bhatia [15].

4.4 Symmetric norms

A norm Φ on $\mathbb{R}^n$ is said to be *symmetric* if Φ satisfies

$$\Phi(a_1, a_2, \dots, a_n) = \Phi(\varepsilon_1 a_{\pi(1)}, \varepsilon_2 a_{\pi(2)}, \dots, \varepsilon_n a_{\pi(n)}) \tag{4.4.1}$$

for every $(a_1, \dots, a_n) \in \mathbb{R}^n$ and for any permutation π on $\{1, \dots, n\}$ and $\varepsilon_i = \pm 1$. This condition is equivalently written as

$$\Phi(a) = \Phi(a_1^*, a_2^*, \dots, a_n^*)$$

for $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, where $(a_1^*, \dots, a_n^*)$ is the decreasing rearrangement of $(|a_1|, \dots, |a_n|)$. A symmetric norm is often called a *symmetric gauge function*. Typical examples of symmetric gauge functions on $\mathbb{R}^n$ are the *ℓ_p-norms* Φ_p, $1 \le p \le \infty$, that are defined by

$$\Phi_p(a) := \begin{cases} (\sum_{i=1}^{n} |a_i|^p)^{1/p} & \text{if } 1 \le p < \infty, \\ \max_{1 \le i \le n} |a_i| & \text{if } p = \infty. \end{cases} \tag{4.4.2}$$

Lemma 4.4.1. *Let Φ be a symmetric norm on $\mathbb{R}^n$.*

(1) *If $a = (a_i)$, $b = (b_i) \in \mathbb{R}^n$ and $|a_i| \le |b_i|$ for $1 \le i \le n$, then $\Phi(a) \le \Phi(b)$.*

(2) *Under the normalization $\Phi(1, 0, \dots, 0) = 1$,*

$$\max_{1 \le i \le n} |a_i| \le \Phi(a) \le \sum_{i=1}^{n} |a_i|, \qquad a = (a_i) \in \mathbb{R}^n,$$

that is, Φ_∞ (resp., Φ_1) is the least (resp., greatest) symmetric gauge function.

Proof. (1) In view of (4.4.1) we may show that

$$\Phi(\alpha a_1, a_2, \dots, a_n) \le \Phi(a_1, a_2, \dots, a_n) \quad \text{for } 0 \le \alpha \le 1.$$

This is seen as follows:

$$\begin{aligned} &\Phi(\alpha a_1, a_2, \dots, a_n) \\ &\quad = \Phi\left(\frac{1+\alpha}{2} a_1 + \frac{1-\alpha}{2}(-a_1), \frac{1+\alpha}{2} a_2 + \frac{1-\alpha}{2} a_2, \dots, \frac{1+\alpha}{2} a_n + \frac{1-\alpha}{2} a_n\right) \\ &\quad \le \frac{1+\alpha}{2} \Phi(a_1, a_2, \dots, a_n) + \frac{1-\alpha}{2} \Phi(-a_1, a_2, \dots, a_n) = \Phi(a_1, a_2, \dots, a_n). \end{aligned}$$

(2) Since (4.4.1) and (1) imply that

$$|a_i| = \Phi(a_i, 0, \dots, 0) \le \Phi(a),$$

the first inequality holds. The second follows since

$$\Phi(a) \le \sum_{i=1}^{n} \Phi(a_i, 0, \dots, 0) = \sum_{i=1}^{n} |a_i|. \qquad \square$$

Lemma 4.4.2. *If $a = (a_i)$, $b = (b_i) \in \mathbb{R}^n$ and $(|a_1|, \dots, |a_n|) \prec_w (|b_1|, \dots, |b_n|)$, then $\Phi(a) \le \Phi(b)$.*

Proof. By Proposition 4.1.3 there exists a $c \in \mathbb{R}^n$ such that

$$(|a_1|, \dots, |a_n|) \le c \prec (|b_1|, \dots, |b_n|).$$

Proposition 4.1.1 says that c is a convex combination of coordinate permutations of $(|b_1|, \dots, |b_n|)$. Lemma 4.4.1 (1) and (4.4.1) imply that $\Phi(a) \le \Phi(c) \le \Phi(b)$. $\square$

Let $\mathcal{H}$ be an n-dimensional Hilbert space. A norm $|||\cdot|||$ on $B(\mathcal{H})$ is said to be *unitarily invariant* if

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

for all $A \in B(\mathcal{H})$ and all unitaries $U, V \in B(\mathcal{H})$. A unitarily invariant norm on $B(\mathcal{H})$ is also called a *symmetric norm*. The following fundamental theorem is due to von Neumann [77].

Theorem 4.4.3. *There is a bijective correspondence between symmetric gauge functions Φ on $\mathbb{R}^n$ and unitarily invariant norms $|||\cdot|||$ on $B(\mathcal{H})$ determined by the formula*

$$|||A||| = \Phi(s(A)), \qquad A \in B(\mathcal{H}). \tag{4.4.3}$$

Proof. Assume that Φ is a symmetric gauge function on $\mathbb{R}^n$. Define $|||\cdot|||$ on $B(\mathcal{H})$ by the formula (4.4.3). Let $A, B \in B(\mathcal{H})$. Since $s(A+B) \prec_w s(A) + s(B)$ by Corollary 4.3.3 (2), it follows from Lemma 4.4.2 that

$$|||A+B||| = \Phi(s(A+B)) \le \Phi(s(A)+s(B)) \le \Phi(s(A)) + \Phi(s(B)) = |||A||| + |||B|||.$$

Also it is clear that $|||A||| = 0$ if and only if $s(A) = 0$ or $A = 0$. For $\alpha \in \mathbb{C}$ we have by Proposition 4.2.1 (2)

$$|||\alpha A||| = \Phi(|\alpha| s(A)) = |\alpha|\,|||A|||.$$

Hence $|||\cdot|||$ is a norm on $B(\mathcal{H})$, which is unitarily invariant since $s(UAV) = s(A)$ for all unitaries U, V.

Conversely, assume that $|||\cdot|||$ is a unitarily invariant norm on $B(\mathcal{H})$. Choose an orthonormal basis $\{e_1, \dots, e_n\}$ of $\mathcal{H}$ and define $\Phi : \mathbb{R}^n \to \mathbb{R}$ by

$$\Phi(a) := \left|\left|\left| \sum_{i=1}^n a_i |e_i\rangle\langle e_i| \right|\right|\right|, \qquad a = (a_i) \in \mathbb{R}^n.$$

Then it is immediate to see that Φ is a norm on $\mathbb{R}^n$. For any permutation π on $\{1, \dots, n\}$ and $\varepsilon_i = \pm 1$, one can define unitaries U, V on $\mathcal{H}$ by $Ue_{\pi(i)} = \varepsilon_i e_i$ and $Ve_{\pi(i)} = e_i$, $1 \le i \le n$, so that

$$\begin{aligned}\Phi(a) &= \left|\left|\left| U\left(\sum_{i=1}^n a_{\pi(i)} |e_{\pi(i)}\rangle\langle e_{\pi(i)}|\right) V^* \right|\right|\right| = \left|\left|\left| \sum_{i=1}^n a_{\pi(i)} |Ue_{\pi(i)}\rangle\langle Ve_{\pi(i)}| \right|\right|\right| \\ &= \left|\left|\left| \sum_{i=1}^n \varepsilon_i a_{\pi(i)} |e_i\rangle\langle e_i| \right|\right|\right| = \Phi(\varepsilon_1 a_{\pi(1)}, \varepsilon_2 a_{\pi(2)}, \dots, \varepsilon_n a_{\pi(n)}).\end{aligned}$$

Hence Φ is a symmetric gauge function. For any $A \in B(\mathcal{H})$ let $A = U|A|$ be the polar decomposition of A and $|A| = \sum_{i=1}^n s_i(A) |u_i\rangle\langle u_i|$ be the Schmidt decomposition of $|A|$ with an orthonormal basis $\{u_1, \dots, u_n\}$. We have a unitary V defined by $Ve_i = v_i$, $1 \le i \le n$. Since

$$A = U|A| = UV\left(\sum_{i=1}^n s_i(A) |e_i\rangle\langle e_i|\right) V^*,$$

we have

$$\Phi(s(A)) = \left|\left|\left| \sum_{i=1}^n s_i(A) |e_i\rangle\langle e_i| \right|\right|\right| = \left|\left|\left| UV\left(\sum_{i=1}^n s_i(A) |e_i\rangle\langle e_i|\right) V^* \right|\right|\right| = |||A|||,$$

and so (4.4.3) holds. Therefore, the assertion is obtained. □

The next proposition summarizes properties of unitarily invariant (or symmetric) norms on $B(\mathcal{H})$.

Proposition 4.4.4. *Let $|||\cdot|||$ be a unitarily invariant norm on $B(\mathcal{H})$ corresponding to a symmetric gauge function Φ on $\mathbb{R}^n$, and $A, B, X, Y \in B(\mathcal{H})$. Then*

(1) $|||A||| = |||A^*|||$.

(2) $|||XAY||| \le \|X\|\,\|Y\|\,|||A|||$.

(3) *If $s(A) \prec_w s(B)$ (in particular, if $|A| \le |B|$), then $|||A||| \le |||B|||$.*

(4) *Under the normalization $\Phi(1, 0, \dots, 0) = 1$ (or $|||P||| = 1$ for a projection of rank one), $\|A\| \le |||A||| \le \|A\|_1$, that is, $\|\cdot\|$ (resp., $\|\cdot\|_1$) is the least (resp., greatest) unitarily invariant norm.*

Proof. By the definition (4.4.3), (1) follows from Proposition 4.2.1 (3). By Proposition 4.2.1 (7) and Lemma 4.4.1 (1) we have (2) as

$$|||XAY||| = \Phi(s(XAY)) \le \Phi(\|X\|\,\|Y\| s(A)) = \|X\|\,\|Y\|\,|||A|||.$$

Moreover, (3) and (4) follow from Lemmas 4.4.2 and 4.4.1 (2), respectively. □

For instance, for $1 \le p \le \infty$, we have the unitarily invariant norm $\|\cdot\|_p$ on $B(\mathcal{H})$ corresponding to the ℓ_p-norm Φ_p in (4.4.2), that is, for $A \in B(\mathcal{H})$,

$$\|A\|_p := \Phi_p(s(A)) = \begin{cases} \{\sum_{i=1}^n s_i(A)^p\}^{1/p} = (\mathrm{Tr}\,|A|^p)^{1/p} & \text{if } 1 \le p < \infty, \\ s_1(A) = \|A\| & \text{if } p = \infty. \end{cases}$$

The norm $\|\cdot\|_p$ is called the *Schatten p-norm*. In particular, $\|A\|_1 = \mathrm{Tr}\,|A|$ is the *trace-norm* of A, $\|A\|_2 = (\mathrm{Tr}\,A^*A)^{1/2}$ is the *Hilbert–Schmidt norm* $\|A\|_{\mathrm{HS}}$ introduced in Section 1.5, and $\|A\|_\infty = \|A\|$ is the *operator norm* of A. Even for $0 < p < 1$, we may define $\|A\|_p$ by the same expression as above for $1 \le p < \infty$ while $\|\cdot\|_p$ is not a norm but a quasi-norm for $0 < p < 1$.

Exercise 4.4.5. For any $A \in B(\mathcal{H})$ with the polar decomposition $A = U|A|$ and for any $u \in \mathcal{H}$, prove that

$$|\langle u, Au\rangle| \le \frac{\langle u, |A|u\rangle + \langle u, U|A|U^*u\rangle}{2}.$$

By summing this for $u = u_1, \dots, u_n$ forming an orthonormal basis, show that

$$|\mathrm{Tr}\,A| \le \|A\|_1.$$

Another important class of unitarily invariant norm is the *Ky Fan norm* $\|\cdot\|_{(k)}$ defined by

$$\|A\|_{(k)} := \sum_{i=1}^k s_i(A) \qquad \text{for } k = 1, \dots, n.$$

Obviously, $\|\cdot\|_{(1)}$ is the operator norm and $\|\cdot\|_{(n)}$ is the trace-norm. In the next proposition we give two variational expressions for the Ky Fan norms, which are sometimes quite useful since the Ky Fan norms are essential in majorization and norm inequalities for matrices. The right-hand side of the second expression is known as the *K-functional* in the real interpolation theory.

Proposition 4.4.6. *For any $A \in B(\mathcal{H})$ and for any $k = 1, \dots, n$,*
(1) $\|A\|_{(k)} = \max\{\|AP\|_1 : P$ *is a projection,* $\mathrm{rank}\,P = k\}$,
(2) $\|A\|_{(k)} = \min\{\|X\|_1 + k\|Y\| : A = X + Y\}$.

Proof. (1) For any projection P of rank k, we have

$$\|AP\|_1 = \sum_{i=1}^n s_i(AP) = \sum_{i=1}^k s_i(AP) \le \sum_{i=1}^k s_i(A)$$

by Proposition 4.2.1 (4) and (7). For the converse, take the polar decomposition $A = U|A|$ with a unitary U and the spectral decomposition $|A| = \sum_{i=1}^n s_i(A)P_i$ with mutually orthogonal projections P_i of rank 1. Let $P := \sum_{i=1}^k P_i$. Then

$$\|AP\|_1 = \|U|A|P\|_1 = \left\| \sum_{i=1}^k s_i(A)P_i \right\|_1 = \sum_{i=1}^k s_i(A) = \|A\|_{(k)}.$$

(2) For any decomposition $A = X + Y$, since $s_i(A) \le s_i(X) + \|y\|$ by Proposition 4.1.6 (10), we have

$$\|A\|_{(k)} \le \sum_{i=1}^k s_i(X) + k\|Y\| \le \|X\|_1 + k\|Y\|.$$

Conversely, with the same notations as in the proof of (1), define

$$X := U\sum_{i=1}^k \{s_i(A) - s_k(A)\}P_i,$$

$$Y := U\left\{ s_k(A)\sum_{i=1}^k P_i + \sum_{i=k+1}^n s_i(A)P_i \right\}.$$

Then $X + Y = A$ and

$$\|X\|_1 = \sum_{i=1}^k s_i(A) - ks_k(A), \quad \|Y\| = s_k(A).$$

Hence $\|X\|_1 + k\|Y\| = \sum_{i=1}^k s_i(A)$. □

The following is a modification of the above expression in (1):

$$\|A\|_{(k)} = \max\{|\mathrm{Tr}(UAP)| : U \text{ a unitary}, P \text{ a projection}, \mathrm{rank}\,P = k\}.$$

Here we show the *Hölder inequality* for matrices to illustrate the usefulness of the majorization technique.

Proposition 4.4.7. *Let $0 < p, p_1, p_2 \le \infty$ and $1/p = 1/p_1 + 1/p_2$. Then*

$$\|AB\|_p \le \|A\|_{p_1}\|B\|_{p_2}, \qquad A, B \in B(\mathcal{H}).$$

Proof. Assume that $0 < p_1, p_2 < \infty$, because the result is obvious by Proposition 4.2.1 (7) when $p_1 = \infty$ or $p_2 = \infty$. Since Corollary 4.3.5 implies that

$$(s_i(AB)^p) \prec_{(\log)} (s_i(A)^p s_i(B)^p),$$

it follows from Proposition 4.1.6 that

$$(s_i(AB)^p) \prec_w (s_i(A)^p s_i(B)^p).$$

Since $(p_1/p)^{-1} + (p_2/p)^{-1} = 1$, the usual Hölder inequality for vectors shows that

$$\begin{aligned}\|AB\|_p &= \left\{\sum_{i=1}^n s_i(AB)^p\right\}^{1/p} \le \left\{\sum_{i=1}^n s_i(A)^p s_i(B)^p\right\}^{1/p} \\ &\le \left\{\sum_{i=1}^n s_i(A)^{p_1}\right\}^{1/p_1} \left\{\sum_{i=1}^n s_i(B)^{p_2}\right\}^{1/p_2} \le \|A\|_{p_1}\|B\|_{p_2}.\end{aligned}$$

□

Exercise 4.4.8. Let $0 < p, p_1, p_2 \le \infty$ and $1/p = 1/p_1 + 1/p_2$. Prove the Hölder inequality for vectors (used in the proof of Proposition 4.4.7):

$$\Phi_p(ab) \le \Phi_{p_1}(a)\Phi_{p_2}(b), \qquad a, b \in \mathbb{R}^n,$$

where $ab = (a_i b_i)$ for $a = (b_i)$, $b = (b_i) \in \mathbb{R}^n$.

Exercise 4.4.9. This is a generalization of the Hölder inequality. Let Φ be a symmetric gauge function on $\mathbb{R}^n$ with the corresponding unitarily invariant norm $|||\cdot|||$ on $B(\mathcal{H})$.

(1) Assume that $1 < p < \infty$ and $1/p + 1/q = 1$. Show that

$$\Phi(a_1b_1, \dots, a_nb_n) \le \Phi(|a_1|^p, \dots, |a_n|^p)^{1/p}\Phi(|b_1|^q, \dots, |b_n|^q)^{1/q}, \quad a, b \in \mathbb{R}^n.$$

(2) For every $1 < p < \infty$ define

$$\Phi^{(p)}(a_1, \dots, a_n) := \Phi(|a_1|^p, \dots, |a_n|^p)^{1/p}, \qquad a \in \mathbb{R}^n.$$

Show that $\Phi^{(p)}$ is a symmetric gauge function and the corresponding unitarily invariant norm is $|||\,|\cdot|^p|||^{1/p}$.

(3) Let p, q be as in (1). Show that

$$|||AB||| \le |||\,|A|^p|||^{1/p}|||\,|B|^q|||^{1/q}, \qquad A, B \in B(\mathcal{H}).$$

(Note that when $|||\cdot||| = \|\cdot\|_r$ with $1 \le r < \infty$, the above becomes $\|AB\|_r \le \|A\|_{pr}\|B\|_{qr}$, the Hölder inequality given in Proposition 4.4.7.)

Corresponding to each symmetric gauge function Φ, define $\Phi' : \mathbb{R}^n \to \mathbb{R}$ by

$$\Phi'(b) := \sup\left\{\sum_{i=1}^n a_i b_i : a = (a_i) \in \mathbb{R}^n,\ \Phi(a) \le 1\right\}, \qquad b = (b_i) \in \mathbb{R}^n. \tag{4.4.4}$$

Exercise 4.4.10. Prove that Φ' defined by (4.4.4) is again a symmetric gauge function on $\mathbb{R}^n$. Moreover, prove that $\Phi'' := (\Phi')'$ is equal to Φ.

The symmetric gauge function Φ' is said to be *dual* to Φ. For example, when $1 \le p \le \infty$ and $1/p + 1/q = 1$, the ℓ_p-norm Φ_p is dual to the ℓ_q-norm Φ_q.

The following is another generalized Hölder inequality, which can be shown as Proposition 4.4.7.

Lemma 4.4.11. *Let Φ, Φ_1 and Φ_2 be symmetric gauge functions with the corresponding unitarily invariant norms $|||\cdot|||$, $|||\cdot|||_1$ and $|||\cdot|||_2$ on $B(\mathcal{H})$, respectively. If*

$$\Phi(ab) \le \Phi_1(a)\Phi_2(b), \qquad a, b \in \mathbb{R}^n,$$

then

$$|||AB||| \le |||A|||_1|||B|||_2, \qquad A, B \in B(\mathcal{H}).$$

In particular, if $|||\cdot|||'$ is the unitarily invariant norm corresponding to Φ' dual to Φ, then

$$\|AB\|_1 \le |||A|||\,|||B|||', \qquad A, B \in B(\mathcal{H}).$$

Proof. By Corollary 4.3.5, Proposition 4.1.6, and Lemma 4.4.2, we have

$$\Phi(s(AB)) \le \Phi(s(A)s(B)) \le \Phi_1(s(A))\Phi_2(s(B)) \le |||A|||_1|||B|||_2,$$

showing the first assertion. For the second part, note by definition of Φ' that $\Phi_1(ab) \le \Phi(a)\Phi'(b)$ for $a, b \in \mathbb{R}^n$. □

Theorem 4.4.12. *Let Φ and Φ' be dual symmetric gauge functions on $\mathbb{R}^n$ with the corresponding norms $|||\cdot|||$ and $|||\cdot|||'$ on $B(\mathcal{H})$, respectively. Then $|||\cdot|||$ and $|||\cdot|||'$ are dual with respect to the duality $(A, B) \mapsto \mathrm{Tr}\,AB$ for $A, B \in B(\mathcal{H})$, that is,*

$$|||B|||' = \sup\{|\mathrm{Tr}\,AB| : A \in B(\mathcal{H}),\ |||A||| \le 1\}, \qquad B \in B(\mathcal{H}). \tag{4.4.5}$$

Proof. First note that any linear functional on $B(\mathcal{H})$ is represented as $A \in B(\mathcal{H}) \mapsto \mathrm{Tr}\,AB$ for some $B \in B(\mathcal{H})$. We write $|||B|||^o$ for the right-hand side of (4.4.5). By Exercise 4.4.5 and Lemma 4.4.11 we have

$$|\mathrm{Tr}\,AB| \le \|AB\|_1 \le |||A|||\,|||B|||'$$

so that $|||B|||^o \le |||B|||'$ for all $B \in B(\mathcal{H})$. On the other hand, let $B = V|B|$ be the polar decomposition and $|B| = \sum_{i=1}^n s_i(B)|v_i\rangle\langle v_i|$ be the Schmidt decomposition of $|B|$. For any $a = (a_i) \in \mathbb{R}^n$ with $\Phi(a) \le 1$, let $A := (\sum_{i=1}^n a_i|v_i\rangle\langle v_i|)V^*$. Then $s(A) = s(\sum_{i=1}^n a_i|v_i\rangle\langle v_i|) = (a_1^*, \dots, a_n^*)$, the decreasing rearrangement of $(|a_1|, \dots, |a_n|)$, and hence $|||A||| = \Phi(s(A)) = \Phi(a) \le 1$. Moreover,

$$\begin{aligned}\mathrm{Tr}\,AB &= \mathrm{Tr}\left(\sum_{i=1}^n a_i|v_i\rangle\langle v_i|\right)\left(\sum_{i=1}^n s_i(B)|v_i\rangle\langle v_i|\right)\\ &= \mathrm{Tr}\left(\sum_{i=1}^n a_i s_i(B)|v_i\rangle\langle v_i|\right) = \sum_{i=1}^n a_i s_i(B)\end{aligned}$$

so that

$$\sum_{i=1}^n a_i s_i(B) \le |\mathrm{Tr}\,AB| \le |||A|||\,|||B|||^o \le |||B|||^o.$$

This implies that $|||B|||' = \Phi'(s(B)) \le |||B|||^o$. □

As special cases we have $\|\cdot\|_p' = \|\cdot\|_q$ when $1 \le p \le \infty$ and $1/p + 1/q = 1$.

The close relation between the (log-)majorization and the unitarily invariant norm inequalities is summarized in the following proposition.

Proposition 4.4.13. *Consider the following conditions for $A, B \in B(\mathcal{H})$. Then*

$$\text{(i)} \iff \text{(ii)} \implies \text{(iii)} \iff \text{(iv)} \iff \text{(v)} \iff \text{(vi)}.$$

- (i) $s(A) \prec_{w(\log)} s(B)$;
- (ii) $|||f(|A|)||| \le |||f(|B|)|||$ *for every unitarily invariant norm* $|||\cdot|||$ *and every continuous non-decreasing function* f *on* $[0, \infty)$ *such that* $f(0) \ge 0$ *and* $f(e^x)$ *is convex;*
- (iii) $s(A) \prec_w s(B)$;
- (iv) $\|A\|_{(k)} \le \|B\|_{(k)}$ *for every* $k = 1, \dots, n$;
- (v) $|||A||| \le |||B|||$ *for every unitarily invariant norm* $|||\cdot|||$;
- (vi) $|||f(|A|)||| \le |||f(|B|)|||$ *for every unitarily invariant norm* $|||\cdot|||$ *and every non-decreasing convex function* f *on* $[0, \infty)$ *such that* $f(0) \ge 0$.

Proof. (i) ⇒ (ii). Let f be as in (ii). By Propositions 4.1.6 and 4.2.1 (11) we have

$$s(f(|A|)) = f(s(A)) \prec_w f(s(B)) = s(f(|B|)). \tag{4.4.6}$$

This implies by Proposition 4.4.4 (3) that $|||f(|A|)||| \le |||f(|B|)|||$ for any unitarily invariant norm.

(ii) ⇒ (i). Take $|||\cdot||| = \|\cdot\|_{(k)}$, the Ky Fan norms, and $f(x) = \log(1 + \varepsilon^{-1}x)$ for $\varepsilon > 0$. Then f satisfies the condition in (ii). Since

$$s_i(f(|A|)) = f(s_i(A)) = \log(\varepsilon + s_i(A)) - \log\varepsilon,$$

the inequality $\|f(|A|)\|_{(k)} \le \|f(|B|)\|_{(k)}$ means that

$$\prod_{i=1}^k(\varepsilon + s_i(A)) \le \prod_{i=1}^k(\varepsilon + s_i(B)).$$

Letting $\varepsilon \searrow 0$ gives $\prod_{i=1}^k s_i(A) \le \prod_{i=1}^k s_i(B)$ and hence (i) follows.

(iii) ⇔ (iv) is trivial by definition of $\|\cdot\|_{(k)}$ and (vi) ⇒ (v) ⇒ (iv) is clear. Finally assume (iii) and let f be as in (vi). Proposition 4.1.3 yields (4.4.6) again, so that (vi) follows. Hence (iii) ⇒ (vi) holds. □

By Theorems 4.3.1, 4.3.2 and Proposition 4.4.13 we have:

Corollary 4.4.14.

(1) *For every* $A, B \in \mathbb{M}_n^{sa}$ *and every unitarily invariant norm* $|||\cdot|||$,

$$|||\mathrm{Diag}(\lambda_1(A) - \lambda_1(B), \dots, \lambda_n(A) - \lambda_n(B))||| \le |||A - B|||.$$

In particular,

$$\left\{\sum_{i=1}^n |\lambda_i(A) - \lambda_i(B)|^p\right\}^{1/p} \le \|A - B\|_p, \qquad 1 \le p < \infty,$$

$$\max_{1\le i\le n} |\lambda_i(A) - \lambda_i(B)| \le \|A - B\| \quad (\textit{Weyl's inequality}).$$

(2) *For every* $A, B \in \mathbb{M}_n$ *and every unitarily invariant norm* $|||\cdot|||$,

$$|||\mathrm{Diag}(s_1(A) - s_1(B), \dots, s_n(A) - s_n(B))||| \le |||A - B|||.$$

In particular,

$$\left\{\sum_{i=1}^n |s_i(A) - s_i(B)|^p\right\}^{1/p} \le \|A - B\|_p, \qquad 1 \le p < \infty,$$

$$\max_{1\le i\le n} |s_i(A) - s_i(B)| \le \|A - B\|.$$

Exercise 4.4.15. For every $A, B \in \mathbb{M}_n^{sa}$ and every unitarily invariant norm $|||\cdot|||$, show that

$$|||A - B||| \le |||\mathrm{Diag}(\lambda_1(A) - \lambda_n(B), \lambda_2(A) - \lambda_{n-1}(B), \dots, \lambda_n(A) - \lambda_1(B))|||.$$

We close the section with an exercise containing examples of norms on $\mathbb{M}_n$ that are not unitarily invariant but yet very important.

Exercise 4.4.16.

(1) The numerical radius $w(\cdot)$ is a norm on $B(\mathcal{H})$ as shown in Proposition 1.5.7 (2). Show that $w(\cdot)$ is not unitarily invariant but invariant under unitary conjugation, i.e., $w(UAU^*) = w(A)$ for all $A, U \in B(\mathcal{H})$ with U unitary.

(2) For each $A \in \mathbb{M}_n$ let $\|A\|_{\mathrm{S}}$ denote the norm of A as the Schur multiplication operator, i.e.,

$$\|A\|_{\mathrm{S}} := \sup\left\{\frac{\|A \circ X\|}{\|X\|} : X \in \mathbb{M}_n,\ X \ne 0\right\},$$

where $A \circ X$ is the Schur product (see Section 1.6). Show that $\|\cdot\|_{\mathrm{S}}$ is a norm on $\mathbb{M}_n$ that is even not invariant under unitary conjugation. (It is sometimes quite difficult to compute the exact value of $\|A\|_{\mathrm{S}}$. See Proposition 5.1.4 for a particular result.)

4.5 Majorizations for sums and differences of positive semidefinite matrices

In the first half of this section, we prove the subadditivity (resp., superadditivity) inequality for $f(A+B)$ and $f(A)+f(B)$ when f is a nonnegative concave (resp., convex) function on $[0,\infty)$ and $A, B \in \mathbb{M}_n$ are positive semidefinite. These inequalities are natural matricial counterparts of elementary inequalities $f(a+b) \le f(a)+f(b)$ (resp., $f(a+b) \ge f(a)+f(b)$) for such a function f and scalars $a, b \ge 0$. When f is a nonnegative concave function on $[0,\infty)$, the famous *Rotfel'd inequality* is

$$\mathrm{Tr}\, f(A+B) \le \mathrm{Tr}\{f(A)+f(B)\}$$

for all $A, B \in \mathbb{M}_n^+$. Below, following [9, 25, 75] let us extend this trace inequality as follows:

$$|||f(A+B)||| \le |||f(A)+f(B)||| \tag{4.5.1}$$

for all $A, B \in \mathbb{M}_n^+$ and for any unitarily invariant norm $|||\cdot|||$, or equivalently (see Proposition 4.4.13),

$$\lambda(f(A+B)) \prec_w \lambda(f(A)+f(B)).$$

We begin with the subadditivity inequality due to Ando and Zhan [9] in the case where f is an operator concave function. The proof was substantially simplified by Uchiyama [75] as presented below.

Theorem 4.5.1. *Let* f *be a nonnegative continuous function on* $[0,\infty)$. *If* f *is operator monotone* (*or operator concave, see Corollary 2.5.4*) *on* $[0,\infty)$, *then* (4.5.1) *holds for all* $A, B \in \mathbb{M}_n^+$ *and for any unitarily invariant norm* $|||\cdot|||$.

The main ingredient of the proof is to show the following lemma.

Lemma 4.5.2. *Let g be a nonnegative continuous function on $[0,\infty)$. If g is non-increasing and $xg(x)$ is non-decreasing, then*

$$\lambda((A+B)g(A+B)) \prec_w \lambda(A^{1/2}g(A+B)A^{1/2}+B^{1/2}g(A+B)B^{1/2})$$

for all $A,B\in\mathbb{M}_n^+$.

Proof. Let $\lambda(A+B)=(\lambda_1,\dots,\lambda_n)$ be the eigenvalue vector arranged in decreasing order and $u_1,\dots,u_n$ be the corresponding eigenvectors forming an orthonormal basis of $\mathbb{C}^n$. For $1\le k\le n$ let P_k be the orthogonal projection onto the subspace spanned by $u_1,\dots,u_k$. Since $xg(x)$ is non-decreasing, it follows that

$$\lambda((A+B)g(A+B))=(\lambda_1 g(\lambda_1),\dots,\lambda_n g(\lambda_n)).$$

Hence, what we need to prove is

$$\operatorname{Tr}(A+B)g(A+B)P_k\le \operatorname{Tr}\{A^{1/2}g(A+B)A^{1/2}+B^{1/2}g(A+B)B^{1/2}\}P_k,$$

since the left-hand side is equal to $\sum_{i=1}^k\lambda_i g(\lambda_i)$ and the right-hand side is less than or equal to $\sum_{i=1}^k\lambda_i(A^{1/2}g(A+B)A^{1/2}+B^{1/2}g(A+B)B^{1/2})$. The above inequality immediately follows by summing the following two:

$$\operatorname{Tr} g(A+B)^{1/2}Ag(A+B)^{1/2}P_k\le \operatorname{Tr} A^{1/2}g(A+B)A^{1/2}P_k, \tag{4.5.2}$$

$$\operatorname{Tr} g(A+B)^{1/2}Bg(A+B)^{1/2}P_k\le \operatorname{Tr} B^{1/2}g(A+B)B^{1/2}P_k. \tag{4.5.3}$$

To prove (4.5.2), we write P_k, $H:=g(A+B)$ and $A^{1/2}$ as

$$P_k=\begin{bmatrix} I_{\mathcal{K}} & 0\\ 0 & 0\end{bmatrix},\quad H=\begin{bmatrix} H_1 & 0\\ 0 & H_2\end{bmatrix},\quad A^{1/2}=\begin{bmatrix} A_{11} & A_{12}\\ A_{12}^* & A_{22}\end{bmatrix}$$

in the form of 2×2 block matrices corresponding to the orthogonal decomposition $\mathbb{C}^n=\mathcal{K}\oplus\mathcal{K}^\perp$ with $\mathcal{K}:=P_k\mathbb{C}^n$. Then

$$P_k g(A+B)^{1/2}Ag(A+B)^{1/2}P_k=\begin{bmatrix} H_1^{1/2}A_{11}^2H_1^{1/2}+H_1^{1/2}A_{12}A_{12}^*H_1^{1/2} & 0\\ 0 & 0\end{bmatrix},$$

$$P_kA^{1/2}g(A+B)A^{1/2}P_k=\begin{bmatrix} A_{11}H_1A_{11}+A_{12}H_2A_{12}^* & 0\\ 0 & 0\end{bmatrix}.$$

Since g is non-increasing, we notice that

$$H_1\le g(\lambda_k)I_{\mathcal{K}},\qquad H_2\ge g(\lambda_k)I_{\mathcal{K}^\perp}.$$

Therefore, we have

$$\operatorname{Tr} H_1^{1/2}A_{12}A_{12}^*H_1^{1/2}=\operatorname{Tr} A_{12}^*H_1A_{12}\le g(\lambda_k)\operatorname{Tr} A_{12}^*A_{12}=g(\lambda_k)\operatorname{Tr} A_{12}A_{12}^*\le \operatorname{Tr} A_{12}H_2A_{12}^*$$

so that

$$\operatorname{Tr}(H_1^{1/2}A_{11}^2H_1^{1/2}+H_1^{1/2}A_{12}A_{12}^*H_1^{1/2})\le \operatorname{Tr}(A_{11}H_1A_{11}+A_{12}H_2A_{12}^*),$$

which shows (4.5.2). (4.5.3) is similarly shown. □

Proof of Theorem 4.5.1. By continuity we may assume that $A,B\in\mathbb{M}_n^+$ are invertible. Let $g(x):=f(x)/x$; then g satisfies the assumptions of Lemma 4.5.2. Hence the lemma implies that

$$\begin{aligned}|||f(A+B)|||\le |||&A^{1/2}(A+B)^{-1/2}f(A+B)(A+B)^{-1/2}A^{1/2}\\ &+B^{1/2}(A+B)^{-1/2}f(A+B)(A+B)^{-1/2}B^{1/2}|||.\end{aligned} \tag{4.5.4}$$

Since $C:=A^{1/2}(A+B)^{-1/2}$ is a contraction, Theorem 2.5.2 implies that

$$A^{1/2}(A+B)^{-1/2}f(A+B)(A+B)^{-1/2}A^{1/2}=Cf(A+B)C^*\le f(C(A+B)C^*)=f(A),$$

and similarly

$$B^{1/2}(A+B)^{-1/2}f(A+B)(A+B)^{-1/2}B^{1/2}\le f(B).$$

Therefore, the right-hand side of (4.5.4) is less than or equal to $|||f(A)+f(B)|||$. □

The following superadditivity inequality obtained in [9] is an immediate corollary of Theorem 4.5.1. The particular case where $g(x)=x^m$, i.e., $|||(A+B)^m|||\ge |||A^m+B^m|||$ for any $m\in\mathbb{N}$ was shown by Bhatia and Kittaneh [18].

Corollary 4.5.3. *Let $g:[0,\infty)\to[0,\infty)$ be an increasing bijective function whose inverse function is operator monotone. Then*

$$|||g(A+B)||| \geq |||g(A)+g(B)||| \tag{4.5.5}$$

for all $A, B \in \mathbb{M}_n^+$ *and* $|||\cdot|||$ *as in Theorem 4.5.1.*

Proof. Let f be the inverse function of g. For every $A, B \in \mathbb{M}_n^+$, Theorem 4.5.1 implies that

$$f(\lambda(A+B)) \prec_w \lambda(f(A)+f(B)).$$

Now, replace A and B by $g(A)$ and $g(B)$, respectively. Then we have

$$f(\lambda(g(A)+g(B))) \prec_w \lambda(A+B).$$

Since f is concave and hence g is convex (and increasing), we have by Proposition 4.1.4 (2)

$$\lambda(g(A)+g(B)) \prec_w g(\lambda(A+B)) = \lambda(g(A+B)),$$

which means by Proposition 4.4.13 that $|||g(A)+g(B)||| \leq |||g(A+B)|||$. □

The above corollary can be extended to the next theorem due to Kosem [56], which is the first main result of this section. The simpler proof below is from [25].

Theorem 4.5.4. *Let* $g : [0,\infty) \to [0,\infty)$ *be a continuous convex function with* $g(0) = 0$. *Then* (4.5.5) *holds for all* A, B *and* $|||\cdot|||$ *as above.*

Proof. First, note that a convex function $g \geq 0$ on $[0,\infty)$ with $g(0) = 0$ is non-decreasing. Let Γ denote the set of all nonnegative functions g on $[0,\infty)$ for which the conclusion of the theorem holds. It is obvious that Γ is closed under pointwise convergence and multiplication by nonnegative scalars. When $f, g \in \Gamma$, for the Ky Fan norms $\|\cdot\|_{(k)}$, $1 \leq k \leq n$, and for $A, B \in \mathbb{M}_n^+$ we have

$$\begin{aligned}
\|(f+g)(A+B)\|_{(k)} &= \|f(A+B)\|_{(k)} + \|g(A+B)\|_{(k)} \\
&\geq \|f(A)+f(B)\|_{(k)} + \|g(A)+g(B)\|_{(k)} \\
&\geq \|(f+g)(A)+(f+g)(B)\|_{(k)},
\end{aligned}$$

where the above equality is guaranteed by the non-decreasingness of f, g and the latter inequality is the triangle inequality. Hence $f+g \in \Gamma$ by Proposition 4.4.13 so that Γ is a convex cone. Notice that any convex function $g \geq 0$ on $[0,\infty)$ with $g(0) = 0$ is the pointwise limit of an increasing sequence of functions of the form $\sum_{l=1}^m c_l \gamma_{a_l}(x)$ with $c_l, a_l > 0$, where γ_a is the angle functions at $a > 0$ given as $\gamma_a(x) := \max\{x-a, 0\}$. Hence it suffices to show that $\gamma_a \in \Gamma$ for all $a > 0$. To do this, for $a, r > 0$ we define

$$h_{a,r}(x) := \frac{1}{2}\left\{\sqrt{(x-a)^2+r} + x - \sqrt{a^2+r}\right\}, \qquad x \geq 0,$$

which is an increasing bijective function on $[0,\infty)$ and whose inverse is

$$x - \frac{r/2}{2x+\sqrt{a^2+r}-a} + \frac{\sqrt{a^2+r}+a}{2}. \tag{4.5.6}$$

Since (4.5.6) is operator monotone on $[0,\infty)$, we have $h_{a,r} \in \Gamma$ by Corollary 4.5.3. Therefore, $\gamma_a \in \Gamma$ since $h_{a,r} \to \gamma_a$ as $r \searrow 0$. □

Exercise 4.5.5. Show that the function $h_{a,r}$ defined in the above proof is increasing and bijective on $[0,\infty)$ and that its inverse function is (4.5.6).

The next subadditivity inequality extending Theorem 4.5.1 was proved by Bourin and Uchiyama [25], which is the second main result.

Theorem 4.5.6. *Let* $f : [0,\infty) \to [0,\infty)$ *be a continuous concave function. Then* (4.5.1) *holds for all* A, B *and* $|||\cdot|||$ *as above.*

Proof. Let λ_i and u_i, $1 \leq i \leq n$, be taken as in the proof of Lemma 4.5.2, and P_k, $1 \leq k \leq n$, be also as there. We may prove that

$$\sum_{i=1}^k f(\lambda_i) \leq \sum_{i=1}^k \lambda_i(f(A)+f(B)), \qquad 1 \leq k \leq n.$$

To do this, it suffices to show that

$$\operatorname{Tr} f(A+B)P_k \leq \operatorname{Tr}\{f(A)+f(B)\}P_k. \tag{4.5.7}$$

Indeed, since f is necessarily non-decreasing, the left-hand side of (4.5.7) is $\sum_{i=1}^k f(\lambda_i)$ and the right-hand side is less than or equal to $\sum_{i=1}^k \lambda_i(f(A)+f(B))$ (see Exercise 4.2.3). Here, note by Exercise 4.5.7 that f is the pointwise limit of a sequence of functions of the form $\alpha + \beta x - g(x)$ where $\alpha \geq 0$, $\beta > 0$, and $g \geq 0$ is a continuous convex function on $[0, \infty)$ with $g(0) = 0$. Hence, to prove (4.5.7), it suffices to show that

$$\operatorname{Tr} g(A+B)P_k \geq \operatorname{Tr}\{g(A) + g(B)\}P_k$$

for any continuous convex function $g \geq 0$ on $[0, \infty)$ with $g(0) = 0$. In fact, this is seen as follows:

$$\operatorname{Tr} g(A+B)P_k = \|g(A+B)\|_{(k)} \geq \|g(A) + g(B)\|_{(k)} \geq \operatorname{Tr}\{g(A) + g(B)\}P_k,$$

where the above equality is due to the non-decreasingness of g and the first inequality follows from Theorem 4.5.4. □

Exercise 4.5.7. Show that a continuous concave function $f \geq 0$ on $[0, \infty)$ is the pointwise limit of a sequence of functions of the form $\alpha + \beta x - \sum_{l=1}^{m} c_l \gamma_{a_l}(x)$ with $\alpha \geq 0$ and $\beta, c_l, a_l > 0$, where γ_a is as given in the proof of Theorem 4.5.4.

Exercise 4.5.8. By slightly modifying the above proofs, extend Theorems 4.5.4 and 4.5.6 to any finite number of matrices $A_1, \ldots, A_m \in \mathbb{M}_n^+$. For instance, the subadditivity inequality of Theorem 4.5.4 is extended as

$$|||f(A_1 + \cdots + A_m)||| \leq |||f(A_1) + \cdots + f(A_m)|||$$

if f satisfies the same assumptions as in Theorem 4.5.4. To do this, first extend Lemma 4.5.2 to $A_1, \ldots, A_m$.

Remark 4.5.9. The subadditivity inequality of Theorem 4.5.4 was further extended by Bourin [24] in such a way that if f is a nonnegative continuous concave function on $[0, \infty)$ then

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

for all normal matrices $A, B \in \mathbb{M}_n$ and for any unitarily invariant norm $|||\cdot|||$. In particular,

$$|||f(|Z|)||| \leq |||f(|A|) + f(|B|)|||$$

when $Z = A + iB$ is the Descartes decomposition of Z.

In the second half of the section, we prove the inequality between norms of $f(|A-B|)$ and $f(A) - f(B)$ (or the weak majorization for their singular values) when f is a nonnegative operator monotone function on $[0, \infty)$ and $A, B \in \mathbb{M}_n^+$. This was proved by Ando [4] long before Ando and Zhan [9] for $f(A) + f(B)$ and $f(A+B)$ presented in the first half.

Theorem 4.5.10. *Let f be a nonnegative continuous function on $[0, \infty)$. If f is operator monotone on $[0, \infty)$, then*

$$|||f(A) - f(B)||| \leq |||f(|A-B|)|||$$

for all $A, B \in \mathbb{M}_n^+$ and for any unitarily invariant norm $|||\cdot|||$, or equivalently,

$$s(f(A) - f(B)) \prec_w s(f(|A-B|)). \tag{4.5.8}$$

When $f(x) = x^\theta$ with $0 < \theta < 1$, the weak majorization (4.5.8) was formerly proved by Birman, Koplienko and Solomyak [23], which gives the generalized Powers–Størmer inequality

$$\|A^\theta - B^\theta\|_{p/\theta} \leq \|A - B\|_p^\theta$$

for all $A, B \in \mathbb{M}_n^+$ if $0 < \theta < 1$ and $\theta \leq p \leq \infty$. The case where $\theta = 1/2$ and $p = 1$ is known as the *Powers–Størmer inequality* [68].

We first prepare simple facts to prove the theorem.

Lemma 4.5.11. *For self-adjoint $X, Y \in \mathbb{M}_n$, let $X = X_+ - X_-$ and $Y = Y_+ - Y_-$ be the Jordan decompositions.*

(1) *If $X \leq Y$ then $s_i(X_+) \leq s_i(Y_+)$ for all i.*

(2) *If $s(X_+) \prec_w s(Y_+)$ and $s(X_-) \prec_w s(Y_-)$, then $s(X) \prec_w s(Y)$.*

Proof. (1) Let Q be the support projection of X_+. Since

$$X_+ = QXQ \leq QYQ \leq QY_+Q,$$

we have $s_i(X_+) \leq s_i(QY_+Q) \leq s_i(Y_+)$ by Proposition 4.2.1 (7).

(2) It is rather easy to see that $s(X)$ is the decreasing rearrangement of the combination of $s(X_+)$ and $s(X_-)$. Hence for each $k \in \mathbb{N}$ we can choose $0 \leq m \leq k$ so that

$$\sum_{i=1}^{k} s_i(X) = \sum_{i=1}^{m} s_i(X_+) + \sum_{i=1}^{k-m} s_i(X_-).$$

Hence

$$\sum_{i=1}^{k} s_i(X) \le \sum_{i=1}^{m} s_i(Y_+) + \sum_{i=1}^{k-m} s_i(Y_-) \le \sum_{i=1}^{k} s_i(Y),$$

as desired. □

Proof of Theorem 4.5.10. First assume that $A \ge B \ge 0$ and let $C := A - B \ge 0$. In view of Proposition 4.4.13, it suffices to prove that

$$\|f(B+C) - f(B)\|_{(k)} \le \|f(C)\|_{(k)}, \qquad 1 \le k \le n. \tag{4.5.9}$$

For each $\lambda \in (0, \infty)$ let

$$h_\lambda(x) = \frac{x}{x+\lambda} = 1 - \frac{\lambda}{x+\lambda},$$

which is increasing on $[0, \infty)$ with $h_\lambda(0) = 0$. According to the integral representation (2.7.5) for f with $a, b \ge 0$ and a positive finite measure m on $(0, \infty)$, we have

$$\begin{aligned} s_i(f(C)) &= f(s_i(C)) \\ &= a + bs_i(C) + \int_{(0,\infty)} \frac{s_i(C)(1+\lambda)}{s_i(C)+\lambda}\, dm(\lambda) \\ &= a + bs_i(C) + \int_{(0,\infty)} (1+\lambda) s_i(h_\lambda(C))\, dm(\lambda), \end{aligned}$$

so that

$$\|f(C)\|_{(k)} \ge b\|C\|_{(k)} + \int_{(0,\infty)} (1+\lambda)\|h_\lambda(C)\|_{(k)}\, dm(\lambda). \tag{4.5.10}$$

On the other hand, since

$$f(B+C) = aI + b(B+C) + \int_{(0,\infty)} (1+\lambda) h_\lambda(B+C)\, dm(\lambda)$$

as well as the analogous expression for $f(B)$, we have

$$f(B+C) - f(B) = bC + \int_{(0,\infty)} (1+\lambda)\{h_\lambda(B+C) - h_\lambda(B)\}\, dm(\lambda),$$

so that

$$\|f(B+C) - f(B)\|_{(k)} \le b\|C\|_{(k)} + \int_{(0,\infty)} (1+\lambda)\|h_\lambda(B+C) - h_\lambda(B)\|_{(k)}\, dm(\lambda). \tag{4.5.11}$$

By (4.5.10) and (4.5.11) it suffices for (4.5.9) to show that

$$\|h_\lambda(B+C) - h_\lambda(B)\|_{(k)} \le \|h_\lambda(C)\|_{(k)}, \qquad \lambda \in (0,\infty),\ 1 \le k \le n.$$

As $h_\lambda(x) = h_1(x/\lambda)$, it is enough to show this inequality for the case $\lambda = 1$ since we may replace B and C by $\lambda^{-1}B$ and $\lambda^{-1}C$, respectively. Thus, what remains to prove is the following:

$$\|(B+I)^{-1} - (B+C+I)^{-1}\|_{(k)} \le \|I - (C+I)^{-1}\|_{(k)}, \qquad 1 \le k \le n. \tag{4.5.12}$$

Since

$$(B+I)^{-1} - (B+C+I)^{-1} = (B+I)^{-1/2} h_1((B+I)^{-1/2} C (B+I)^{-1/2})(B+I)^{-1/2}$$

and $\|(B+I)^{-1/2}\| \le 1$, we obtain

$$\begin{aligned} s_i((B+I)^{-1} - (B+C+I)^{-1}) &\le s_i(h_1((B+I)^{-1/2} C (B+I)^{-1/2})) \\ &= h_1(s_i((B+I)^{-1/2} C (B+I)^{-1/2})) \\ &\le h_1(s_i(C)) = s_i(I - (C+I)^{-1}) \end{aligned}$$

by repeated use of Proposition 4.2.1 (7). Therefore, (4.5.12) is proved.

Next, let us prove the assertion in the general case $A, B \ge 0$. Since $0 \le A \le B + (A-B)_+$, it follows that

$$f(A) - f(B) \le f(B + (A-B)_+) - f(B),$$

which implies by Lemma 4.5.11 (1) that

$$\|(f(A) - f(B))_+\|_{(k)} \le \|f(B + (A-B)_+) - f(B)\|_{(k)}.$$

Applying (4.5.9) to $B + (A - B)_+$ and B, we have

$$\|f(B + (A - B)_+) - f(B)\|_{(k)} \le \|f((A - B)_+)\|_{(k)}.$$

Therefore,

$$s((f(A) - f(B))_+) \prec_w s(f((A - B)_+)). \tag{4.5.13}$$

Exchanging the role of A, B gives

$$s((f(A) - f(B))_-) \prec_w s(f((A - B)_-)). \tag{4.5.14}$$

Here, we may assume that $f(0) = 0$ since f can be replaced by $f - f(0)$. Then it is immediate to see that

$$f((A - B)_+)f((A - B)_-) = 0, \quad f((A - B)_+) + f((A - B)_-) = f(|A - B|).$$

Hence $s(f(A) - f(B)) \prec_w s(f(|A - B|))$ follows from (4.5.13) and (4.5.14) thanks to Lemma 4.5.11 (2). □

The following is an immediate corollary of Theorem 4.5.10, whose proof is similar to that of Corollary 4.5.3.

Corollary 4.5.12. *Let $g : [0, \infty) \to [0, \infty)$ be an increasing bijective function whose inverse function is operator monotone. Then*

$$|||g(A) - g(B)||| \ge |||g(|A - B|)|||$$

for all A, B and $||| \cdot |||$ as above.

In [11], Audenaert and Aujla pointed out that Theorem 4.5.10 is not true in the case where $f : [0, \infty) \to [0, \infty)$ is a general continuous concave function and that Corollary 4.5.12 is not true in the case where $g : [0, \infty) \to [0, \infty)$ is a general continuous convex function.

4.6 Majorizations of Golden–Thompson type and complementary Golden–Thompson type

We begin with providing a machinery of antisymmetric tensors, which is quite useful in deriving log-majorization results. Let $\mathcal{H}$ be an n-dimensional Hilbert space as before. For each $k \in \mathbb{N}$ let $\mathcal{H}^{\otimes k}$ denote the k-fold tensor product of $\mathcal{H}$, which is the n^k-dimensional Hilbert space with respect to the inner product defined by

$$\langle x_1 \otimes \cdots \otimes x_k, y_1 \otimes \cdots \otimes y_k \rangle := \prod_{i=1}^{k} \langle x_i, y_i \rangle.$$

For $x_1, \dots, x_k \in \mathcal{H}$ define $x_1 \wedge \cdots \wedge x_k \in \mathcal{H}^{\otimes k}$ by

$$x_1 \wedge \cdots \wedge x_k := \frac{1}{\sqrt{k!}} \sum_{\pi} (\operatorname{sgn} \pi) x_{\pi(1)} \otimes \cdots \otimes x_{\pi(k)}, \tag{4.6.1}$$

where π runs over all permutations on $\{1, \dots, k\}$ and $\operatorname{sgn} \pi = \pm 1$ accordingly as π is even or odd. The subspace of $\mathcal{H}^{\otimes k}$ spanned by $\{x_1 \wedge \cdots \wedge x_k : x_i \in \mathcal{H}\}$ is called the k-fold *antisymmetric tensor product* of $\mathcal{H}$ and denoted by $\mathcal{H}^{\wedge k}$.

Lemma 4.6.1.

(1) *$x_1 \wedge \cdots \wedge x_i \wedge \cdots \wedge x_j \wedge \cdots \wedge x_k = -x_1 \wedge \cdots \wedge x_j \wedge \cdots \wedge x_i \wedge \cdots \wedge x_k$, where x_i and x_j are interchanged for any two distinct i, j. Hence $x_1 \wedge \cdots \wedge x_k = 0$ if $x_i = x_j$ for some distinct i, j.*

(2) *$\langle x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k \rangle = \det[\langle x_i, y_j \rangle]_{i,j=1}^{k}$.*

(3) *$x_1 \wedge \cdots \wedge x_k \ne 0$ if and only if $\{x_1, \dots, x_k\}$ is linearly independent.*

(4) *The linear extension of the map $x_1 \otimes \cdots \otimes x_k \mapsto \frac{1}{\sqrt{k!}} x_1 \wedge \cdots \wedge x_k$ is the projection of $\mathcal{H}^{\otimes k}$ onto $\mathcal{H}^{\wedge k}$.*

(5) *If $\{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathcal{H}$, then $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$ is an orthonormal basis of $\mathcal{H}^{\wedge k}$. Hence $\dim \mathcal{H}^{\wedge k} = \binom{n}{k}$ for $1 \le k \le n$ and $\mathcal{H}^{\wedge k} = \{0\}$ for $k > n$.*

Proof. (1) is obvious by definition (4.6.1). (2) is readily seen as

$$\begin{aligned}
\langle x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k \rangle &= \frac{1}{k!} \sum_{\pi, \tau \in S_k} (\operatorname{sgn} \pi)(\operatorname{sgn} \tau) \prod_{i=1}^{k} \langle x_{\pi(i)}, y_{\tau(i)} \rangle \\
&= \frac{1}{k!} \sum_{\pi, \tau \in S_k} (\operatorname{sgn} \pi\tau^{-1}) \prod_{i=1}^{k} \langle x_{\pi\tau^{-1}(i)}, y_i \rangle \\
&= \sum_{\pi \in S_k} (\operatorname{sgn} \pi) \prod_{i=1}^{k} \langle x_{\pi(i)}, y_i \rangle = \det[\langle x_i, y_j \rangle]_{i,j=1}^{k},
\end{aligned}$$

and (3) follows from (2) since $\{x_1, \dots, x_k\}$ is linearly independent if and only if $\det[\langle x_i, x_j \rangle]_{i,j=1}^{k} \ne 0$.

Let P be the linear operator in question in (4). Repeated use of (1) yields that

$$\begin{aligned}
P^2(x_1 \otimes \cdots \otimes x_k) &= \frac{1}{(k!)^{3/2}} \sum_{\pi} (\operatorname{sgn} \pi) x_{\pi(1)} \wedge \cdots \wedge x_{\pi(k)} \\
&= \frac{1}{(k!)^{3/2}} \sum_{\pi} (\operatorname{sgn} \pi)^2 x_1 \wedge \cdots \wedge x_k \\
&= \frac{1}{\sqrt{k!}} x_1 \wedge \cdots \wedge x_k = P(x_1 \otimes \cdots \otimes x_k).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\langle P(x_1 \otimes \cdots \otimes x_k), y_1 \otimes \cdots \otimes y_k \rangle &= \frac{1}{k!} \sum_{\pi} (\operatorname{sgn} \pi) \prod_{i=1}^{k} \langle x_{\pi(i)}, y_i \rangle \\
&= \frac{1}{k!} \sum_{\pi} (\operatorname{sgn} \pi^{-1}) \prod_{i=1}^{k} \langle x_i, y_{\pi^{-1}(i)} \rangle \\
&= \langle x_1 \otimes \cdots \otimes x_k, P(y_1 \otimes \cdots \otimes y_k) \rangle.
\end{aligned}$$

Hence (4) follows. For the last (5), it is clear that $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$ spans $\mathcal{H}^{\wedge k}$. The orthonormality of this set is immediately seen from (2). $\square$

For each $A \in B(\mathcal{H})$ and $k \in \mathbb{N}$, the k-fold tensor product $A^{\otimes k} \in B(\mathcal{H}^{\otimes k})$ is given as

$$A^{\otimes k}(x_1 \otimes \cdots \otimes x_k) := Ax_1 \otimes \cdots \otimes Ax_k.$$

Since $\mathcal{H}^{\wedge k}$ is invariant for $A^{\otimes k}$, the *antisymmetric tensor power* $A^{\wedge k}$ of A can be defined as $A^{\wedge k} = A^{\otimes k}|_{\mathcal{H}^{\wedge k}}$; in fact,

$$A^{\wedge k}(x_1 \wedge \cdots \wedge x_k) = Ax_1 \wedge \cdots \wedge Ax_k. \tag{4.6.2}$$

In matrix theory $A^{\wedge k}$ is usually called the kth *compound* of A. By Lemma 4.6.1 note that $\mathcal{H}^{\wedge n} = \mathbb{C}$ and the scalar $A^{\wedge n}$ is equal to

$$A^{\wedge n} = \langle e_1 \wedge \cdots \wedge e_n, A^{\wedge n}(e_1 \wedge \cdots \wedge e_n) \rangle = \det[\langle e_i, Ae_j \rangle]_{i,j=1}^n = \det A,$$

where $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathcal{H}$.

Assume that $\mathcal{H} = \mathbb{C}^n$ and $A \in B(\mathbb{C}^n) = \mathbb{M}_n$, and take the orthonormal basis $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$ of $\mathcal{H}^{\wedge k} = (\mathbb{C}^n)^{\wedge k}$ where $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{C}^n$. Then $A^{\wedge k}$ is represented as the $\binom{n}{k} \times \binom{n}{k}$ matrix whose entries are

$$\langle e_{i_1} \wedge \cdots \wedge e_{i_k}, A^{\wedge k}(e_{j_1} \wedge \cdots \wedge e_{j_k}) \rangle = \det[\langle e_{i_l}, Ae_{j_m} \rangle]_{l,m=1}^k = \det A\begin{pmatrix} i_1, \ldots, i_k \\ j_1, \ldots, j_k \end{pmatrix},$$

where $A\binom{i_1,\ldots,i_k}{j_1,\ldots,j_k}$ is the submatrix of A consisting of rows $i_1 < \cdots < i_k$ and columns $j_1 < \cdots < j_k$. This is indeed the usual definition of the kth compound of A in matrix theory.

The following are elementary properties of antisymmetric tensor powers.

Lemma 4.6.2. *Let $X, X_j, Y, A \in B(\mathcal{H})$ and $1 \le k \le n$.*

(1) $(X^*)^{\wedge k} = (X^{\wedge k})^*$.
(2) $(XY)^{\wedge k} = (X^{\wedge k})(Y^{\wedge k})$ (*sometimes called the Binet–Cauchy theorem, see* [63]).
(3) *If* $\|X_j - X\| \to 0$, *then* $\|X_j^{\wedge k} - X^{\wedge k}\| \to 0$.
(4) *If* $A \ge 0$, *then* $A^{\wedge k} \ge 0$ *and* $(A^p)^{\wedge k} = (A^{\wedge k})^p$ *for all* $p > 0$.
(5) $|X|^{\wedge k} = |X^{\wedge k}|$.

Proof. (1) and (2) are the restrictions of the corresponding formulas $(X^*)^{\otimes k} = (X^{\otimes k})^*$ and $(XY)^{\otimes k} = (X^{\otimes k})(Y^{\otimes k})$ to $\mathcal{H}^{\wedge k}$. For (3) it suffices to show the corresponding convergences for $A^{\otimes k}$, which are readily verified. If $A \ge 0$ then $A^{\wedge k} = ((A^{1/2})^{\wedge k})^*((A^{1/2})^{\wedge k}) \ge 0$ by (1) and (2). When p is rational, the second assertion of (4) is immediate from (2). Then (3) implies the assertion for general $p > 0$. Finally (5) follows from (1), (2), and (4). $\square$

The following lemma supplies an important technique in the majorization theory for matrices.

Lemma 4.6.3. *For every $A \in B(\mathcal{H})$ and every $k = 1, \ldots, n$,*

$$\prod_{i=1}^{k} s_i(A) = s_1(A^{\wedge k}) \ (= \|A^{\wedge k}\|).$$

Proof. By Lemma 4.6.2 (5) we may assume that $A \ge 0$. Then there exists an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathcal{H}$ such that $Au_i = s_i(A)u_i$ for all i. Thanks to (4.6.2) we have

$$A^{\wedge k}(u_{i_1} \wedge \cdots \wedge u_{i_k}) = \left\{ \prod_{j=1}^{k} s_{i_j}(A) \right\} u_{i_1} \wedge \cdots \wedge u_{i_k},$$

and so $\{u_{i_1} \wedge \cdots \wedge u_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$ is a complete set of eigenvectors of $A^{\wedge k}$. Hence the assertion follows. □

Before going into the main part of this section, let us prove the *Weyl majorization theorem*, showing the usefulness of the antisymmetric tensor technique.

Theorem 4.6.4. *Let $A \in B(\mathcal{H})$ and $\lambda_1(A), \ldots, \lambda_n(A)$ be the eigenvalues of A arranged as $|\lambda_1(A)| \ge \cdots \ge |\lambda_n(A)|$ with counting algebraic multiplicities. Then*

$$(|\lambda_1(A)|, \ldots, |\lambda_n(A)|) \prec_{w(\log)} s(A),$$

that is,

$$\prod_{i=1}^{k} |\lambda_i(A)| \le \prod_{i=1}^{k} s_i(A), \qquad 1 \le k \le n.$$

Proof. If λ is an eigenvalue of A with algebraic multiplicity m, then there exists a set $\{y_1, \ldots, y_m\}$ of independent vectors such that

$$Ay_j - \lambda y_j \in \operatorname{span}\{y_1, \ldots, y_{j-1}\}, \qquad 1 \le j \le m.$$

Hence one can choose independent vectors $x_1, \ldots, x_n$ such that $Ax_i = \lambda_i(A)x_i + z_i$ with $z_i \in \operatorname{span}\{x_1, \ldots, x_{i-1}\}$ for $1 \le i \le n$. Then it is readily checked that

$$A^{\wedge k}(x_1 \wedge \cdots \wedge x_n) = Ax_1 \wedge \cdots \wedge Ax_k = \left\{ \prod_{i=1}^{k} \lambda_i(A) \right\} x_1 \wedge \cdots \wedge x_n$$

and $x_1 \wedge \cdots \wedge x_n \ne 0$, implying that $\prod_{i=1}^{k} \lambda_i(A)$ is an eigenvalue of $A^{\wedge k}$. Hence Lemma 4.6.3 yields that

$$\left| \prod_{i=1}^{k} \lambda_i(A) \right| \le \|A^{\wedge k}\| = \prod_{i=1}^{k} s_i(A).$$ □

The first main result of this section is the following log-majorization due to Araki [10] (also shown in [78]).

Theorem 4.6.5. *For every $A, B \in B(\mathcal{H})^+$,*

$$s((A^{1/2}BA^{1/2})^r) \prec_{w(\log)} s(A^{r/2}B^rA^{r/2}), \qquad r \ge 1, \tag{4.6.3}$$

or equivalently

$$s((A^{p/2}B^pA^{p/2})^{1/p}) \prec_{w(\log)} s((A^{q/2}B^qA^{q/2})^{1/q}), \qquad 0 < p \le q. \tag{4.6.4}$$

Proof. We can pass to the limit from $A + \varepsilon I$ and $B + \varepsilon I$ as $\varepsilon \searrow 0$ by Proposition 4.2.1 (10). So we may assume that A and B are invertible. First let us show that

$$\|(A^{1/2}BA^{1/2})^r\| \le \|A^{r/2}B^rA^{r/2}\|, \qquad r \ge 1. \tag{4.6.5}$$

To do so, it suffices to show that $A^{r/2}B^rA^{r/2} \le I$ implies $A^{1/2}BA^{1/2} \le I$, equivalently $B^r \le A^{-r}$ implies $B \le A^{-1}$. But this is just the Löwner–Heinz inequality. For every $k = 1, \ldots, n$, since Lemma 4.6.2 shows that

$$((A^{1/2}BA^{1/2})^r)^{\wedge k} = ((A^{\wedge k})^{1/2}(B^{\wedge k})(A^{\wedge k})^{1/2})^r,$$
$$(A^{r/2}B^rA^{r/2})^{\wedge k} = (A^{\wedge k})^{r/2}(B^{\wedge k})^r(A^{\wedge k})^{r/2},$$

it follows from (4.6.5) with $A^{\wedge k}, B^{\wedge k}$ instead of A, B that

$$\|((A^{1/2}BA^{1/2})^r)^{\wedge k}\| \le \|(A^{r/2}B^rA^{r/2})^{\wedge k}\|.$$

This means thanks to Lemma 4.6.3 that

$$\prod_{i=1}^{k} s_i((A^{1/2}BA^{1/2})^r) \le \prod_{i=1}^{k} s_i(A^{r/2}B^rA^{r/2}).$$

Hence (4.6.3) is proved. If we replace A, B by A^p, B^p and take $r = q/p$, then

$$s((A^{p/2}B^pA^{p/2})^{q/p}) \prec_{w(\log)} s(A^{q/2}B^qA^{q/2}),$$

which implies (4.6.4) by Proposition 4.2.1 (11). □

Theorem 4.6.5 and Proposition 4.4.13 yield:

Corollary 4.6.6. *Let $A, B \in B(\mathcal{H})^+$ and $|||\cdot|||$ be any unitarily invariant norm. If f is a continuous non-decreasing function on $[0,\infty)$ such that $f(0) \geq 0$ and $f(e^t)$ is convex, then*

$$|||f((A^{1/2}BA^{1/2})^r)||| \leq |||f(A^{r/2}B^rA^{r/2})|||, \qquad r \geq 1.$$

In particular,

$$|||(A^{1/2}BA^{1/2})^r||| \leq |||A^{r/2}B^rA^{r/2}|||, \qquad r \geq 1.$$

The following convergence lemma is a kind of the *Lie–Trotter formula*. Its usual form is

$$\lim_{N\to\infty}(e^{H/N}e^{K/N})^N = e^{H+K}$$

for self-adjoint operators $H, K \in B(\mathcal{H})$. Concerning the Lie–Trotter formula, the real difficulty appears when H, K are unbounded operators in an infinite-dimensional Hilbert space (see [48] for example), while the finite-dimensional case is easy to show.

Lemma 4.6.7. *For every self-adjoint $H, K \in B(\mathcal{H})$,*

$$\lim_{r\to 0}(e^{rH/2}e^{rK}e^{rH/2})^{1/r} = e^{H+K}.$$

Proof. Since $(e^{rH/2}e^{rK}e^{rH/2})^{1/r} = (e^{-rH/2}e^{-rK}e^{-rH/2})^{-1/r}$, we may consider only the case $r \searrow 0$. For $0 < r < 1$ let $A(r) := e^{rH/2}e^{rK}e^{rH/2}$ and $B(r) := e^{r(H+K)}$, and $1/r = m + s$ with $m = m(r) \in \mathbb{N}$ and $s = s(r) \in [0,1)$. Since

$$\|A(r)\| \leq \|e^{rH/2}\|\,\|e^{rK}\|\,\|e^{rH/2}\| \leq e^{r(\|H\|+\|K\|)}$$

and the same inequality holds for $\|B(r)\|$, we have

$$\|A(r)^{1/r} - A(r)^m\| \leq \|A(r)\|^m\|A(r)^s - I\| \leq e^{\|H\|+\|K\|}\|A(r)^s - I\| \longrightarrow 0$$

and similarly $\|e^{H+K} - B(r)^m\| \to 0$ as $r \searrow 0$. Hence it suffices to prove that $\|A(r)^m - B(r)^m\| \to 0$ as $r \searrow 0$. Since

$$A(r) = \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{rH}{2}\right)^k\sum_{k=0}^{\infty}\frac{(rK)^k}{k!}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{rH}{2}\right)^k = I + r(H+K) + o(r)$$

as well as $B(r) = I + r(H+K) + o(r)$, we have

$$\begin{aligned}\|A(r)^m - B(r)^m\| &\leq m\|A(r) - B(r)\|(\max\{\|A(r)\|, \|B(r)\|\})^{m-1}\\ &\leq \frac{1}{r}\|A(r) - B(r)\|e^{\|H\|+\|K\|} \longrightarrow 0,\end{aligned}$$

as required. □

The next corollary is the *Golden–Thompson inequality* strengthened to the form of log-majorization.

Corollary 4.6.8. *For every self-adjoint $H, K \in B(\mathcal{H})$,*

$$s(e^{H+K}) \prec_{w(\log)} s((e^{rH/2}e^{rK}e^{rH/2})^{1/r}), \qquad r > 0.$$

Hence, for every unitarily invariant norm $|||\cdot|||$,

$$|||e^{H+K}||| \leq |||(e^{rH/2}e^{rK}e^{rH/2})^{1/r}|||, \qquad r > 0,$$

and the above right-hand side decreases to $|||e^{H+K}|||$ as $r \searrow 0$. In particular,

$$|||e^{H+K}||| \leq |||e^{H/2}e^Ke^{H/2}||| \leq |||e^He^K|||. \tag{4.6.6}$$

Proof. The log-majorization follows by letting $p \searrow 0$ in (4.6.4) thanks to the above lemma. The second assertion follows from the first and Proposition 4.4.13. Thanks to Proposition 4.2.1 (3) and Theorem 4.6.5 the second inequality of (4.6.6) is seen as

$$|||e^He^K||| = |||\,|e^Ke^H|\,||| = |||(e^He^{2K}e^H)^{1/2}||| \geq |||e^{H/2}e^Ke^{H/2}|||. \qquad \square$$

The specialization of (4.6.6) to the trace-norm $\|\cdot\|_1$ is the celebrated *Golden–Thompson trace inequality*

$$\operatorname{Tr} e^{H+K} \leq \operatorname{Tr} e^He^K$$

established independently in [34, 72, 73]. It was shown in [71] that $\operatorname{Tr} e^{H+K} \leq \operatorname{Tr}(e^{H/n}e^{K/n})^n$ for every $n \in \mathbb{N}$. The extension (4.6.6) was given in [59, 74]. Also (4.6.6) for the operator norm is known as *Segal's inequality* (see [70, p. 260]).

In the rest of this section we study log-majorizations and norm inequalities involving power operator means $A \#_\alpha B$ for $A, B \in B(\mathcal{H})^+$, where $0 \le \alpha \le 1$ (see Example 3.3.1 (4)). The log-majorization in the next theorem is due to Ando and Hiai [8], which is considered as complementary to Theorem 4.6.5.

Theorem 4.6.9. *For every $A, B \in B(\mathcal{H})^+$,*

$$s(A^r \#_\alpha B^r) \prec_{w(\log)} s((A \#_\alpha B)^r), \qquad r \ge 1, \tag{4.6.7}$$

or equivalently

$$s((A^p \#_\alpha B^p)^{1/p}) \prec_{w(\log)} s((A^q \#_\alpha B^q)^{1/q}), \qquad p \ge q > 0. \tag{4.6.8}$$

Proof. First assume that both A and B are invertible. For every $k = 1, \dots, n$, it is easily verified from Lemma 4.6.2 that

$$(A^r \#_\alpha B^r)^{\wedge k} = (A^{\wedge k})^r \#_\alpha (B^{\wedge k})^r,$$
$$((A \#_\alpha B)^r)^{\wedge k} = ((A^{\wedge k}) \#_\alpha (B^{\wedge k}))^r.$$

So it suffices to show that

$$\|A^r \#_\alpha B^r\| \le \|(A \#_\alpha B)^r\|, \qquad r \ge 1, \tag{4.6.9}$$

because (4.6.7) follows from Lemma 4.6.3 by taking $A^{\wedge k}, B^{\wedge k}$ instead of A, B in (4.6.9). To show (4.6.9), we may prove that $A \#_\alpha B \le I$ implies $A^r \#_\alpha B^r \le I$. When $1 \le r \le 2$, let us write $r = 2 - \varepsilon$ with $0 \le \varepsilon \le 1$. Let $C := A^{-1/2} B A^{-1/2}$. Suppose that $A \#_\alpha B \le I$. Then $C^\alpha \le A^{-1}$ and

$$A \le C^{-\alpha}, \tag{4.6.10}$$

so that thanks to $0 \le \varepsilon \le 1$

$$A^{1-\varepsilon} \le C^{-\alpha(1-\varepsilon)}. \tag{4.6.11}$$

Now we have

$$\begin{aligned} A^r \#_\alpha B^r &= A^{1-\frac{\varepsilon}{2}} \{A^{-1+\frac{\varepsilon}{2}} B \cdot B^{-\varepsilon} \cdot B A^{-1+\frac{\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}} \\ &= A^{1-\frac{\varepsilon}{2}} \{A^{-\frac{1-\varepsilon}{2}} C A^{1/2} (A^{-1/2} C^{-1} A^{-1/2})^\varepsilon A^{1/2} C A^{-\frac{1-\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}} \\ &= A^{1/2} \{A^{1-\varepsilon} \#_\alpha [C(A \#_\varepsilon C^{-1})C]\} A^{1/2} \\ &\le A^{1/2} \{C^{-\alpha(1-\varepsilon)} \#_\alpha [C(C^{-\alpha} \#_\varepsilon C^{-1})C]\} A^{1/2} \end{aligned}$$

by using (4.6.10), (4.6.11), and the joint monotonicity of power means (see Definition 3.1.2 (i)). Since

$$C^{-\alpha(1-\varepsilon)} \#_\alpha [C(C^{-\alpha} \#_\varepsilon C^{-1})C] = C^{-\alpha(1-\varepsilon)(1-\alpha)} [C(C^{-\alpha(1-\varepsilon)} C^{-\varepsilon})C]^\alpha = C^\alpha,$$

we have

$$A^r \#_\alpha B^r \le A^{1/2} C^\alpha A^{1/2} = A \#_\alpha B \le I.$$

Therefore (4.6.7) is proved when $1 \le r \le 2$. When $r > 2$, write $r = 2^m s$ with $m \in \mathbb{N}$ and $1 \le s \le 2$. Repeating the above argument we have

$$\begin{aligned} s(A^r \#_\alpha B^r) &\prec_{w(\log)} s(A^{2^{m-1}s} \#_\alpha B^{2^{m-1}s})^2 \\ &\vdots \\ &\prec_{w(\log)} s(A^s \#_\alpha B^s)^{2^m} \\ &\prec_{w(\log)} s(A \#_\alpha B)^r. \end{aligned}$$

For general $A, B \in B(\mathcal{H})^+$ let $A_\varepsilon := A + \varepsilon I$ and $B_\varepsilon := B + \varepsilon I$ for $\varepsilon > 0$. Since

$$A^r \#_\alpha B^r = \lim_{\varepsilon \searrow 0} A_\varepsilon^r \#_\alpha B_\varepsilon^r \quad \text{and} \quad (A \#_\alpha B)^r = \lim_{\varepsilon \searrow 0} (A_\varepsilon \#_\alpha B_\varepsilon)^r,$$

we have (4.6.7) by the above case and Proposition 4.2.1 (10). Finally, (4.6.8) readily follows from (4.6.7) as in the last part of the proof of Theorem 4.6.5. □

By Theorem 4.6.9 and Proposition 4.4.13 we have:

Corollary 4.6.10. *Let $A, B \in B(\mathcal{H})^+$ and $\|\cdot\|$ be any unitarily invariant norm. If f is a continuous non-decreasing function on $[0, \infty)$ such that $f(0) \ge 0$ and $f(e^t)$ is convex, then*

$$\|f(A^r \#_\alpha B^r)\| \le \|f((A \#_\alpha B)^r)\|, \qquad r \ge 1.$$

In particular,

$$\|A^r \#_\alpha B^r\| \le \|(A \#_\alpha B)^r\|, \qquad r \ge 1.$$

The next exercise is a variant of the Lie–Trotter formula. The proof is a modification of that of Lemma 4.6.7.

Exercise 4.6.11. For every self-adjoint $H, K \in B(\mathcal{H})$,

$$\lim_{r \to 0} (e^{rH} \#_\alpha e^{rK})^{1/r} = e^{(1-\alpha)H + \alpha K}.$$

By Theorem 4.6.9, Exercise 4.6.11 and Proposition 4.4.13, the Golden–Thompson type log-majorization in Corollary 4.6.8 is complemented as follows:

Corollary 4.6.12. *For every self-adjoint $H, K \in B(\mathcal{H})$,*

$$s((e^{rH} \#_\alpha e^{rK})^{1/r}) \prec_{w(\log)} s(e^{(1-\alpha)H + \alpha K}), \qquad r > 0.$$

Hence, for every unitarily invariant norm $|||\cdot|||$,

$$|||(e^{rH} \#_\alpha e^{rK})^{1/r}||| \le |||e^{(1-\alpha)H + \alpha K}|||, \qquad r > 0,$$

and the above left-hand side increases to $|||e^{(1-\alpha)H+\alpha K}|||$ as $r \searrow 0$.

Specializing to trace inequality we have

$$\mathrm{Tr}(e^{rH} \#_\alpha e^{rK})^{1/r} \le \mathrm{Tr}\, e^{(1-\alpha)H + \alpha K}, \qquad r > 0,$$

which was first proved in [42]. The following logarithmic trace inequalities are also known for every $A, B \in B(\mathcal{H})^+$ and every $r > 0$:

$$\frac{1}{r} \mathrm{Tr}\, A \log B^{r/2} A^r B^{r/2} \le \mathrm{Tr}\, A(\log A + \log B) \le \frac{1}{r} \mathrm{Tr}\, A \log A^{r/2} B^r A^{r/2},$$

$$\frac{1}{r} \mathrm{Tr}\, A \log (A^r \# B^r)^2 \le \mathrm{Tr}\, A(\log A + \log B).$$

See [8, 42] for details on these logarithmic trace inequalities.

5. Means for Matrices and Their Norm Inequalities

5.1 Means for matrices and their comparison

For matrices H, K, X with $H, K \ge 0$, the norm inequality

$$|||H^{1/2} X K^{1/2}||| \le \frac{1}{2} |||HX + XK||| \tag{5.1.1}$$

for any unitarily invariant norm $|||\cdot|||$ was established by Bhatia and Davis [16] and is known as the *matrix arithmetic-geometric inequality*. To prove this, the case where $H = K$ and H is diagonal (with eigenvalues $\lambda_1, \dots, \lambda_n$) is essential due to the 2×2 matrix trick (see the discussion after Proposition 5.1.4 below) and the unitary invariance. Then it is plain to see that

$$H^{1/2} X H^{1/2} = \left[\frac{2\sqrt{\lambda_i \lambda_j}}{\lambda_i + \lambda_j} \right] \circ \left(\frac{1}{2} (HX + XH) \right),$$

where $\circ$ means the Schur (or Hadamard) product (see Section 1.6). As shown in [45, 64], the above equality is quite useful to prove (5.1.1), and a crucial point here is the positive semidefiniteness of the multiplier matrix $[2\sqrt{\lambda_i \lambda_j}/(\lambda_i + \lambda_j)]$. The usefulness of this approach was further exemplified in [19, 80] for example. On the other hand, in [52] (see also [38]) Kosaki observed that

$$H^{1/2} X K^{1/2} = \int_{-\infty}^{\infty} H^{it} (HX + XK) K^{-it} \frac{dt}{2\cosh(\pi t)},$$

which immediately implies (5.1.1) since the density function here is positive with total mass $1/2$. The positive semidefiniteness of multiplier matrices in the former approach and the positivity of density functions in the latter are related via the Bochner theorem in Fourier analysis as one can easily imagine, and a systematic study of means for matrices (also for Hilbert space operators) was made in [39] (also [40]) by unifying the two approaches. For further developments in this directions see [55]. The present chapter is a survey on means for matrices mostly based on [39].

In this section we first introduce a certain class of binary means (for positive scalars) in an axiomatic fashion and then obtain a general norm comparison result for the corresponding matrix means, which will play a fundamental role in the rest.

Let $M(x, y)$ be a positive real function on $(0, \infty) \times (0, \infty)$, and the continuity is always assumed. A *symmetric homogeneous mean* is such an M satisfying

(1) $M(x,y) = M(y,x)$,
(2) $M(\alpha x, \alpha y) = \alpha M(x,y)$ for all $\alpha > 0$,
(3) $M(x,y)$ is non-decreasing in x and y,
(4) $\min\{x,y\} \le M(x,y) \le \max\{x,y\}$.

We denote by $\mathfrak{M}$ the set of all such symmetric homogeneous means (for positive scalars). We set $f(x) = M(x,1)$, and then $M(x,y) = yf(x/y)$ by homogeneity and f is a continuous function on $(0,\infty)$ satisfying

(a) $f(x) = xf(x^{-1})$,
(b) f is non-decreasing,
(c) $f(1) = 1$ and $f(x) \le x$ for $x \ge 1$.

Conversely, when such an f is given, M defined by $M(x,y) = yf(x/y)$ belongs to $\mathfrak{M}$. Indeed, (1) follows from (a) and (2) is clear by definition. The properties (a) and (b) imply (3), and (a)–(c) altogether imply (4).

Thanks to the non-decreasingness one can automatically extend the domain of $M \in \mathfrak{M}$ to $[0,\infty) \times [0,\infty)$ as follows:

$$M(0,y) = \lim_{x \searrow 0} M(x,y), \quad M(x,0) = \lim_{y \searrow 0} M(x,y),$$
$$M(0,0) = \lim_{y \searrow 0} M(0,y) = \lim_{x \searrow 0} M(x,0).$$

For $M \in \mathfrak{M}$ and $H, K \in \mathbb{M}_n^+ = B(\mathbb{C}^n)^+$ we consider the mean of the left multiplication L_H by H and the right multiplication R_K by K associated with M. Since L_H and R_K are commuting positive linear operators on the Hilbert space $(\mathbb{M}_n, \langle\cdot,\cdot\rangle_{\mathrm{HS}})$ (see Exercise 1.5.6 (2)), one can define a positive linear operator $M(L_H, R_K)$ on $\mathbb{M}_n$ via functional calculus, which will be denoted simply by $M(H,K)$. More explicitly, if $H = \sum_{i=1}^n \lambda_i P_i$ is the spectral decomposition with the eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding orthogonal projections $P_1, \dots, P_n$ of rank 1 and if $K = \sum_{j=1}^n \mu_j Q_j$ is similarly taken, then $M(H,K)$ is given by

$$M(H,K)X := \sum_{i,j=1}^n M(\lambda_i, \mu_j) P_i X Q_j, \qquad X \in \mathbb{M}_n. \tag{5.1.2}$$

This means that with the diagonalization

$$H = U\mathrm{Diag}(\lambda_1, \dots, \lambda_n)U^*, \quad K = V\mathrm{Diag}(\mu_1, \dots, \mu_n)V^*$$

via unitary matrices U, V, we have

$$M(H,K)X = U([M(\lambda_i, \mu_j)]_{ij} \circ (U^*XV))V^*, \tag{5.1.3}$$

where $[M(\lambda_i, \mu_j)]_{ij}$ is the matrix with the (i,j)-entry $M(\lambda_i, \mu_j)$ and $\circ$ is the Schur product.

Exercise 5.1.1. Show the expression (5.1.3) from definition (5.1.2).

Exercise 5.1.2. Let $M \in \mathfrak{M}$ and define

$$M^{(-)}(x,y) := M(x^{-1}, y^{-1})^{-1}, \qquad x, y > 0.$$

Show that $M^{(-)} \in \mathfrak{M}$ and that, for each $H, K > 0$ in $\mathbb{M}_n$, $M^{(-)}(H^{-1}, K^{-1})$ is the inverse of $M(H,K)$ as operators on M_n. Hence $M(H,K)X = Y$ is equivalent to $M^{(-)}(H^{-1}, K^{-1})Y = X$.

Positive operators on $\mathbb{M}_n$ defined as above (i.e., via the left and right multiplications) were treated in [43, 65] to study certain Riemannian metrics on matrix spaces. In fact, arbitrary nonnegative real functions on $[0,\infty) \times [0,\infty)$ works in the above definition, but the restriction of M to $\mathfrak{M}$ is convenient for our exposition on means for matrices.

When $X = I$ (the identity matrix), the matrix $M(H,K)I$ can be regarded as a certain mean of $H, K \ge 0$, but it is not necessarily positive semidefinite and is different from operator means in the sense of Chapter 3. For instance, for the geometric mean $M(x,y) = \sqrt{xy}$ we have $M(H,K)I = H^{1/2}K^{1/2}$ while the geometric operator mean for $H, K > 0$ is given by $H^{1/2}(H^{-1/2}KH^{-1/2})^{1/2}H^{1/2}$. We will adopt the convention $H^0 = I$ for any $H \ge 0$, and write H^{is} ($s \in \mathbb{R}$) only for $H > 0$. So H^{is} ($s \in \mathbb{R}$) are well-defined and form a continuous one-parameter group of unitary matrices.

Theorem 5.1.3. *For $M, N \in \mathfrak{M}$ the following conditions are equivalent:*

(i) *there exists a symmetric probability measure ν on $\mathbb{R}$ such that*

$$M(H,K)X = \int_{-\infty}^{\infty} H^{is}(N(H,K)X)K^{-is}\, d\nu(s) \tag{5.1.4}$$

for all matrices H, K, X of any size with $H, K > 0$;

(ii) $|||M(H,K)X||| \le |||N(H,K)X|||$ *for all matrices H, K, X of any size with $H, K \ge 0$ and for any unitarily invariant norm $|||\cdot|||$;*

(iii) $\|M(H,H)X\| \le \|N(H,H)X\|$ *for all matrices H, X of any size with $H \ge 0$;*

(iv) *the matrix $[M(x_i, x_j)/N(x_i, x_j)]_{1 \le i,j \le n}$ is positive semidefinite for any $x_1, \dots, x_n > 0$ with any $n \in \mathbb{N}$;*

(v) *the function $M(e^t, 1)/N(e^t, 1)$ is positive definite on $\mathbb{R}$, where the positive definiteness of a real continuous function ϕ on $\mathbb{R}$ means that $[\phi(t_i - t_j)]_{1 \le i,j \le n}$ is positive semidefinite for any $t_1, \dots, t_n \in \mathbb{R}$ with any $n \in \mathbb{N}$.*

In the above, the measure ν in (i) *is the representing one for $M(e^t,1)/N(e^t,1)$ in the Bochner theorem i.e., $M(e^t,1)/N(e^t,1) = \int_{-\infty}^{\infty} e^{its}\,d\nu(s)$.*

Proof. (i) $\Rightarrow$ (ii). The inequality in (ii) is obvious from (i) if $H, K > 0$. For general $H, K \ge 0$ we may take the limit of the inequality for $H+\varepsilon I, K+\varepsilon I$ as $\varepsilon \searrow 0$.

(ii) $\Rightarrow$ (iii) is trivial.

(iii) $\Rightarrow$ (iv). Let $x_1,\dots,x_n > 0$ and $a_{ij} := M(x_i,x_j)/N(x_i,x_j)$. Then $A := [a_{ij}]$ is a Hermitian matrix with the diagonals $a_{ii} = 1$. Applying condition (iii) to $H = \mathrm{Diag}(x_1,\dots,x_n)$ and using (5.1.3), we have $\|A\circ X\| \le \|X\|$ for all $X \in \mathbb{M}_n$. It is immediate to see that

$$\langle X, A\circ Y\rangle_{\mathrm{HS}} = \langle A\circ X, Y\rangle_{\mathrm{HS}}, \qquad X, Y \in \mathbb{M}_n. \tag{5.1.5}$$

Since the operator norm $\|\cdot\|$ and the trace norm $\|\cdot\|_1$ are dual norms of each other with respect to $\langle\cdot,\cdot\rangle_{\mathrm{HS}}$ (see Theorem 4.4.12), we have

$$|\langle X, A\circ Y\rangle_{\mathrm{HS}}| \le \|A\circ X\|\,\|Y\|_1 \le \|X\|\,\|Y\|_1$$

so that $\|A\circ Y\|_1 \le \|Y\|_1$ for all $Y \in \mathbb{M}_n$. We specialize Y to the matrix with all entries equal to 1. Then we obtain $\|A\|_1 \le n$ due to $A\circ Y = A$ and $\|Y\|_1 = n$. Let $\alpha_1,\dots,\alpha_n$ be the real eigenvalues of A (whose nonnegativity is to be shown), and we notice

$$\sum_{i=1}^n |\alpha_i| = \|A\|_1 \le n = \mathrm{Tr}\,A = \sum_{i=1}^n \alpha_i,$$

where $n = \mathrm{Tr}\,A$ follows from $a_{ii} = 1$. This forces all the α_i's to be nonnegative.

(iv) $\Rightarrow$ (v) is immediate from

$$\left[\frac{M(e^{t_i-t_j},1)}{N(e^{t_i-t_j},1)}\right] = \left[\frac{M(e^{t_i},e^{t_j})}{N(e^{t_i},t^{t_j})}\right] \quad \text{for } t_1,\dots,t_n \in \mathbb{R}.$$

(v) $\Rightarrow$ (i). Due to the Bochner theorem there exists a probability measure ν on $\mathbb{R}$ such that $M(e^t,1)/N(e^t,1) = \int_{-\infty}^{\infty} e^{its}\,d\nu(s)$ for $t \in \mathbb{R}$. Since $M(e^t,1)/N(e^t,1) = M(e^{-t},1)/N(e^{-t},1)$, it is clear that ν is a symmetric measure. For $H, K > 0$, with the notations in (5.1.2) we compute

$$\begin{aligned}
M(H,K)X &= \sum_{k,l=1}^n M(\lambda_k,\mu_l)P_kXQ_l \\
&= \sum_{k,l=1}^n \mu_l M(e^{\log\lambda_k - \log\mu_l},1)P_kXQ_l \\
&= \sum_{k,l=1}^n \mu_l N(e^{\log\lambda_k - \log\mu_l},1)\left(\int_{-\infty}^{\infty}\left(\frac{\lambda_k}{\mu_l}\right)^{is} d\nu(s)\right)P_kXQ_l \\
&= \int_{-\infty}^{\infty}\sum_{k,l=1}^n \left(\frac{\lambda_k}{\mu_l}\right)^{is} N(\lambda_k,\mu_l)P_kXQ_l\,d\nu(s) \\
&= \int_{-\infty}^{\infty} H^{is}(N(H,K)X)K^{-is}\,d\nu(s),
\end{aligned}$$

implying (i). $\square$

In the following proposition, we present more established results in the background of the above theorem. For $A \in \mathbb{M}_n$ we define the *Schur multiplication operator* S_A on the Hilbert space $\mathbb{M}_n$ by

$$S_A(X) := A\circ X, \qquad X \in \mathbb{M}_n,$$

and let $\|S_A\|_{(|||\cdot|||,|||\cdot|||)}$ denote the norm of the Schur multiplication by A with respect to a norm $|||\cdot|||$ on $\mathbb{M}_n$, i.e.,

$$\|S_A\|_{(|||\cdot|||,|||\cdot|||)} := \sup_{X\ne 0}\frac{|||A\circ X|||}{|||X|||}.$$

In particular, we write $\|S_A\|_{(\infty,\infty)}$ for the norm of the Schur multiplication with respect to the operator norm $\|\cdot\|$.

Proposition 5.1.4. *Let $A \in \mathbb{M}_n$ and $|||\cdot|||$ be an arbitrary unitarily invariant norm. Then*

(1) $\|S_A\|_{(|||\cdot|||,|||\cdot|||)} \le \|S_A\|_{(\infty,\infty)}$.

(2) *If $A = [a_{ij}]$ is positive semidefinite, then* $\|S_A\|_{(|||\cdot|||,|||\cdot|||)} = \max_{1\le i\le n} a_{ii}$.

Proof. (1) Set $\gamma := \|S_A\|_{(\infty,\infty)}$. Notice as (5.1.5) that

$$\langle X, A\circ Y\rangle_{\mathrm{HS}} = \langle \overline{A}\circ X, Y\rangle_{\mathrm{HS}}, \qquad X, Y \in \mathbb{M}_n,$$

where $\overline{A} := [\overline{a_{ij}}]_{ij}$ for $A = [a_{ij}]_{ij}$. Then, as in the proof of (iii) $\Rightarrow$ (iv) of Theorem 5.1.3, we have

$$\begin{aligned}\|S_A\|_{(\|\cdot\|_1,\|\cdot\|_1)} &= \max\{|\langle X, A\circ Y\rangle_{\mathrm{HS}}| : \|X\|\le 1,\ \|Y\|_1\le 1\}\\ &= \max\{|\langle \overline{A}\circ X, Y\rangle_{\mathrm{HS}}| : \|X\|\le 1,\ \|Y\|_1\le 1\} = \|S_{\overline{A}}\|_{(\infty,\infty)}.\end{aligned}$$

Moreover, notice that $A\mapsto\overline{A}$ is an isometry with respect to $\|\cdot\|$ (indeed, it is so with respect to any unitarily invariant norm), so $\|\overline{A}\circ X\| = \|\overline{\overline{A}\circ X}\| = \|A\circ\overline{X}\|$. Hence $\|S_A\|_{(\|\cdot\|_1,\|\cdot\|_1)} = \|S_{\overline{A}}\|_{(\infty,\infty)} = \gamma$. For any $k=1,\dots,n$ and any decomposition $X=Y+Z$ so that $A\circ X = A\circ Y + A\circ Z$, it follows from Proposition 4.4.6 (2) that

$$\|A\circ X\|_{(k)} \le \|A\circ Y\|_1 + k\|A\circ Z\| \le \gamma(\|Y\|_1 + k\|Z\|).$$

By Proposition 4.4.6 (2) again we have $\|A\circ X\|_{(k)} \le \gamma\|X\|_{(k)}$ for $1\le k\le n$. Thanks to Proposition 4.4.13 this implies that $|||A\circ X||| \le \gamma|||X|||$ for any unitarily invariant norm $|||\cdot|||$, so $\|S_A\|_{(|||\cdot|||,|||\cdot|||)}\le\gamma$.

(2) Assume that $A\ge 0$. Since $S_A:\mathbb{M}_n\to\mathbb{M}_n$ is a positive map by the Schur product theorem (Theorem 1.6.3), Proposition 1.7.6 shows that

$$\|S_A\|_{(\infty,\infty)} = \|S_A(I)\| = \|\mathrm{Diag}(a_{11},\dots,a_{nn})\| = \max_{1\le i\le n} a_{ii}.$$

Hence it follows from (1) that $\|S_A\|_{(|||\cdot|||,|||\cdot|||)} \le \max_{1\le i\le n} a_{ii}$. On the other hand, for the matrix units E_{ii} we have

$$a_{ii}|||E_{ii}||| = |||S_A(E_{ii})||| \le \|S_A\|_{(|||\cdot|||,|||\cdot|||)}|||E_{ii}|||$$

so that $a_{ii}\le\|S_A\|_{(|||\cdot|||,|||\cdot|||)}$ for $1\le i\le n$. $\square$

Together with the famous 2×2 matrix trick the above fact (1) is used to show that (iii) implies (ii) in Theorem 5.1.3. In fact, for any $x_1,\dots,x_n>0$, (iii) implies thanks to the above (1) that

$$\left\|S_{\left[\frac{M(x_i,x_j)}{N(x_i,x_j)}\right]}\right\|_{(|||\cdot|||,|||\cdot|||)} \le \left\|S_{\left[\frac{M(x_i,x_j)}{N(x_i,x_j)}\right]}\right\|_{(\infty,\infty)} \le 1,$$

which shows that $|||M(H,H)X|||\le|||N(H,H)X|||$ for all $H,X\in\mathbb{M}_n$ with $H\ge 0$. For $H,K,X\in\mathbb{M}_n$ with $H,K\ge 0$ consider $\tilde{H} := \begin{bmatrix} H & 0\\ 0 & K\end{bmatrix}$, $\tilde{X} := \begin{bmatrix} 0 & X\\ 0 & 0\end{bmatrix}$ instead of H,X. Since $M(\tilde{H},\tilde{H})\tilde{X} = \begin{bmatrix} 0 & M(H,K)X\\ 0 & 0\end{bmatrix}$, we have

$$|||M(H,K)X||| = |||M(\tilde{H},\tilde{H})\tilde{X}||| \le |||N(\tilde{H},\tilde{H})\tilde{X}||| = |||N(H,K)X|||.$$

Also, the above (2) together with the 2×2 trick shows that (iv) implies (ii) in Theorem 5.1.3.

For $M,N\in\mathfrak{M}$ we write $M\preceq N$ if M,N satisfy the equivalent conditions in Theorem 5.1.3. It is a partial order in $\mathfrak{M}$ and preserved under taking the pointwise limit: If $M_n,N_n\in\mathfrak{M}$ converge pointwise to $M,N\in\mathfrak{M}$ respectively, then one has $M\preceq N$ whenever $M_n\preceq N_n$ for all n. Also, note that $M\preceq N$ is equivalent to $N^{(-)}\preceq M^{(-)}$. Of course, $M\preceq N$ implies the simple order $M\le N$, i.e., $M(x,y)\le N(x,y)$ for all $x,y>0$. Actually $M\preceq N$ is strictly stronger than $M\le N$ as will be seen in examples in Sections 5.2 and 5.3. On the other hand, the next exercise shows that the simple order $M\le N$ is related to an estimate in the Hilbert–Schmidt norm $\|\cdot\|_{\mathrm{HS}}$, which may reveal why inequalities for this norm are easy to hold and sometimes very easy to show.

Exercise 5.1.5. Let M,N be general nonnegative real functions on $[0,\infty)\times[0,\infty)$. Then prove that the following conditions are equivalent:

(i) $\|M(H,K)X\|_{\mathrm{HS}}\le\|N(H,K)X\|_{\mathrm{HS}}$ for all matrices H,K,X with $H,K\ge 0$;
(ii) $M(x,y)\le N(x,y)$ for all $x,y\ge 0$.

A kernel function $M:(0,\infty)\times(0,\infty)\to(0,\infty)$ is called a *positive definite kernel* if $[M(x_i,x_j)]_{1\le i,j\le n}$ is positive semidefinite for any $x_1,\dots,x_n>0$ with any n. If $M,N\in\mathfrak{M}$ satisfies $M\preceq N$ and N is a positive definite kernel, then so is M. This is an immediate consequence of Theorem 5.1.3 (iv) and the Schur product theorem. The next proposition says that the geometric mean G is the largest in the order $\preceq$ among means in $\mathfrak{M}$ that are positive definite kernels.

Proposition 5.1.6. *The following conditions are equivalent for $M\in\mathfrak{M}$:*

(i) *$M(H,H)X\ge 0$ for every $H,X\in B(\mathcal{H})$ with $H,X\ge 0$;*
(ii) *M is a positive definite kernel;*
(iii) *$M\preceq G$.*

If this is the case, then $|||M(H,K)X|||\le\sqrt{\|H\|\,\|K\|}\cdot|||X|||$ for all H,K,X with $H,K\ge 0$ and for any unitarily invariant norm.

Proof. (i) $\Rightarrow$ (ii). Let X be the matrix with all entries equal to 1, and set $H=\mathrm{Diag}(x_1,\dots,x_n)$ with $x_1,\dots,x_n\ge 0$. Then $M(H,H)X = [M(x_i,x_j)]_{1\le i,j\le n}$. Hence (i) implies the positive semidefiniteness of $[M(x_i,x_j)]$.

(ii) $\Leftrightarrow$ (iii). For any $x_1,\dots,x_n>0$ notice that

$$\left[\frac{M(x_i,x_j)}{G(x_i,x_j)}\right] = \mathrm{Diag}(x_1^{-1/2},\dots,x_n^{-1/2})[M(x_i,x_j)]\mathrm{Diag}(x_1^{-1/2},\dots,x_n^{-1/2})$$

and

$$[M(x_i,x_j)] = \mathrm{Diag}(x_1^{1/2},\dots,x_n^{1/2})\left[\frac{M(x_i,x_j)}{G(x_i,x_j)}\right]\mathrm{Diag}(x_1^{1/2},\dots,x_n^{1/2}),$$

which show that $[M(x_i,x_j)] \ge 0$ if and only if $[M(x_i,x_j)/G(x_i,x_j)] \ge 0$. Hence (ii) $\Leftrightarrow$ (iii) holds.

(iii) $\Rightarrow$ (i). Assume (iii), so we have the measure ν representing the ratio $M(e^x,1)/G(e^x,1)$. Then, for $H > 0$ Theorem 5.1.3 implies that

$$M(H,H)X = \int_{-\infty}^{\infty} H^{is}(H^{1/2}XH^{1/2})H^{-is}\,d\nu(s),$$

which is positive semidefinite if so is X. Hence (i) holds for all $H,X \ge 0$ by continuity. Furthermore, by Proposition 4.4.4 (2) we have

$$|||M(H,K)X||| \le |||H^{1/2}XK^{1/2}||| \le \sqrt{\|H\|\,\|K\|}\cdot|||X|||$$

for any unitarily invariant norm. $\square$

When H is a matrix with eigenvalues $x_1,\dots,x_n \ge 0$, $M(H,H)$ is essentially equal to the Schur multiplication by $[M(x_i,x_j)]_{1\le i,j\le n}$ (up to unitary conjugation, see (5.1.3)). So one may consider the property (i) above as a generalization of the Schur product theorem. For $M \in \mathfrak{M}$, as in the proof of (ii) $\Leftrightarrow$ (iii) above, we see also that $1/M$ is a positive definite kernel if and only if $G \preceq M$.

According to Theorem 5.1.3, to obtain a norm inequality between matrix means it is crucial to show the positive definiteness of the related function on $\mathbb{R}$. Among important classes of such functions are the following ratios of hyperbolic functions:

$$\frac{\sinh(at)}{\sinh(bt)},\quad \frac{\cosh(at)}{\cosh(bt)},\qquad 0 \le a < b.$$

Indeed, it is well known that these have the following inverse Fourier transforms:

$$\frac{\sinh(at)}{\sinh(bt)} = \int_{-\infty}^{\infty} e^{its}\,\frac{\sin(\frac{\pi a}{b})}{2b(\cosh(\frac{\pi}{b}s)+\cos(\frac{\pi a}{b}))}\,ds, \tag{5.1.6}$$

$$\frac{\cosh(at)}{\cosh(bt)} = \int_{-\infty}^{\infty} e^{its}\,\frac{\cos(\frac{\pi a}{2b})\cosh(\frac{\pi}{2b}s)}{b(\cosh(\frac{\pi}{b}s)+\cos(\frac{\pi a}{b}))}\,ds. \tag{5.1.7}$$

The function $t/\sinh(\frac{t}{2})$ is also positive definite with the inverse Fourier transform

$$\frac{t}{\sinh(\frac{t}{2})} = \int_{-\infty}^{\infty} e^{its}\,\frac{\pi}{\cosh^2(\pi s)}\,ds. \tag{5.1.8}$$

The proofs of the formulas (5.1.6)–(5.1.8) of Fourier transforms are given in Appendix A.5 for the convenience of the reader.

5.2 Norm inequalities for A-L-G interpolating means

We will apply the general result in the previous section to several typical examples of symmetric homogeneous means, and this method proves quite useful to obtain various norm inequalities refining the matrix arithmetic-geometric mean inequality. Throughout this section and next, let H,K,X be matrices with $H,K \ge 0$.

For $\alpha \in \mathbb{R}$ and $x,y > 0$ we set

$$M_\alpha(x,y) := \begin{cases} \dfrac{\alpha-1}{\alpha}\cdot\dfrac{x^\alpha - y^\alpha}{x^{\alpha-1}-y^{\alpha-1}} & \text{if } x \ne y,\\ x & \text{if } x = y.\end{cases}$$

In particular, we have

$$\begin{aligned}
M_2(x,y) &= A(x,y) := \frac{x+y}{2} && \text{(arithmetic mean)},\\
M_1(x,y) &= L(x,y) := \frac{x-y}{\log x - \log y} && \Big(= \lim_{\alpha\to1} M_\alpha(x,y)\Big)\quad \text{(logarithmic mean)},\\
M_{1/2}(x,y) &= G(x,y) := \sqrt{xy} && \text{(geometric mean)},\\
M_0(x,y) &= \frac{\log x - \log y}{y^{-1}-x^{-1}} && \Big(= \lim_{\alpha\to0} M_\alpha(x,y)\Big),\\
M_{-1}(x,y) &= H(x,y) := \frac{2}{x^{-1}+y^{-1}} && \text{(harmonic mean)}.
\end{aligned}$$

Moreover, we may set

$$M_\infty(x,y) := \max\{x,y\} \quad \Big(= \lim_{\alpha\to\infty} M_\alpha(x,y)\Big),$$
$$M_{-\infty}(x,y) := \min\{x,y\} \quad \Big(= \lim_{\alpha\to-\infty} M_\alpha(x,y)\Big).$$

Thus, $\{M_\alpha\}_{-\infty\le\alpha\le\infty}$ is a one-parameter family of means in $\mathfrak{M}$, which we call *A-L-G interpolating means* due to the above interpolation of the means A, L, and G. For any $x, y > 0$ fixed, $M_\alpha(x,y)$ is continuous in $\alpha \in [-\infty,\infty]$ and $M_\alpha(x,y) \le M_\beta(x,y)$ if $-\infty \le \alpha < \beta \le \infty$. Note that $M_{1-\alpha} = M_\alpha^{(-)}$, i.e., $M_{1-\alpha}(x,y) = M_\alpha(x^{-1},y^{-1})^{-1}$ for any $\alpha \in [-\infty,\infty]$. The next theorem says that if $\alpha < \beta$ then the order $M_\alpha \preceq M_\beta$ actually holds true more strongly than $M_\alpha \le M_\beta$.

Theorem 5.2.1. *If $-\infty \le \alpha < \beta \le \infty$, then $M_\alpha \preceq M_\beta$ and hence*

$$|||M_\alpha(H,K)X||| \le |||M_\beta(H,K)X|||$$

for any unitarily invariant norm $|||\cdot|||$.

Proof. Since $M_{1-\alpha} = M_\alpha^{(-)}$ (in particular, $M_{1/2} = G = G^{(-)}$), we may restrict ourselves to the case $1/2 \le \alpha < \beta \le \infty$. When $1/2 \le \alpha < \beta < \infty$, we have

$$\begin{aligned}\frac{M_\alpha(e^{2t},1)}{M_\beta(e^{2t},1)} &= \frac{(\alpha-1)\beta}{\alpha(\beta-1)} \cdot \frac{(e^{2\alpha t}-1)(e^{2(\beta-1)t}-1)}{(e^{2(\alpha-1)t}-1)(e^{2\beta t}-1)} \\ &= \frac{(\alpha-1)\beta}{\alpha(\beta-1)} \cdot \frac{(e^{\alpha t}-e^{-\alpha t})(e^{(\beta-1)t}-e^{-(\beta-1)t})}{(e^{(\alpha-1)t}-e^{-(\alpha-1)t})(e^{\beta t}-e^{-\beta t})} \\ &= \frac{(\alpha-1)\beta}{\alpha(\beta-1)} \cdot \frac{\sinh(\alpha t)\sinh((\beta-1)t)}{\sinh((\alpha-1)t)\sinh(\beta t)}\end{aligned}$$

with the conventions

$$\frac{\alpha-1}{\sinh((\alpha-1)t)} = \frac{1}{t} \quad \text{for } \alpha = 1 \quad \text{and} \quad \frac{\sinh((\beta-1)t)}{\beta-1} = t \quad \text{for } \beta = 1.$$

If $1/2 \le \alpha < \beta \le 1$, then

$$\frac{M_\alpha(e^{2t},1)}{M_\beta(e^{2t},1)} = \frac{(1-\alpha)\beta}{\alpha(1-\beta)} \cdot \frac{\sinh(\alpha t)}{\sinh(\beta t)} \cdot \frac{\sinh((1-\beta)t)}{\sinh((1-\alpha)t)}$$

is positive definite thanks to (5.1.6) (and (5.1.8) for $\beta = 1$). Hence $M_\alpha \preceq M_\beta$ by Theorem 5.1.3. On the other hand, when $1 < \alpha < \beta < \infty$ we notice that

$$\begin{aligned}&\frac{\sinh(\alpha t)\sinh((\beta-1)t)}{\sinh((\alpha-1)t)\sinh(\beta t)} - 1 \\ &\quad = \frac{\sinh((\alpha-1)t+t)\sinh((\beta-1)t) - \sinh((\alpha-1)t)\sinh((\beta-1)t+t)}{\sinh((\alpha-1)t)\sinh(\beta t)} \\ &\quad = \frac{\sinh t\{\cosh((\alpha-1)t)\sinh((\beta-1)t) - \sinh((\alpha-1)t)\cosh((\beta-1)t)\}}{\sinh((\alpha-1)t)\sinh(\beta t)} \\ &\quad = \frac{\sinh t}{\sinh(\beta t)} \cdot \frac{\sinh((\beta-\alpha)t)}{\sinh((\alpha-1)t)}.\end{aligned}$$

If $1 < \alpha < \beta < 2\alpha - 1$ (hence $0 < \beta - \alpha < \alpha - 1$), then the above expression shows that $M_\alpha(e^{2t},1)/M_\beta(e^{2t},1)$ is positive definite and hence $M_\alpha \preceq M_\beta$. In the general case ($1 < \alpha < \beta < \infty$), we may choose $\alpha = \alpha_0 < \alpha_1 < \cdots < \alpha_m = \beta$ satisfying $\alpha_k < 2\alpha_{k-1} - 1$ ($1 \le k \le m$) to conclude $M_\alpha \preceq M_\beta$. Finally, the result when $1 = \alpha < \beta < \infty$ or $1 < \alpha < \beta = \infty$ can be obtained from the above case by taking the limit as $\alpha \to 1$ or $\beta \to \infty$. □

The means M_α are of particular interest when $\alpha = n/(n-1)$ ($n = 2, 3, \dots$) and when $\alpha = m/(m+1)$ ($m = 1, 2 \dots$). Since

$$M_{n/(n-1)}(x,y) = \frac{1}{n} \cdot \frac{x^{n/(n-1)} - y^{n/(n-1)}}{x^{1/(n-1)} - y^{1/(n-1)}} = \frac{1}{n}\sum_{k=0}^{n-1} x^{k/(n-1)} y^{(n-1-k)/(n-1)},$$
$$M_{m/(m+1)}(x,y) = \frac{1}{m} \cdot \frac{x^{m/(m+1)} - y^{m/(m+1)}}{y^{-1/(m+1)} - x^{-1/(m+1)}} = \frac{1}{m}\sum_{k=1}^{m} x^{k/(m+1)} y^{(m+1-k)/(m+1)},$$

we have

$$M_{n/(n-1)}(H,K)X = \frac{1}{n}\sum_{k=0}^{n-1} H^{k/(n-1)}XK^{(n-1-k)/(n-1)} \qquad (n = 2,3,\ldots), \tag{5.2.1}$$

$$M_{m/(m+1)}(H,K)X = \frac{1}{m}\sum_{k=1}^{m} H^{k/(m+1)}XK^{(m+1-k)/(m+1)} \quad (m = 1,2,\ldots). \tag{5.2.2}$$

On the other hand, since $L(x,y) = \int_0^1 x^t y^{1-t}\,dt$, we have

$$L(H,K)X = \int_0^1 H^t XK^{1-t}\,dt. \tag{5.2.3}$$

Both (5.2.1) and (5.2.2) being considered as Riemann sums for the integral (5.2.3), it is straightforward to see that

$$\begin{aligned}\int_0^1 H^t XK^{1-t}\,dt &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} H^{k/(n-1)}XK^{(n-1-k)/(n-1)}\\ &= \lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^{m} H^{k/(m+1)}XK^{(m+1-k)/(m+1)}.\end{aligned}$$

As a direct consequence of Theorem 5.2.1 we have the next result. Indeed, this was shown in [38] for Hilbert space operators.

Corollary 5.2.2. *The inequalities*

$$\begin{aligned}|||H^{1/2}XK^{1/2}||| &\le \frac{1}{m}\left|\left|\left|\sum_{k=1}^{m} H^{k/(m+1)}XK^{(m+1-k)/(m+1)}\right|\right|\right| \le \left|\left|\left|\int_0^1 H^t XK^{1-t}\,dt\right|\right|\right|\\ &\le \frac{1}{n}\left|\left|\left|\sum_{k=0}^{n-1} H^{k/(n-1)}XK^{(n-1-k)/(n-1)}\right|\right|\right| \le \frac{1}{2}\,|||HX+XK|||\end{aligned}$$

hold for each integers $m \ge 1$ and $n \ge 2$ and for any unitarily invariant norm. Furthermore,

$$\frac{1}{m}\left|\left|\left|\sum_{k=1}^{m} H^{k/(m+1)}XK^{(m+1-k)/(m+1)}\right|\right|\right| \quad \textit{increases to} \quad \left|\left|\left|\int_0^1 H^t XK^{1-t}\,dt\right|\right|\right| \quad \textit{as } m\to\infty$$

and

$$\frac{1}{n}\left|\left|\left|\sum_{k=0}^{n-1} H^{k/(n-1)}XK^{(n-1-k)/(n-1)}\right|\right|\right| \quad \textit{decreases to} \quad \left|\left|\left|\int_0^1 H^t XK^{1-t}\,dt\right|\right|\right| \quad \textit{as } n\to\infty.$$

When $-\infty \le \alpha < \beta \le \infty$, Theorems 5.1.3 and 5.2.1 say that $M_\alpha(e^t,1)/M_\beta(e^t,1)$ is the Fourier transform of some symmetric probability measure $\nu_{\alpha,\beta}$ on $\mathbb{R}$ and the integral expression

$$M_\alpha(H,K)X = \int_{-\infty}^{\infty} H^{is}(M_\beta(H,K)X)K^{-is}\,d\nu_{\alpha,\beta}(s) \tag{5.2.4}$$

holds for every $H,K > 0$ and X. Typical examples are

Corollary 5.2.3. *Assume that $H,K > 0$. For $1/2 < \alpha < \infty$,*

$$H^{1/2}XK^{1/2} = \int_{-\infty}^{\infty} H^{is}(M_\alpha(H,K)X)K^{-is}\,\frac{\sin(\pi\frac{\alpha-1}{\alpha})}{(\alpha-1)(\cosh(\frac{2\pi}{\alpha}s)+\cos(\pi\frac{\alpha-1}{\alpha}))}\,ds. \tag{5.2.5}$$

For $1 < \alpha < \infty$,

$$\begin{aligned}&\int_0^1 H^t XK^{1-t}\,dt\\ &\quad = \int_{-\infty}^{\infty} H^{is}(M_\alpha(H,K)X)K^{-is}\,\frac{\alpha}{2\pi(\alpha-1)}\log\left(\frac{\cosh(\frac{2\pi}{\alpha}s)-\cos(\frac{2\pi}{\alpha})}{\cosh(\frac{2\pi}{\alpha}s)-1}\right)ds.\end{aligned} \tag{5.2.6}$$

Proof. For $1/2 < \alpha < \infty$ the formula (5.1.6) shows that

$$\frac{G(e^t,1)}{M_\alpha(e^t,1)} = \frac{\alpha}{\alpha-1}\cdot\frac{e^{t/2}(e^{(\alpha-1)t}-1)}{e^{\alpha t}-1} = \frac{\alpha}{\alpha-1}\cdot\frac{\sinh(\frac{\alpha-1}{2}t)}{\sinh(\frac{\alpha}{2}t)}$$
$$= \int_{-\infty}^{\infty} e^{its}\frac{\sin(\pi\frac{\alpha-1}{\alpha})}{(\alpha-1)(\cosh(\frac{2\pi}{\alpha}s)+\cos(\pi\frac{\alpha-1}{\alpha}))}\,ds. \tag{5.2.7}$$

When $1<\alpha\le 2$, we compute

$$\frac{L(e^t,1)}{M_\alpha(e^t,1)} = \frac{\alpha}{\alpha-1}\cdot\frac{(e^t-1)(e^{(\alpha-1)t}-1)}{t(e^{\alpha t}-1)} = \frac{2\alpha}{\alpha-1}\cdot\frac{\sinh(\frac{t}{2})\sinh(\frac{\alpha-1}{2}t)}{t\sinh(\frac{\alpha}{2}t)}$$
$$= \frac{2\alpha}{\alpha-1}\int_0^{\frac{\alpha-1}{2}}\frac{\sinh(\frac{t}{2})\cosh(ut)}{\sinh(\frac{\alpha}{2}t)}\,du$$
$$= \frac{\alpha}{\alpha-1}\int_0^{\frac{\alpha-1}{2}}\frac{\sinh((\frac12+u)t)+\sinh((\frac12-u)t)}{\sinh(\frac{\alpha}{2}t)}\,du$$
$$= \frac{1}{\alpha-1}\int_0^{\frac{\alpha-1}{2}}du\int_{-\infty}^{\infty}e^{its}\left(\frac{\sin(\pi\frac{1+2u}{\alpha})}{\cosh(\frac{2\pi}{\alpha}s)+\cos(\pi\frac{1+2u}{\alpha})}\right.$$
$$\left.+\frac{\sin(\pi\frac{1-2u}{\alpha})}{\cosh(\frac{2\pi}{\alpha}s)+\cos(\pi\frac{1-2u}{\alpha})}\right)ds$$
$$= \frac{1}{\alpha-1}\int_{-\infty}^{\infty}e^{its}\,ds\int_{-\frac{\alpha-1}{2}}^{\frac{\alpha-1}{2}}\frac{\sin(\pi\frac{1+2u}{\alpha})}{\cosh(\frac{2\pi}{\alpha}s)+\cos(\pi\frac{1+2u}{\alpha})}\,du$$
$$= \frac{\alpha}{2\pi(\alpha-1)}\int_{-\infty}^{\infty}e^{its}\,ds\int_{-1}^{-\cos(\frac{2\pi}{\alpha})}\frac{1}{\cosh(\frac{2\pi}{\alpha}s)+x}\,dx$$
$$= \frac{\alpha}{2\pi(\alpha-1)}\int_{-\infty}^{\infty}e^{its}\log\frac{\cosh(\frac{2\pi}{\alpha}s)-\cos(\frac{2\pi}{\alpha})}{\cosh(\frac{2\pi}{\alpha}s)-1}\,ds.$$

Here, we have used (5.1.6) and the Fubini theorem in the above fifth and sixth equalities. When $2<\alpha<\infty$, we can modify the above computation and arrive at the same integral. Hence (5.2.5) and (5.2.6) follow from Theorem 5.1.3. □

Note that the density function in (5.2.5) is bounded and continuous while that in (5.2.6) has a singularity at $s=0$. Moreover, from the proof of Theorem 5.2.1, it is easy to see that when $1<\alpha<\beta<\infty$ the representing measure $\nu_{\alpha,\beta}$ in (5.2.4) has an atom with the mass $(\alpha-1)\beta/\alpha(\beta-1)$ at $s=0$ as well as a continuous part represented as the convolution of a finite number of density functions similar to that in (5.2.5).

Example 5.2.4. This is an example due to T. Ando and D. Petz, showing the difference between the two orders $\le$ and $\preceq$. Let $AH := (A+H)/2$, i.e., the average of the arithmetic and the harmonic means, and let us compare it with the geometric mean G. Consider the function

$$f(t) := \frac{G(e^{2t},1)}{AH(e^{2t},1)} = \frac{2}{\cosh t+(\cosh t)^{-1}}.$$

It is clear that $f(t)\le 1$ for all $t\in\mathbb{R}$, so $G\le AH$. However, $G\preceq AH$ fails to hold. The following reasoning is due to H. Kosaki. Suppose that $1/(\cosh t+(\cosh t)^{-1})$ is positive definite; then its product with $1/\cosh t$

$$\frac{1}{\cosh^2 t+1} = \frac{2}{\cosh(2t)+3}$$

is also positive definite. But [19, Theorem 5.1] (also [14, 5.6.6]) says that the above function is not positive definite. It is worth noting that a more general fact was obtained in [55, Theorem 7.10].

5.3 Norm inequalities for Heinz-type means and binomial means

In this section we deal with the following classes of means in $\mathfrak{M}$:

$$A_\alpha(x,y) = A_{1-\alpha}(x,y) := \frac12(x^\alpha y^{1-\alpha}+x^{1-\alpha}y^\alpha)\quad \text{for } 0\le\alpha\le 1,$$
$$B_\alpha(x,y) := \left(\frac{x^\alpha+y^\alpha}{2}\right)^{1/\alpha}\quad \text{for } -\infty\le\alpha\le\infty.$$

Notice that $\{A_\alpha\}_{0\le\alpha\le 1}$ is a family of means in $\mathfrak{M}$ interpolating the arithmetic mean $A_0 = A$ and the geometric mean $A_{1/2} = G$. For matrices $H, K \ge 0$ and X we have

$$A_\alpha(H,K)X = \frac{1}{2}(H^\alpha XK^{1-\alpha} + H^{1-\alpha}XK^\alpha).$$

For $0 < \alpha < \beta \le 1/2$ we have

$$\frac{A_\beta(e^t,1)}{A_\alpha(e^t,1)} = \frac{e^{\beta t} + e^{(1-\beta)t}}{e^{\alpha t} + e^{(1-\alpha)t}} = \frac{\cosh((\frac{1}{2}-\beta)t)}{\cosh((\frac{1}{2}-\alpha)t)},$$

which is positive definite by (5.1.7). This implies

Proposition 5.3.1. *The inequality*

$$|||H^\alpha XK^{1-\alpha} + H^{1-\alpha}XK^\alpha||| \le |||HX + XK||| \tag{5.3.1}$$

holds for any $0 \le \alpha \le 1$ *and for any unitarily invariant norm, and moreover* $|||H^\alpha XK^{1-\alpha} + H^{1-\alpha}XK^\alpha|||$ *is monotonically decreasing in* $\alpha \in [0, 1/2]$.

The following "difference" version is also known:

$$|||H^\alpha XK^{1-\alpha} - H^{1-\alpha}XK^\alpha||| \le |2\alpha - 1| \cdot |||HX - XK|||. \tag{5.3.2}$$

The inequalities (5.3.1) and (5.3.2) for the operator norm were formerly shown by Heinz [36], so we call A_α's *Heinz-type means*. Those inequalities for unitarily invariant norms were given in [16, 17]. See [52] for the proof based on the Poisson integral formula and [19] (also [64]) for the above lines of proof based on the Schur multiplier (as in the next exercise). Furthermore, the asymmetric $H^\alpha XK^{1-\alpha}$ can be treated (though it does not fit the setting described in Section 5.1) as discussed in [39, 52], and a further development on the subject is found in [55, Chapter 4].

Exercise 5.3.2. Prove (5.3.2) in the following way:

(1) When $1/2 < \alpha < 1$ and $H = \mathrm{Diag}(x_1, \dots, x_n)$ with $x_i > 0$, show that

$$H^\alpha XH^{1-\alpha} - H^{1-\alpha}XH^\alpha = \left[x_i^{1-\alpha}\, \frac{x_i^{2\alpha-1} - x_j^{2\alpha-1}}{x_i - x_j}\, x_j^{1-\alpha} \right] \circ (HX - XH),$$

where $(x_i^{2\alpha-1} - x_j^{2\alpha-1})/(x_i - x_j)$ is understood as $2\alpha - 1$ if $x_i = x_j$.

(2) When $1/2 < \alpha < 1$ and $x_i > 0$, apply (5.1.6) to prove that $[(x_i^{2\alpha-1} - x_j^{2\alpha-1})/(x_i - x_j)]$ is positive semidefinite.

(3) Use Proposition 5.1.4 (2), the 2×2 matrix trick, and continuity to prove (5.3.2).

Proposition 5.3.3. *The inequalities*

$$\begin{aligned} \frac{1}{1-2\alpha} \left|\left|\left| \int_\alpha^{1-\alpha} H^t XK^{1-t}\, dt \right|\right|\right| &\le \frac{1}{2} |||H^\alpha XK^{1-\alpha} + H^{1-\alpha}XK^\alpha||| \\ &\le \frac{1}{2\alpha} \left|\left|\left| \int_0^\alpha (H^t XK^{1-t} + H^{1-t}XK^t)\, dt \right|\right|\right| \end{aligned} \tag{5.3.3}$$

hold for any $\alpha \in (0, 1/2)$ *and for any unitarily invariant norm. Moreover, each of the above three terms is monotonically decreasing in* $\alpha \in (0, 1/2)$.

Proof. For $0 < \alpha < 1/2$ define L_α and $\tilde{L}_\alpha$ in $\mathfrak{M}$ by

$$L_\alpha(x,y) = \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} A_t(x,y)\, dt, \quad \tilde{L}_\alpha(x,y) = \frac{1}{\alpha} \int_0^\alpha A_t(x,y)\, dt$$

so that the first and third terms of (5.3.3) are $|||L_\alpha(H,K)X|||$ and $|||\tilde{L}_\alpha(H,K)X|||$, respectively. The decreasingness of the second term of (5.3.3) was seen in Proposition 5.3.1, from which the first inequality is easy:

$$|||L_\alpha(H,K)X||| \le \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} |||A_t(H,K)X|||\, dt \le |||A_\alpha(H,K)X|||.$$

The second inequality follows from the positive definiteness of the function

$$\begin{aligned} \frac{A_\alpha(e^{2t},1)}{\tilde{L}_\alpha(e^{2t},1)} &= \frac{\alpha(e^{2\alpha t} + e^{2(1-\alpha)t})}{\int_0^\alpha (e^{2ut} + e^{2(1-u)t})\, du} = \frac{2\alpha t(e^{2\alpha t} + e^{2(1-\alpha)t})}{(e^{2t} - 1) - (e^{2(1-\alpha)t} - e^{2\alpha t})} \\ &= \frac{2\alpha t \cosh((1-2\alpha)t)}{\sinh t - \sinh((1-2\alpha)t)} = \frac{\alpha t}{\sinh(\alpha t)} \cdot \frac{\cosh((1-2\alpha)t)}{\cosh((1-\alpha)t)}. \end{aligned}$$

Next, when $0 < \alpha < \beta < 1/2$, we have

$$\frac{L_\beta(e^{2t},1)}{L_\alpha(e^{2t},1)} = \frac{1-2\alpha}{1-2\beta} \cdot \frac{e^{2(1-\beta)t} - e^{2\beta t}}{e^{2(1-\alpha)t} - e^{2\alpha t}} = \frac{1-2\alpha}{1-2\beta} \cdot \frac{\sinh((1-2\beta)t)}{\sinh((1-2\alpha)t)}.$$

Since it is positive definite, the first term in (5.3.3) is decreasing in $\alpha \in (0,1/2)$. We compute

$$\frac{\tilde{L}_\beta(e^{2t},1)}{\tilde{L}_\alpha(e^{2t},1)} = \frac{\alpha}{\beta} \cdot \frac{(e^{2t}-1) - (e^{2(1-\beta)t} - e^{2\beta t})}{(e^{2t}-1) - (e^{2(1-\alpha)t} - e^{2\alpha t})} = \frac{\alpha}{\beta} \cdot \frac{\sinh(\beta t)\cosh((1-\beta)t)}{\sinh(\alpha t)\cosh((1-\alpha)t)}$$

and

$$\begin{aligned}
&\frac{\sinh(\beta t)\cosh((1-\beta)t)}{\sinh(\alpha t)\cosh((1-\alpha)t)} - 1 \\
&\quad = \frac{\sinh(\beta t)\cosh(\alpha t + (1-\alpha-\beta)t) - \sinh(\alpha t)\cosh(\beta t + (1-\alpha-\beta)t)}{\sinh(\alpha t)\cosh((1-\alpha)t)} \\
&\quad = \frac{\{\sinh(\beta t)\cosh(\alpha t) - \cosh(\beta t)\sinh(\alpha t)\}\cosh((1-\alpha-\beta)t)}{\sinh(\alpha t)\cosh((1-\alpha)t} \\
&\quad = \frac{\sinh((\beta-\alpha)t)}{\sinh(\alpha t)} \cdot \frac{\cosh((1-\alpha-\beta)t)}{\cosh((1-\alpha)t)}.
\end{aligned}$$

Thus, $\tilde{L}_\beta(e^{2t},1)/\tilde{L}_\alpha(e^{2t},1)$ is positive definite for $0 < \alpha < \beta < 2\alpha$, and the decreasingness of the third term in $\alpha \in (0,1/2)$ is seen as in the proof of Theorem 5.2.1. ☐

We further take the limits of (5.3.3) as $\alpha \searrow 0$ or $\alpha \nearrow 1/2$ to have

$$\begin{aligned}
|||H^{1/2}XK^{1/2}||| &\le \frac{1}{1-2\alpha}\left|\left|\left|\int_\alpha^{1-\alpha} H^tXK^{1-t}\,dt\right|\right|\right| \le \left|\left|\left|\int_0^1 H^tXK^{1-t}\,dt\right|\right|\right| \\
&\le \frac{1}{2\alpha}\left|\left|\left|\int_0^\alpha (H^tXK^{1-t} + H^{1-t}XK^t)\,dt\right|\right|\right| \le \frac{1}{2}\,|||HX + XK|||.
\end{aligned}$$

After the relations $G \preceq L \preceq A$ and $G \preceq A_\alpha \preceq A$ for $0 \le \alpha \le 1$ are known, it is natural to question what is the relation between the logarithmic mean L and the Heinz-type means A_α. This was settled by Drissi [31] as follows.

Theorem 5.3.4. *If $0 \le \alpha \le 1$, then $A_\alpha \preceq L$ holds if and only if $1/4 \le \alpha \le 3/4$.*

Proof. Set $\beta := 2\alpha - 1$ and define

$$f_\beta(t) := \frac{A_\alpha(e^{2t},1)}{L(e^{2t},1)} = \frac{t\cosh(\beta t)}{\sinh t}.$$

Let us determine the range of β for which f_β is positive definite. Since f_β is not bounded if $|\beta| \ge 1$, it suffices to consider the case $|\beta| < 1$. Then by (5.1.6) we have

$$\frac{\sinh(\beta t)}{\sinh t} = \int_{-\infty}^{\infty} e^{its}\,\frac{\sin(\pi\beta)}{2(\cosh(\pi s) + \cos(\pi\beta))}\,ds. \tag{5.3.4}$$

Compute the derivative

$$\frac{d}{d\beta}\,\frac{\sin(\pi\beta)}{\cosh(\pi s) + \cos(\pi\beta)} = \frac{\pi(\cos(\pi\beta)\cosh(\pi s) + 1)}{(\cosh(\pi s) + \cos(\pi\beta))^2}.$$

Hence the Lebesgue convergence theorem can be used to differentiate the right-hand side of (5.3.4) so that

$$\frac{t\cosh(\beta t)}{\sinh t} = \int_{-\infty}^{\infty} e^{its}\,\frac{\pi(\cos(\pi\beta)\cosh(\pi s) + 1)}{2(\cosh(\pi s) + \cos(\pi\beta))^2}\,ds.$$

This implies that f_β is positive definite if and only if $\cos(\pi\beta) \ge 0$ or $|\beta| \le 1/2$, equivalently $1/4 \le \alpha \le 3/4$. ☐

On the other hand, the range of α for which $A_\alpha \le L$ holds was also determined in [31], as stated in the following exercise. Thus we have a one-parameter family of examples showing the difference between $\le$ and $\preceq$.

Exercise 5.3.5. Verify that $A_\alpha \le L$ holds if and only if $\frac{1}{2}(1 - \frac{1}{\sqrt{3}}) \le \alpha \le \frac{1}{2}(1 + \frac{1}{\sqrt{3}})$.

We next consider the means $B_\alpha(x,y) = ((x^\alpha + y^\alpha)/2)^{1/\alpha}$ for $\alpha \in [-\infty, \infty]$. Here, $B_1 = A$ is the arithmetic mean and $B_{-1} = H$ is the harmonic mean, and B_α for the special values $\alpha = 0, \pm\infty$ are understood as

$$B_0(x,y) = G(x,y) = \sqrt{xy} \quad \Bigl(= \lim_{\alpha\to 0} B_\alpha(x,y)\Bigr),$$
$$B_\infty(x,y) = \max\{x,y\} \quad \Bigl(= \lim_{\alpha\to \infty} B_\alpha(x,y)\Bigr),$$
$$B_{-\infty}(x,y) = \min\{x,y\} \quad \Bigl(= \lim_{\alpha\to -\infty} B_\alpha(x,y)\Bigr).$$

It is clear that $B_\alpha^{(-)} = B_{-\alpha}$ for any $\alpha \in [-\infty,\infty]$. From the concavity of t^γ $(t>0)$ for $0<\gamma<1$ one readily checks that $B_\alpha(x,y)$ is monotonically increasing in $\alpha \in [-\infty,\infty]$. When $\alpha = 1/n$ $(n=1,2,\ldots)$, $B_{1/n}$ has the binomial expansion

$$B_{1/n}(H,K)X = \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}H^{k/n}XK^{(n-k)/n}, \tag{5.3.5}$$

so we call B_α's *binomial means*.

Although the family B_α was studied in [39, 40], the basic monotonicity conjecture $B_\alpha \preceq B_\beta$ for $-\infty \le \alpha < \beta \le \infty$ was unsettled there except some special cases. But, in [53] Kosaki finally solved the problem by proving a much stronger result that the ratio

$$\frac{B_\alpha(e^{2t},1)}{B_\beta(e^{2t},1)} = \frac{(\cosh \alpha t)^{1/\alpha}}{(\cosh \beta t)^{1/\beta}}$$

is *infinitely divisible*, i.e., $((\cosh \alpha t)^{1/\alpha}/(\cosh \beta t)^{1/\beta})^r$ is positive definite for each $r>0$ if $\alpha<\beta$. (The infinite divisibility of $M_\alpha(e^{2t},1)/M_\beta(e^{2t},1)$ for $\alpha<\beta$ was also proved in [53].) Thus, we state

Theorem 5.3.6. *If $-\infty \le \alpha < \beta \le \infty$, then $B_\alpha \preceq B_\beta$ and hence*

$$|||B_\alpha(H,K)X||| \le |||B_\beta(H,K)X|||$$

for any unitarily invariant norm $|||\cdot|||$.

Since $B_0 = G$ and $B_1 = A$, Theorem 5.3.6 together with (5.3.5) implies the next corollary except the convergence of $2^{-n}|||\sum_{k=0}^{n}\binom{n}{k}H^{k/n}XK^{(n-k)/n}|||$ to $|||H^{1/2}XK^{1/2}|||$.

Corollary 5.3.7. *For every positive integer n, the inequalities*

$$|||H^{1/2}XK^{1/2}||| \le \frac{1}{2^n}\left|\left|\left|\sum_{k=0}^{n}\binom{n}{k}H^{k/n}XK^{(n-k)/n}\right|\right|\right| \le \frac{1}{2}|||HX+XK|||$$

hold for any unitarily invariant norm. Furthermore,

$$\frac{1}{2^n}\left|\left|\left|\sum_{k=0}^{n}\binom{n}{k}H^{k/n}XK^{(n-k)/n}\right|\right|\right| \quad \text{decreases to} \quad |||H^{1/2}XK^{1/2}||| \quad \text{as } n\to\infty.$$

Exercise 5.3.8. Prove the convergence stated in the last of Corollary 5.3.7.

For $n=1,2,\ldots$, both means $M_{(n+1)/n}$ and $B_{1/n}$ have similar forms as convex combinations of $x^{k/n}y^{(n-k)/n}$, $k=0,1,\ldots,n$. In fact, the next proposition asserts that $B_{1/n} \preceq M_{(n+1)/n}$ holds true.

Proposition 5.3.9. *For every positive integer n and unitarily invariant norm,*

$$\frac{1}{2^n}\left|\left|\left|\sum_{k=0}^{n}\binom{n}{k}H^{k/n}XK^{(n-k)/n}\right|\right|\right| \le \frac{1}{n+1}\left|\left|\left|\sum_{k=0}^{n}H^{k/n}XK^{(n-k)/n}\right|\right|\right|.$$

Proof. We compute

$$\begin{aligned}\frac{B_{1/n}(e^{2t},1)}{M_{(n+1)/n}(e^{2t},1)} - \frac{n+1}{2^n} &= (n+1)\,\frac{\cosh^n(\frac{t}{n})\sinh(\frac{t}{n})}{\sinh(\frac{n+1}{n}t)} - \frac{n+1}{2^n}\\ &= (n+1)\,\frac{(e^{\frac{t}{n}}+e^{-\frac{t}{n}})^n(e^{\frac{t}{n}}-e^{-\frac{t}{n}}) - (e^{\frac{n+1}{n}t}-e^{-\frac{n+1}{n}t})}{2^{n+1}\sinh(\frac{n+1}{n}t)}.\end{aligned} \tag{5.3.6}$$

The numerator in the latter expression is equal to

$$
\begin{aligned}
&\sum_{k=0}^{n}\binom{n}{k}e^{\frac{n-2k}{n}t}(e^{\frac{t}{n}}-e^{-\frac{t}{n}})-(e^{\frac{n+1}{n}t}-e^{-\frac{n+1}{n}t})\\
&=\sum_{k=0}^{n}\binom{n}{k}e^{\frac{n+1-2k}{n}t}-\sum_{k=1}^{n+1}\binom{n}{k-1}e^{\frac{n+1-2k}{n}t}-(e^{\frac{n+1}{n}t}-e^{-\frac{n+1}{n}t})\\
&=\sum_{k=1}^{n}\left\{\binom{n}{k}-\binom{n}{k-1}\right\}e^{\frac{n+1-2k}{n}t}\\
&=\sum_{k=1}^{[\frac{n+1}{2}]}\left\{\binom{n}{k}-\binom{n}{k-1}\right\}(e^{\frac{n+1-2k}{n}t}-e^{-\frac{n+1-2k}{n}t})\\
&=2\sum_{k=1}^{[\frac{n+1}{2}]}\left\{\binom{n}{k}-\binom{n}{k-1}\right\}\sinh\Big(\frac{n+1-2k}{n}t\Big).
\end{aligned}
$$

Now it is clear that (5.3.6) is positive definite and so is $B_{1/n}(e^{2t},1)/M_{(n+1)/n}(e^{2t},1)$. □

Example 5.3.10. Here we note that $L \preceq B_{1/3}$ fails while $L \le B_{1/3}$. Looking at the Taylor expansions

$$
L(e^t,1)=\sum_{n=0}^{\infty}\frac{t^n}{(n+1)!} \quad \text{and} \quad B_\alpha(e^t,1)=1+\frac{t}{2}+\frac{1+\alpha}{8}t^2+\cdots,
$$

we notice that $L \le B_\alpha$ does not hold when $1/6 > (1+\alpha)/8$, i.e., $\alpha < 1/3$. But $L \le B_{1/3}$ is valid, which was shown in [61]. But here is another short proof. Indeed, it suffices to show that $L(e^t,1) \le B_{1/3}(e^t,1)$ for all $t \ge 0$. This can be directly checked because

$$
B_{1/3}(e^t,1)=\frac{1}{8}(e^t+3e^{2t/3}+3e^{t/3}+1)=1+\sum_{n=1}^{\infty}\frac{1}{8n!}\left(1+\frac{2^n}{3^{n-1}}+\frac{1}{3^{n-1}}\right)t^n,
$$

$$
\frac{1}{(n+1)!}\le\frac{1}{8n!}\left(1+\frac{2^n}{3^{n-1}}+\frac{1}{3^{n-1}}\right) \quad \text{for } 1\le n\le 6,
$$

and $1/(n+1)! \le 1/8n!$ is clear for $n \ge 7$. Hence we have $L \le B_\alpha$ when (and only when) $\alpha \ge 1/3$. However, the function

$$
\frac{L(e^{2t},1)}{B_{1/3}(e^{2t},1)}=\frac{\sinh t}{t\cosh^3(\frac{t}{3})}
$$

is not positive definite. This was confirmed by a numerical computation in [39] but a theoretical proof for this non-positive definiteness was also obtained by Kosaki [54, Remark 3].

On the other hand, we observe

$$
\frac{L(e^{2t},1)}{B_{1/2}(e^{2t},1)}=\frac{\sinh t}{t\cosh^2(\frac{t}{2})}=\frac{2\sinh(\frac{t}{2})}{t\cosh(\frac{t}{2})}=2\int_0^{1/2}\frac{\cosh(ut)}{\cosh(\frac{t}{2})}\,du,
$$

and it is positive definite. Hence $L \preceq B_{1/2}$ is valid, which together with Proposition 5.3.9 for $n=2$ implies that

$$
\left|\left|\left|\int_0^1 H^tXK^{1-t}\,dt\right|\right|\right|\le\frac{1}{4}\,|||HX+XK+2H^{\frac12}XK^{\frac12}|||\le\frac{1}{3}\,|||HX+XK+H^{\frac12}XK^{\frac12}|||.
$$

Exercise 5.3.11. We say that $M \in \mathfrak{M}$ is *operator monotone* if $M(x,1)$ is an operator monotone functions on $(0,\infty)$ in the sense of Definition 2.1.3. Show the following facts:

(1) M_α is operator monotone when (and only when) $-1 \le \alpha \le 2$,
(2) A_α is operator monotone for any $0 \le \alpha \le 1$,
(3) B_α is operator monotone when (and only when) $-1 \le \alpha \le 1$.

5.4 Integral expressions for solutions to Lyapunov type matrix equations

Actual computations so far in this chapter have direct relevance to integral expressions for solutions to certain matrix equations (typically the Lyapunov equation). In this section we collect some integral formulas in this connection.

First, for given matrices H, K, X with $H, K > 0$, we consider the following algebraic equations in Y:

$$\frac{1}{n}\sum_{k=0}^{n-1} H^{k/(n-1)} Y K^{(n-1-k)/(n-1)} = X, \tag{5.4.1}$$

$$\frac{1}{m}\sum_{k=1}^{m} H^{k/(m+1)} Y K^{(m+1-k)/(m+1)} = X. \tag{5.4.2}$$

The limit case of these equations is

$$\int_0^1 H^t Y K^{1-t}\,dt = X. \tag{5.4.3}$$

Proposition 5.4.1. *Assume that $H, K > 0$. Then the above equations (5.4.1) for $n \ge 2$, (5.4.2) for $m \ge 2$, and (5.4.3) have unique solutions $Y = Y_n$, $Y = \tilde{Y}_m$, and $Y = Y_\infty$ respectively, and furthermore they are expressed as*

$$Y_n = \int_{-\infty}^{\infty} H^{is} H^{-1/2} X K^{-1/2} K^{-is} \frac{(n-1)\sin(\frac{\pi}{n})}{\cosh(\pi\frac{2(n-1)}{n}s) + \cos(\frac{\pi}{n})}\,ds, \tag{5.4.4}$$

$$\tilde{Y}_m = \int_{-\infty}^{\infty} H^{is} H^{-1/2} X K^{-1/2} K^{-is} \frac{(m+1)\sin(\frac{\pi}{m})}{\cosh(\pi\frac{2(m+1)}{m}s) + \cos(\frac{\pi}{m})}\,ds,$$

$$Y_\infty = \int_{-\infty}^{\infty} H^{is} H^{-1/2} X K^{-1/2} K^{-is} \frac{\pi}{2\cosh^2(\pi s)}\,ds. \tag{5.4.5}$$

Here, the inequalities $|||Y_n||| \le |||Y_\infty||| \le |||\tilde{Y}_m|||$ hold for any unitarily invariant norm, and $|||Y_n|||$ increases to $|||Y_\infty|||$ as $n \to \infty$ while $|||\tilde{Y}_m|||$ decreases to $|||Y_\infty|||$ as $m \to \infty$.

Proof. Since the equations (5.4.1)–(5.4.3) are

$$M_{n/(n-1)}(H,K)Y = X, \quad M_{m/(m+1)}(H,K)Y = X, \quad L(H,K)Y = X,$$

they have unique solutions respectively given as

$$\begin{aligned} Y_n &= M^{(-)}_{n/(n-1)}(H^{-1},K^{-1})X = M_{-1/(n-1)}(H^{-1},K^{-1})X,\\ \tilde{Y}_m &= M^{(-)}_{m/(m+1)}(H^{-1},K^{-1})X = M_{1/(m+1)}(H^{-1},K^{-1})X,\\ Y_\infty &= L^{(-)}(H^{-1},K^{-1})X = M_0(H^{-1},K^{-1})X. \end{aligned}$$

Noting that

$$\frac{M_\alpha(e^t,1)}{G(e^t,1)} = \frac{G(e^t,1)}{M_{1-\alpha}(e^t,1)} \quad \text{for } \alpha < 1/2,$$

we obtain, thanks to (5.2.7), the integral expressions of Y_n and $\tilde{Y}_m$ by Theorem 5.1.3. That of Y_∞ is obtained by (5.1.8) since

$$\frac{M_0(e^t,1)}{G(e^t,1)} = \frac{t}{2\sinh(\frac{t}{2})}.$$

The remaining assertions on norm inequalities are immediate from Theorem 5.2.1. □

Secondly, for $\alpha \in \mathbb{R}$ with $\alpha \ne 1$ we set $f(x) := x^{(\alpha-1)/\alpha}$, $x > 0$. Notice that

$$M_\alpha(x,y)^{-1} = \frac{\alpha}{\alpha-1}\cdot\frac{x^{\alpha-1}-y^{\alpha-1}}{x^\alpha - y^\alpha} = \frac{\alpha}{\alpha-1} f^{[1]}(x^\alpha, y^\alpha) \quad x,y > 0.$$

Hence, for $H > 0$ it follows from (5.1.3) and (2.3.9) that

$$M_\alpha(H,H)^{-1}X = \frac{\alpha}{\alpha-1} D(f(H^\alpha))(X) \tag{5.4.6}$$

for all X. For $1 < \alpha < \infty$, since $0 < (\alpha-1)/\alpha < 1$, $f(x) = x^{(\alpha-1)/\alpha}$ has the integral expression

$$f(x) = \frac{\sin(\pi\frac{\alpha-1}{\alpha})}{\pi}\int_0^\infty \frac{xt^{-1/\alpha}}{z+t}\,dt = a + \frac{\sin(\pi\frac{\alpha-1}{\alpha})}{\pi}\int_0^\infty \left(\frac{t}{t^2+1} - \frac{1}{x+t}\right) t^{(\alpha-1)/\alpha}\,dt$$

with

$$a = \frac{\sin(\pi\frac{\alpha-1}{\alpha})}{\pi}\int_0^\infty \left(\frac{1}{t} - \frac{t}{t^2+1}\right) t^{(\alpha-1)/\alpha}\,dt.$$

Therefore, we obtain

$$f(H^\alpha + uX) = aI + \frac{\sin(\pi\frac{\alpha-1}{\alpha})}{\pi}\int_0^\infty \left(\frac{t}{t^2+1}I - (H^\alpha + tI + uX)^{-1}\right)t^{(\alpha-1)/\alpha}\,dt$$

for $|u|$ small enough. Since

$$\frac{d}{du}(H^\alpha + tI + uX)^{-1}\bigg|_{u=0} = -(H^\alpha + tI)^{-1}X(H^\alpha + tI)^{-1},$$

from the above integral formula we have

$$D(f(H^\alpha))(X) = \frac{\sin(\pi\frac{\alpha-1}{\alpha})}{\pi}\int_0^\infty (H^\alpha + tI)^{-1}X(H^\alpha + tI)^{-1}t^{(\alpha-1)/\alpha}\,dt.$$

Thanks to (5.4.6) this computation with the multiple constant $\alpha/(\alpha-1)$ gives the unique solution to the equation $M_\alpha(H,H)Y = X$. Thus, the usual 2×2 matrix trick shows

Proposition 5.4.2. *If $1 < \alpha < \infty$ and $H,K > 0$, then the unique solution to the equation $M_\alpha(H,K)Y = X$ is expressed as*

$$Y = \frac{\alpha\sin(\pi\frac{\alpha-1}{\alpha})}{\pi(\alpha-1)}\int_0^\infty (H^\alpha + tI)^{-1}X(K^\alpha + tI)^{-1}t^{(\alpha-1)/\alpha}\,dt.$$

In particular, as an alternative form of the solution (5.4.4) to the equation (5.4.1) given in Proposition 5.4.1 we have

$$Y_n = \frac{n\sin(\pi/n)}{\pi}\int_0^\infty (H^{n/(n-1)} + tI)^{-1}X(K^{n/(n-1)} + tI)^{-1}t^{1/n}\,dt. \tag{5.4.7}$$

Also, set $g(x) := \log x$, $x > 0$, so that

$$L(x,y)^{-1} = g^{[1]}(x,y), \qquad x,y > 0.$$

Since

$$L(H,H)^{-1}X = D(g(H))(X)$$

and

$$g(x) = \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t}\right)dt,$$

it follows as Proposition 5.4.2 that the solution (5.4.5) to the equation (5.4.3) also admits another integral expression

$$Y_\infty = \int_0^\infty (H + tI)^{-1}X(K + tI)^{-1}\,dt, \tag{5.4.8}$$

and it is the limit of (5.4.7) as $n \to \infty$.

Finally, let $M \in \mathfrak{M}$ be an operator monotone mean (as defined in Exercise 5.3.11), so by Theorem 2.7.11, $M(x,1)$ admits the representation

$$M(x,1) = a + bx + \int_0^\infty \frac{x}{x+t}\,dm(t), \qquad x \ge 0,$$

where $a,b \ge 0$ and m is a positive measure on $(0,\infty)$ such that $\int_0^\infty (1+t)^{-1}\,dm(t) < +\infty$. But note that the symmetry $M(x,1) = xM(x^{-1},1)$ forces $a = b$ and $dm(t) = t\,dm(t^{-1})$. Hence we have the integral expression

$$M(x,y) = a(x+y) + \int_0^\infty \frac{xy}{x+ty}\,dm(t), \qquad x,y \ge 0,$$

where a and m satisfy $2a + \int_0^\infty (1+t)^{-1}\,dm(t) = 1$. By noticing that

$$\frac{xy}{x+ty} = \int_0^\infty e^{-sx}xye^{-sty}\,ds = \int_0^\infty e^{-stx^{-1}}e^{-sy^{-1}}\,ds, \qquad x,y,\lambda > 0,$$

we observe that

$$\begin{aligned} M(H,K)X &= a(HX + XK) + \int_0^\infty\int_0^\infty e^{-sH}HXKe^{-stK}\,ds\,dm(t) \\ &= a(HX + XK) + \int_0^\infty\int_0^\infty e^{-stH^{-1}}Xe^{-sK^{-1}}\,ds\,dm(t) \end{aligned}$$

for all matrices H,K,X with $H,K > 0$. Assume that the measure m has the density $\varphi(t)$. We set $\psi(s,t) = s^{-1}\varphi(t/s)$, $s,t > 0$, so that $\psi(s,t) = \psi(t,s)$. In this case we have

$$
\begin{aligned}
M(H,K)X &= a(HX+XK) + \int_0^\infty \int_0^\infty e^{-sH} HXK e^{-tK} \psi(s,t)\, ds\, dt \\
&= a(HX+XK) + \int_0^\infty \int_0^\infty e^{-sH^{-1}} X e^{-tK^{-1}} \psi(s,t)\, ds\, dt.
\end{aligned} \tag{5.4.9}
$$

Since the harmonic mean $M_{-1} = H$ is the case where $a = 0$ and $m = 2\delta_1$, we have

$$
H(H,K)X = 2\int_0^\infty e^{-sH} HXK e^{-sK}\, ds = 2\int_0^\infty e^{-sH^{-1}} X e^{-sK^{-1}}\, ds.
$$

By this together with (5.4.4) and (5.4.7) for $n = 2$, the unique solution to the famous *Lyapunov equation* $HY + YK = X$ with $H, K > 0$ has the three different integral expressions as follows:

$$
\begin{aligned}
Y &= \int_0^\infty e^{-tH} X e^{-tK}\, dt \\
&= \int_{-\infty}^\infty H^{it} H^{-1/2} X K^{-1/2} K^{-it} \, \frac{1}{2\cosh(\pi t)}\, dt \\
&= \frac{1}{\pi}\int_0^\infty (H^2 + tI)^{-1} X (K^2 + tI)^{-1} t^{1/2}\, dt \\
\biggl(&= \frac{2}{\pi}\int_0^\infty (H^2 + t^2 I)^{-1} X (K^2 + t^2 I)^{-1} t^2\, dt \biggr).
\end{aligned}
$$

As further examples of the integral expression of the form (5.4.9) we have

Proposition 5.4.3. *For every $H, K > 0$ and X,*

$$
\begin{aligned}
L(H,K)X &= \int_0^\infty \int_0^\infty e^{-sH} HXK e^{-tK} \frac{s+t}{st((\log \frac{t}{s})^2 + \pi^2)}\, ds\, dt \\
&= \int_0^\infty \int_0^\infty e^{-sH^{-1}} X e^{-tK^{-1}} \frac{s+t}{st((\log \frac{t}{s})^2 + \pi^2)}\, ds\, dt,
\end{aligned}
$$

$$
M_0(H,K)X = \int_0^\infty \int_0^\infty e^{-sH} HXK e^{-tK} \frac{ds\, dt}{s+t} = \int_0^\infty \int_0^\infty e^{-sH^{-1}} X e^{-tK^{-1}} \frac{ds\, dt}{s+t}.
$$

Proof. Since

$$
\begin{aligned}
L(x,1) &= \int_0^1 x^\alpha\, d\alpha = \int_0^1 \left(\frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \frac{x t^{\alpha-1}}{x+t}\, dt \right) d\alpha \\
&= \int_0^\infty \frac{x}{x+t} \left(\int_0^1 \frac{\sin(\pi\alpha)}{\pi t^{1-\alpha}}\, d\alpha \right) dt = \int_0^\infty \frac{x}{x+t} \left(\frac{t+1}{t((\log t)^2 + \pi^2)} \right) dt,
\end{aligned}
$$

the integral expressions in (5.4.9) for L are given with $a = 0$ and

$$
\psi(s,t) = \frac{s+t}{st((\log \frac{t}{s})^2 + \pi^2)}.
$$

On the other hand, since

$$
M_0(x,1) = \frac{\log x}{1 - x^{-1}} = \frac{x}{x-1} \int_0^\infty \left(\frac{1}{1+t} - \frac{1}{x+t} \right) dt = \int_0^\infty \frac{x}{(x+t)(1+t)}\, dt,
$$

the expressions in (5.4.9) for M_0 are given with $a = 0$ and $\psi(s,t) = (s+t)^{-1}$. □

The second expression of the above proposition implies that the solution Y_∞ to the equation (5.4.3) admits, besides (5.4.5) and (5.4.8), one more integral expression

$$
Y_\infty = \int_0^\infty \int_0^\infty e^{-sH} X e^{-tK} \frac{ds\, dt}{s+t}.
$$

Appendix

A.1 Converse to Taylor's theorem

In this section let $\mathcal{X}$ and $\mathcal{Y}$ be general Banach spaces and f be a map from an open subset $\mathcal{U}$ of $\mathcal{X}$ into $\mathcal{Y}$. Then f is said to be *Fréchet differentiable* at a point $a \in \mathcal{U}$ if there exists a $Df(a) \in B(\mathcal{X}, \mathcal{Y})$ such that

$$\frac{\|f(a+x)-f(a)-Df(a)x\|}{\|x\|}\longrightarrow 0 \quad \text{as } x\in\mathcal{X},\ \|x\|\to 0.$$

The higher degree Fréchet differentiability of f is also inductively defined as stated in Section 2.3. For $m\in\mathbb{N}$, f is said to be C^m on $\mathcal{U}$ if f is m times Fréchet differentiable and $D^m f:\mathcal{U}\to B(\mathcal{X}^m,\mathcal{Y})$ is norm-continuous. The next lemma, called converse to Taylor's theorem, provides a useful technique to prove the C^m of f. For the proof see [1, pp. 6–9].

Lemma A.1.1. *Let $\mathcal{X},\mathcal{Y}$ be Banach spaces, and let an open convex set $\mathcal{U}\subset\mathcal{X}$ and a map $f:\mathcal{U}\to\mathcal{Y}$ be given. Moreover, let $m\ge 0$ and $\Phi_k:\mathcal{U}\to B_s(\mathcal{X}^k,\mathcal{Y})$, $k=0,1,\dots,m$, be given, where $\Phi_0:\mathcal{U}\to\mathcal{Y}$ and for $k\ge 1$, $B_s(\mathcal{X}^k,\mathcal{Y})$ denotes the set of symmetric (under permutation of arguments) bounded k-multilinear maps from $\mathcal{X}^k$ to $\mathcal{Y}$. For every $a\in\mathcal{U}$ and $x\in\mathcal{X}$ such that $a+x\in\mathcal{U}$, define $R(a,x)\in\mathcal{Y}$ by*

$$f(a+x)=\sum_{k=0}^{m}\frac{1}{k!}\,\Phi_k(a)(x^{(k)})+R(a,x),$$

where $x^{(k)}$ denotes the k times $x,\dots,x$. Assume:
(a) *for each $k=0,1,\dots,m$, Φ_k is norm-continuous,*
(b) *$\|R(a,x)\|/\|x\|^m\to 0$ as $(a,x)\in\mathcal{U}\times\mathcal{X}$, $(a,x)\to(b,0)$ for each $b\in\mathcal{U}$.*
Then f is C^m on $\mathcal{U}$ and $D^k f=\Phi_k$ for all $k=0,1,\dots,m$.

A.2 Regularization of functions

Choose and fix a smooth (i.e., C^∞) function φ on $\mathbb{R}$ such that φ is supported on $[-1,1]$, i.e., $\varphi(x)=0$ outside $[-1,1]$, $\varphi(x)\ge 0$ and $\varphi(x)=\varphi(-x)$ for all $x\in\mathbb{R}$, and $\int_{-1}^{1}\varphi(x)\,dx=1$. Let f be a real measurable function on an open interval (a,b) assumed to be locally integrable in the sense that $\int_c^d|f(x)|\,dx<+\infty$ for any closed interval $[c,d]$ inside (a,b). For each small $\varepsilon>0$ we define a function f_ε on $(a+\varepsilon,b-\varepsilon)$ by

$$f_\varepsilon(x):=\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}\varphi\Big(\frac{x-t}{\varepsilon}\Big)f(t)\,dt=\int_{-1}^{1}\varphi(t)f(x-\varepsilon t)\,dt,\qquad x\in(a+\varepsilon,b-\varepsilon),\tag{A.2.1}$$

which we call the *regularization* of f of order ε. In fact, we have:

Lemma A.2.1.
(1) *f_ε is C^∞ on $(a+\varepsilon,b-\varepsilon)$ for every $\varepsilon>0$.*
(2) *If f is continuous at $x_0\in(a,b)$, then $f_\varepsilon(x_0)$ converges as $\varepsilon\searrow 0$ to $f(x_0)$.*
(3) *If f is continuous on (a,b), then f_ε converges as $\varepsilon\searrow 0$ to f uniformly on any closed interval inside (a,b).*
(4) *If f is absolutely continuous on any closed interval inside (a,b), then $(f')_\varepsilon=f'_\varepsilon$, i.e., the regularization $(f')_\varepsilon$ of f' is the derivative of f_ε, and moreover $f'_\varepsilon(x)$ converges as $\varepsilon\searrow 0$ to $f'(x)$ almost everywhere on (a,b).*

Proof. The proofs of (1)–(3) are easy and may be left to exercises.

(4) Recall that f is absolutely continuous on $[c,d]$ if and only if f is differentiable almost everywhere on $[c,d]$ with integrable derivative f'. In this case, $f(x)=f(c)+\int_c^x f'(t)\,dt$ for all $x\in[c,d]$. Since

$$\frac{d}{dt}\Big\{\varphi\Big(\frac{x-t}{\varepsilon}\Big)f(t)\Big\}=-\frac{1}{\varepsilon}\varphi'\Big(\frac{x-t}{\varepsilon}\Big)f(t)+\varphi\Big(\frac{x-t}{\varepsilon}\Big)f'(t)$$

for almost everywhere $t\in[c,d]$, we have

$$\begin{aligned}(f')_\varepsilon(x)&=\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}\varphi\Big(\frac{x-t}{\varepsilon}\Big)f'(t)\,dt=\frac{1}{\varepsilon^2}\int_{x-\varepsilon}^{x+\varepsilon}\varphi'\Big(\frac{x-t}{\varepsilon}\Big)f(t)\,dt\\&=\frac{1}{\varepsilon}\frac{d}{dx}\int_{x-\varepsilon}^{x+\varepsilon}\varphi\Big(\frac{x-t}{\varepsilon}\Big)f(t)\,dt=f'_\varepsilon(x)\end{aligned}\tag{A.2.2}$$

for every $x\in[c+\varepsilon,d-\varepsilon]$. Since $[c,d]$ is an arbitrary closed interval inside (a,b), $(f')_\varepsilon=f'_\varepsilon$ on $(a+\varepsilon,b-\varepsilon)$.

For the latter assertion, it suffices to prove that $f'_\varepsilon(x)$ converges as $\varepsilon\searrow 0$ to $f'(x)$ at any differentiable point $x\in(a,b)$ for f. If x is such a point, then

$$f(t)=f(x)+f'(x)(t-x)+\theta(t),$$

where $\theta(t)=o(|x-t|)$. Then, by (A.2.2),

$$\begin{aligned}f'_\varepsilon(x)&=\frac{1}{\varepsilon^2}\int_{x-\varepsilon}^{x+\varepsilon}\varphi'\Big(\frac{x-t}{\varepsilon}\Big)f(t)\,dt\\&=\frac{f'(x)}{\varepsilon^2}\int_{x-\varepsilon}^{x+\varepsilon}\varphi'\Big(\frac{x-t}{\varepsilon}\Big)(t-x)\,dt+\frac{1}{\varepsilon^2}\int_{x-\varepsilon}^{x+\varepsilon}\varphi'\Big(\frac{x-t}{\varepsilon}\Big)\theta(t)\,dt.\end{aligned}$$

The first term of the last expression is

$$-f'(x)\int_{-1}^{1}\varphi'(t)t\,dt = f'(x).$$

The second term is dominated by

$$\frac{1}{\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}\left|\varphi'\left(\frac{x-t}{\varepsilon}\right)\right|dt\cdot\sup_{|t-x|<\varepsilon}\frac{|\theta(t)|}{\varepsilon} = \int_{-1}^{1}|\varphi'(t)|\,dt\cdot\sup_{|t-x|<\varepsilon}\frac{|\theta(t)|}{\varepsilon}\longrightarrow 0 \quad \text{as } \varepsilon\searrow 0.$$

Hence $f_\varepsilon'(x)\to f'(x)$ as $\varepsilon\searrow 0$. □

Exercise A.2.2. Show (1)–(3) of the above lemma.

A.3 C^2 of 2-convex functions: Proof of Theorem 2.4.2

This section is devoted to the proof of Theorem 2.4.2 based on the original paper of Kraus [57] and also Ando's English translation. We begin with the following general criteria for C^1 and C^2 functions in terms of their second and third divided differences, which will be useful in the proof below.

Lemma A.3.1. *Let f be a real-valued function on (a,b), and let $a<c<d<b$.*
(i) *If $f^{[2]}(x_1,x_2,x_3)$ is uniformly bounded when $x_1<x_2<x_3$ run over $[c,d]$, then f is C^1 on (c,d).*
(ii) *If $f^{[3]}(x_1,x_2,x_3,x_4)$ is uniformly bounded when $x_1<x_2<x_3<x_4$ run over $[c,d]$, then f is C^2 on (c,d).*

Proof. (i) Assume that

$$K := \sup\{|f^{[2]}(x_1,x_2,x_3)| : c\le x_1<x_2<x_3\le d\} < +\infty.$$

Let $\xi\in(c,d)$ and $\delta,\delta'\in(c-\xi,d-\xi)$ with $\delta,\delta'\ne 0$ and $\delta\ne\delta'$. Then

$$\left|\frac{f(\xi+\delta)-f(\xi)}{\delta}-\frac{f(\xi+\delta')-f(\xi)}{\delta'}\right| = |f^{[2]}(\xi+\delta,\xi,\xi+\delta')(\delta-\delta')|\le K|\delta-\delta'|.$$

This implies that

$$f'(\xi)=\lim_{\delta\to 0}\frac{f(\xi+\delta)-f(\xi)}{\delta}$$

exists. Moreover, for every $\xi<\eta$ in (c,d), since

$$f^{[2]}(\xi,x,\eta)=\frac{\frac{f(x)-f(\xi)}{x-\xi}-\frac{f(x)-f(\eta)}{x-\eta}}{\xi-\eta}$$

for every $x\in(a,b)$ with ξ,x,η distinct, we have

$$\lim_{x\to\xi}f^{[2]}(\xi,x,\eta)=\frac{f'(\xi)-\frac{f(\xi)-f(\eta)}{\xi-\eta}}{\xi-\eta},\qquad \lim_{y\to\eta}f^{[2]}(\xi,y,\eta)=\frac{\frac{f(\xi)-f(\eta)}{\xi-\eta}-f'(\eta)}{\xi-\eta}.$$

Therefore,

$$\left|\frac{f'(\xi)-f'(\eta)}{\xi-\eta}\right| = \lim_{x\to\xi,\,y\to\eta}|f^{[2]}(\xi,x,\eta)+f^{[2]}(\xi,y,\eta)|\le 2K$$

so that $|f'(\xi)-f'(\eta)|\le 2K|\xi-\eta|$ for all $\xi,\eta\in(c,d)$, implying the C^1 of f on (c,d).

(ii) Assume that

$$K := \sup\{|f^{[3]}(x_1,x_2,x_3,x_4)| : c\le x_1<x_2<x_3<x_4\le d\} < +\infty.$$

For any choices of $d'\in(c,d)$ and $\lambda\in(d',d)$, apply (i) to the function $f^{[1]}(x,\lambda)=(f(x)-f(\lambda))/(x-\lambda)$ to see that f is C^1 on (c,d'). Since $d'\in(c,d)$ is arbitrary, f is C^1 on (c,d). Hence for each $\xi\in(c,d)$, one can define the function g_ξ on (a,b) by

$$g_\xi(x) := \begin{cases} f^{[1]}(x,\xi) & \text{if } x\in(a,b),\ x\ne\xi,\\ f'(\xi) & \text{if } x=\xi.\end{cases}$$

For every $x_1,x_2,x_3\in(a,b)$ with x_1,x_2,x_3,ξ distinct, notice

$$g_\xi^{[2]}(x_1,x_2,x_3)=f^{[3]}(x_1,x_2,x_3,\xi),$$
$$g_\xi^{[2]}(x_1,x_2,\xi)=\lim_{x\to\xi}f^{[3]}(x_1,x_2,x,\xi).$$

Therefore,

$$\sup\{|g_\xi^{[2]}(x_1,x_2,x_3)| : c\le x_1<x_2<x_3\le d\}\le K$$

so that g_ξ is C^1 on (c,d) by (i). We now define

$$h(\xi) := g_\xi'(\xi) = \lim_{x\to\xi} \frac{f^{[1]}(x,\xi) - f'(\xi)}{x-\xi}, \qquad \xi \in (c,d).$$

Note that

$$h(\xi) = \lim_{x\to\xi}\lim_{y\to\xi} \frac{f^{[1]}(x,\xi) - f^{[1]}(y,\xi)}{x-y} = \lim_{x\to\xi}\lim_{y\to\xi} f^{[2]}(x,y,\xi). \tag{A.3.1}$$

Moreover, for every $\xi < \eta$ in (c,d), we notice

$$\begin{aligned}
&|f^{[2]}(x,y,\xi) - f^{[2]}(x',y',\eta)| \\
&\quad \le |f^{[2]}(x,y,\xi) - f^{[2]}(x',y,\xi)| + |f^{[2]}(x',y,\xi) - f^{[2]}(x',y',\xi)| + |f^{[2]}(x',y',\xi) - f^{[2]}(x',y',\eta)| \\
&\quad \le K|x-x'| + K|y-y'| + K|\xi-\eta|.
\end{aligned} \tag{A.3.2}$$

Letting $y \to \xi$, $y' \to \eta$ and then $x \to \xi$, $x' \to \eta$, we have

$$|h(\xi) - h(\eta)| \le 3K|\xi - \eta|, \qquad \xi,\eta \in (c,d)$$

so that h is continuous on (c,d). Letting

$$r(\xi,\delta) := \frac{f^{[1]}(\xi+\delta,\xi) - f'(\xi)}{\delta} - h(\xi)$$

for $\xi \in (c,d)$ and $\delta \in (c-\xi, d-\xi)$, $\delta \ne 0$, one can write

$$f(\xi+\delta) = f(\xi) + f'(\xi)\delta + h(\xi)\delta^2 + r(\xi,\delta)\delta^2.$$

Notice that

$$\begin{aligned}
r(\xi,\delta) &= \lim_{x'\to\xi}\left\{\frac{f^{[1]}(\xi+\delta,\xi) - f^{[1]}(x',\xi)}{\xi+\delta-x'} - h(\xi)\right\} \\
&= \lim_{x'\to\xi}\{f^{[2]}(x',\xi+\delta,\xi) - h(\xi)\} \\
&= \lim_{x\to\xi}\lim_{y\to\xi}\lim_{x'\to\xi}\{f^{[2]}(x',\xi+\delta,\xi) - f^{[2]}(x,y,\xi)\}
\end{aligned}$$

thanks to (A.3.1) and that

$$|f^{[2]}(x',\xi+\delta,\xi) - f^{[2]}(x,y,\xi)| \le K|x'-x| + K|\xi+\delta-y|$$

as in (A.3.2). Hence $|r(\xi,\delta)| \le K\delta$ for all $\xi \in (c,d)$ and all $\delta \in (c-\xi, d-\xi)$, $\delta \ne 0$. This implies (see Lemma A.1.1) that f is C^2 on (c,d) with $f''(\xi) = 2h(\xi)$. □

Throughout the rest of this section, assume that f is a conditionally 2-convex function on (a,b) as stated in Theorem 2.4.2.

Lemma A.3.2. *Let $\xi_1 < \eta_1 < \xi < \eta_2 < \xi_2$ be given in (a,b) with constraint*

$$(\xi_1 - \xi)(\eta_2 - \xi) + (\xi_2 - \xi)(\eta_1 - \xi) - (\xi_1 - \xi)(\xi_2 - \xi) \ge 0. \tag{A.3.3}$$

Then

$$\begin{aligned}
&\det\begin{bmatrix} f^{[2]}(\xi_1,\xi,\eta_1) & f^{[2]}(\eta_1,\xi,\eta_2) \\ f^{[2]}(\xi_1,\xi,\xi_2) & f^{[2]}(\xi_2,\xi,\eta_2) \end{bmatrix}(\xi_1-\eta_1)(\xi_2-\eta_2) \\
&\qquad + f^{[2]}(\xi_1,\xi,\xi_2) f^{[2]}(\eta_1,\xi,\eta_2)(\eta_1-\xi)(\eta_2-\xi) \le 0,
\end{aligned} \tag{A.3.4}$$

or equivalently,

$$\begin{aligned}
&f^{[2]}(\xi_1,\xi,\eta_1) f^{[2]}(\xi_2,\xi,\eta_2)(\xi_1-\eta_1)(\xi_2-\eta_2) \\
&\quad + f^{[2]}(\xi_1,\xi,\xi_2) f^{[2]}(\eta_1,\xi,\eta_2)\{(\xi_1-\xi)(\eta_2-\xi) + (\xi_2-\xi)(\eta_1-\xi) - (\xi_1-\xi)(\xi_2-\xi)\} \le 0.
\end{aligned} \tag{A.3.5}$$

Proof. First it is immediate to verify the equivalence between (A.3.4) and (A.3.5).

Step 1. We notice that it is enough to prove the lemma in the situation where $\xi = 0 \in (a,b)$ and $f(\xi_1) = f(\xi_2) = 0$. In fact, define the function $g(x) := f(x+\xi) + \alpha x + \beta$ on $(a-\xi, b-\xi)$, where $\alpha, \beta \in \mathbb{R}$ can be determined so that we have $g(\xi_1 - \xi) = g(\xi_2 - \xi) = 0$. Since g is conditionally 2-convex on $(a-\xi, b-\xi)$ and $f^{[2]}(x,\xi,y) = g^{[2]}(x-\xi, 0, y-\xi)$ for all $x, y \in (a,b)$ with x, ξ, y distinct, one can reduce the proof of the lemma to the stated situation. Thus, in the rest of the proof, we assume that

$$\xi_1 < \eta_1 < 0 < \eta_2 < \xi_2 \quad \text{in } (a,b), \tag{A.3.6}$$

$$f(\xi_1) = f(\xi_2) = 0 \tag{A.3.7}$$

with constraint

$$\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2 \geq 0. \tag{A.3.8}$$

Step 2. We show that there exist $a_1, a_2, b_1, b_2 \in \mathbb{R}$ and $\lambda \in (0,1)$ such that

$$a_1^2 + a_2^2 = b_1^2 + b_2^2 = 1, \tag{A.3.9}$$

$$\begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} = \lambda\xi_1 \begin{bmatrix} a_1^2 & a_1a_2 \\ a_1a_2 & a_2^2 \end{bmatrix} + (1-\lambda)\xi_2 \begin{bmatrix} b_1^2 & b_1b_2 \\ b_1b_2 & b_2^2 \end{bmatrix}. \tag{A.3.10}$$

In fact, (A.3.10) means that

$$\lambda\xi_1 a_1^2 + (1-\lambda)\xi_2 b_1^2 = \eta_1, \tag{A.3.11}$$

$$\lambda\xi_1 a_2^2 + (1-\lambda)\xi_2 b_2^2 = \eta_2, \tag{A.3.12}$$

$$\lambda\xi_1 a_1a_2 + (1-\lambda)\xi_2 b_1b_2 = 0. \tag{A.3.13}$$

It follows from (A.3.9), (A.3.11) and (A.3.12) that $\lambda\xi_1 + (1-\lambda)\xi_2 = \eta_1 + \eta_2$ and hence

$$\lambda = \frac{\xi_2 - \eta_1 - \eta_2}{\xi_2 - \xi_1}, \tag{A.3.14}$$

which is in $(0,1)$ thanks to (A.3.6). From (A.3.9) (A.3.13) and (A.3.11) we have

$$\lambda^2\xi_1^2 a_1^2(1-a_1^2) = (1-\lambda)^2\xi_2^2 b_1^2(1-b_1^2)$$

and

$$\eta_1^2 - 2\lambda\xi_1\eta_1 a_1^2 + \lambda^2\xi_1^2 a_1^4 = (1-\lambda)^2\xi_2^2 b_1^4$$

so that

$$\lambda\xi_1(2\eta_1 - \lambda\xi_1)a_1^2 + (1-\lambda)^2\xi_2^2 b_1^2 = \eta_1^2.$$

Solving this and (A.3.11) as a pair of linear equations for a_1^2 and a_2^2, and applying (A.3.14), we have

$$a_1^2 = \frac{\eta_1(\xi_1\eta_1 + \xi_2\eta_2 - \xi_1\xi_2)}{\xi_1(\eta_2 - \eta_1)(\xi_2 - \eta_1 - \eta_2)}, \tag{A.3.15}$$

$$b_1^2 = \frac{-\eta_1(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2)}{\xi_2(\eta_2 - \eta_1)(\eta_1 + \eta_2 - \xi_1)}. \tag{A.3.16}$$

Hence, by (A.3.9),

$$a_2^2 = \frac{\eta_2(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2)}{-\xi_1(\eta_2 - \eta_1)(\xi_2 - \eta_1 - \eta_2)}, \tag{A.3.17}$$

$$b_2^2 = \frac{\eta_2(\xi_1\eta_1 + \xi_2\eta_2 - \xi_1\xi_2)}{\xi_2(\eta_2 - \eta_1)(\eta_1 + \eta_2 - \xi_1)}. \tag{A.3.18}$$

By assumptions (A.3.6) and (A.3.8), the right-hand sides of (A.3.15)–(A.3.18) are all nonnegative. So one can fix a_1, a_2, b_1, b_2 satisfying (A.3.15)–(A.3.18) with positive sign, for which (A.3.11)–(A.3.13) are really satisfied so that (A.3.9) and (A.3.10) hold.

Step 3. Put

$$A := \xi_1 \begin{bmatrix} a_1^2 & a_1a_2 \\ a_1a_2 & a_2^2 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix} \begin{bmatrix} \xi_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix},$$

$$B := \xi_2 \begin{bmatrix} b_1^2 & b_1b_2 \\ b_1b_2 & b_2^2 \end{bmatrix} = \begin{bmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{bmatrix} \begin{bmatrix} \xi_2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{bmatrix},$$

where $\begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix}$ and $\begin{bmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{bmatrix}$ are unitaries. Then $A \leq 0 \leq B$ and (A.3.10) means that

$$\begin{bmatrix} \eta_1 & 0 \\ 0 & \eta_2 \end{bmatrix} = \lambda A + (1-\lambda)B.$$

Hence the conditional 2-convexity of f implies, thanks to (A.3.7), that

$$\begin{aligned}
\begin{bmatrix} f(\eta_1) & 0 \\ 0 & f(\eta_2) \end{bmatrix} &\le \lambda f(A) + (1-\lambda) f(B) \\
&= \lambda \begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & f(0) \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_2 & -a_1 \end{bmatrix} + (1-\lambda) \begin{bmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & f(0) \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_2 & -b_1 \end{bmatrix} \\
&= f(0) \begin{bmatrix} \lambda a_2^2 + (1-\lambda) b_2^2 & -\{\lambda a_1 a_2 + (1-\lambda) b_1 b_2\} \\ -\{\lambda a_1 a_2 + (1-\lambda) b_1 b_2\} & \lambda a_1^2 + (1-\lambda) b_2^2 \end{bmatrix}.
\end{aligned}$$

By taking the determinant of the difference of the right-hand and the left-hand sides, we have $\beta_{11}\beta_{22} - \beta_{12}^2 \ge 0$, where

$$\begin{aligned}
\beta_{11} &:= f(0)\{\lambda a_2^2 + (1-\lambda) b_2^2\} - f(\eta_1), \\
\beta_{22} &:= f(0)\{\lambda a_1^2 + (1-\lambda) b_1^2\} - f(\eta_2), \\
\beta_{12} &:= f(0)\{\lambda a_1 a_2 + (1-\lambda) b_1 b_2\}.
\end{aligned}$$

Direct computations using (A.3.9), (A.3.14), (A.3.15) and (A.3.17) yield

$$\begin{aligned}
\beta_{11}\xi_1\xi_2(\eta_2-\eta_1) &= f(0)\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_1) - f(\eta_1)\xi_1\xi_2(\eta_2-\eta_1), \\
\beta_{22}\xi_1\xi_2(\eta_2-\eta_1) &= f(0)\eta_1(\xi_1\eta_2 + \xi_2\eta_2 - \xi_1\xi_2) - f(\eta_2)\xi_1\xi_2(\eta_2-\eta_1).
\end{aligned}$$

Also, use (A.3.13)–(A.3.15) and (A.3.17) to obtain

$$\beta_{12}^2\xi_1^2\xi_2^2(\eta_2-\eta_1)^2 = f(0)^2\eta_1\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_2)(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2).$$

Therefore,

$$\begin{aligned}
0 &\le (\beta_{11}\beta_{22} - \beta_{12}^2)\xi_1^2\xi_2^2(\eta_2-\eta_1)^2 \\
&= \{f(0)\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_1) - f(\eta_1)\xi_1\xi_2(\eta_2-\eta_1)\} \\
&\quad \times \{f(0)\eta_1(\xi_1\eta_2 + \xi_2\eta_2 - \xi_1\xi_2) - f(\eta_2)\xi_1\xi_2(\eta_2-\eta_1)\} \\
&\quad - f(0)^2\eta_1\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_2)(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2) \\
&= f(0)^2\xi_1\xi_2\eta_1\eta_2(\eta_2-\eta_1)^2 - f(0)f(\eta_1)\xi_1\xi_2\eta_1(\eta_2-\eta_1)(\xi_1\eta_2 + \xi_2\eta_2 - \xi_1\xi_2) \\
&\quad - f(0)f(\eta_2)\xi_1\xi_2\eta_2(\eta_2-\eta_1)(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_1) + f(\eta_1)f(\eta_2)\xi_1^2\xi_2^2(\eta_2-\eta_1)^2.
\end{aligned}$$

Since $\xi_1\xi_2(\eta_2-\eta_1) < 0$, we have

$$\begin{aligned}
&f(0)^2\eta_1\eta_2(\eta_2-\eta_1) - f(0)f(\eta_1)\eta_1(\xi_1\eta_2 + \xi_2\eta_2 - \xi_1\xi_2) \\
&\quad - f(0)f(\eta_2)\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_1) + f(\eta_1)f(\eta_2)\xi_1\xi_2(\eta_2-\eta_1) \le 0.
\end{aligned} \tag{A.3.19}$$

Step 4. What we have to prove is (A.3.5) with $\xi = 0$, i.e.,

$$f^{[2]}(\xi_1,0,\eta_1)f^{[2]}(\xi_2,0,\eta_2)(\xi_1-\eta_1)(\xi_2-\eta_2) + f^{[2]}(\xi_1,0,\xi_2)f^{[2]}(\eta_1,0,\eta_2)(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2) \le 0.$$

By (A.3.7) this means that

$$\begin{aligned}
&\left\{\frac{-f(0)}{\xi_1} - \frac{f(\eta_1)-f(0)}{\eta_1}\right\}\left\{\frac{-f(0)}{\xi_2} - \frac{f(\eta_2)-f(0)}{\eta_2}\right\} \\
&\quad + \frac{\frac{-f(0)}{\xi_1} - \frac{-f(0)}{\xi_2}}{\xi_1-\xi_2} \cdot \frac{\frac{f(\eta_1)-f(0)}{\eta_1} - \frac{f(\eta_2)-f(0)}{\eta_2}}{\eta_1-\eta_2}(\xi_1\eta_2 + \xi_2\eta_1 - \xi_1\xi_2) \le 0,
\end{aligned}$$

which is rewritten as

$$\begin{aligned}
&\frac{1}{\xi_1\xi_2\eta_1\eta_2(\eta_2-\eta_1)}\{f(0)^2\eta_1\eta_2(\eta_2-\eta_1) - f(0)f(\eta_1)\eta_1(\xi_1\eta_2 + \xi_2\eta_2 - \xi_1\xi_2) \\
&\qquad - f(0)f(\eta_2)\eta_2(\xi_1\xi_2 - \xi_1\eta_1 - \xi_2\eta_1) + f(\eta_1)f(\eta_2)\xi_1\xi_2(\eta_2-\eta_1)\} \le 0.
\end{aligned}$$

Since $\xi_1\xi_2\eta_1\eta_2(\eta_2-\eta_1) > 0$, this is nothing but (A.3.19). □

Lemma A.3.3. *If $f^{[2]}(\xi_1,\xi_2,\xi_3) = 0$ for some distinct ξ_1,ξ_2,ξ_3 in (a,b), then f is linear on (a,b).*

Proof. First, note that $f^{[2]}(x,y,z) \ge 0$ for all distinct x,y,z in (a,b) since f is convex in the usual sense. Assume that $f^{[2]}(\xi_1,\eta_1,\xi) = 0$ for some $\xi_1 < \eta_1 < \xi$ in (a,b). Then f is linear on $[\xi_1,\xi]$ and so $f^{[2]}(\xi_1,\eta_1,\xi) = 0$ for all $\eta_1 \in (\xi_1,\xi)$. For any η_2,ξ_2 with $\xi < \eta_2 < \xi_2$, the left-hand side of (A.3.3) is positive if η_1 is sufficiently close to ξ, since it is positive when $\eta_1 = \xi$. Since (A.3.5) holds for such $\eta_1 \in (\xi_1,\xi)$ by Lemma A.3.1, we have

$$f^{[2]}(\xi_1,\xi,\xi_2)f^{[2]}(\eta_1,\xi,\eta_2) \le 0$$

so that $f^{[2]}(\xi_1,\xi,\xi_2) = 0$ or $f^{[2]}(\eta_1,\xi,\eta_2) = 0$. When $f^{[2]}(\eta_1,\xi,\eta_2) = 0$, f is linear on $[\eta_1,\eta_2]$ and hence f is linear on $[\xi_1,\eta_2]$. This is the case also when $f^{[2]}(\xi_1,\xi,\xi_2) = 0$ and so f is linear on $[\xi_1,\xi_2]$. Since $\eta_2 < \xi_2$ can be arbitrarily close to b, it follows that f is linear on $[\xi_1, b)$. From $f^{[2]}(\xi, \eta_2, \xi_2) = 0$ where $\xi < \eta_2 < \xi_2$, it can be similarly shown that f is linear on $(a, \xi_2]$. Thus f is linear on the whole (a, b). □

From now on, we assume that $f^{[2]}(x, y, z) > 0$ for all distinct x, y, z in (a, b), i.e., f is strictly convex on $(a.b)$.

Lemma A.3.4. *f is C^1 on (a, b).*

Proof. By Lemma A.3.1 (i) it suffices to prove that, for any $\xi_0 \in (a, b)$, there exist ξ_1', ξ_2' such that $\xi_1' < \xi_0 < \xi_2'$ and $f^{[2]}(\eta_1, \xi, \eta_2)$ is uniformly bounded when $\eta_1 < \xi < \eta_2$ run over $[\xi_1', \xi_2']$. First choose ξ_1, ξ_2 with $a < \xi_1 < \xi_0 < \xi_2 < b$. Since the left-hand side of (A.3.3) is $(\xi_0 - \xi_1)(\xi_2 - \xi_0) > 0$ when $\eta_1 = \eta_2 = \xi = \xi_0$, one can choose ξ_1', ξ_2' such that $\xi_1 < \xi_1' < \xi_0 < \xi_2' < \xi_2$ and

$$(\xi_1 - \xi)(\eta_2 - \xi) + (\xi_2 - \xi)(\eta_1 - \xi) - (\xi_1 - \xi)(\xi_2 - \xi) \geq C \tag{A.3.20}$$

for any choice of $\eta_1 \leq \xi \leq \eta_2$ in $[\xi_1', \xi_2']$, where $C > 0$ is a constant. Then it follows from (A.3.5) that for any $\eta_1 < \xi < \eta_2$ in $[\xi_1', \xi_2']$,

$$\begin{aligned} 0 < f^{[2]}(\eta_1, \xi, \eta_2) &\leq \frac{f^{[2]}(\xi_1, \xi, \eta_1) f^{[2]}(\xi_2, \xi, \eta_2)}{f^{[2]}(\xi_1, \xi, \xi_2)} \\ &\quad \times \frac{(\eta_1 - \xi_1)(\xi_2 - \eta_2)}{(\xi_1 - \xi)(\eta_2 - \xi) + (\xi_2 - \xi)(\eta_1 - \xi) - (\xi_1 - \xi)(\xi_2 - \xi)} \\ &\leq \frac{\{f^{[1]}(\eta_1, \xi) - f^{[1]}(\xi_1, \xi)\}\{f^{[1]}(\xi, \xi_2) - f^{[1]}(\xi, \eta_2)\}}{f^{[2]}(\xi_1, \xi, \xi_2) C}. \end{aligned} \tag{A.3.21}$$

Since f is strictly convex in the usual sense, it is obvious that $f^{[1]}(\xi_1, \xi_1') < f^{[1]}(\eta_1, \xi) < f^{[1]}(\xi_2', \xi_2)$ and $f^{[1]}(\xi_1, \xi_1') < f^{[1]}(\xi, \eta_2) < f^{[1]}(\xi, \xi_2)$. Furthermore, $f^{[2]}(\xi_1, \xi, \xi_2)$ attains the minimum value γ when ξ runs over $[\xi_1', \xi_2']$, which is positive. Therefore,

$$\begin{aligned} &\sup\{f^{[2]}(\eta_1, \xi, \eta_2) : \xi_1' \leq \eta_1 < \xi < \eta_2 \leq \xi_2'\} \\ &\quad \leq \frac{\{f^{[1]}(\xi_2', \xi_2) - f^{[1]}(\xi_1, \xi)\}\{f^{[1]}(\xi, \xi_2) - f^{[1]}(\xi_1, \xi_1')\}}{\gamma C} < +\infty. \end{aligned}$$ □

Lemma A.3.5. *f is C^2 on (a, b).*

Proof. Since the C^1 of f has been shown in Lemma A.3.4, we can extend the second divided difference $f^{[2]}$ to $(x, y, z) \in (a, b)^3$, in which two of the variables coincide but the third is different from them, as follows:

$$f^{[2]}(\xi, \xi, \eta) := \lim_{x \to \xi} f^{[2]}(x, \xi, \eta) = \frac{f'(\xi) - \frac{f(\xi) - f(\eta)}{\xi - \eta}}{\xi - \eta}.$$

Let $[\xi_1, \xi_2]$ be any closed interval included in (a, b). For any $\xi < \eta_2$ in (ξ_1, ξ_2), the left-hand side of (A.3.3) is positive if $\eta_1 \in (\xi_1, \xi)$ is sufficiently close to ξ, since it is positive when $\eta_1 = \xi$. Since (A.3.21) holds for all such $\eta_1 < \xi$, letting $\eta_1 \nearrow \xi$ in (A.3.21) we have

$$f^{[2]}(\xi, \xi, \eta_2) \leq \frac{f^{[2]}(\xi_1, \xi, \xi) f^{[2]}(\xi, \eta_2, \xi_2)}{f^{[2]}(\xi_1, \xi, \xi_2)}$$

so that

$$f^{[2]}(\xi, \xi, \eta_2) - f^{[2]}(\xi_1, \xi, \xi) \leq f^{[2]}(\xi_1, \xi, \xi) \frac{f^{[2]}(\eta_2, \xi, \xi_2) - f^{[2]}(\xi_1, \xi, \xi_2)}{f^{[2]}(\xi_1, \xi, \xi_2)}.$$

Dividing this by $\eta_2 - \xi_1 > 0$ yields

$$f^{[3]}(\xi_1, \xi, \xi, \eta_2) \leq f^{[2]}(\xi_1, \xi, \xi) \frac{f^{[3]}(\xi_1, \xi, \eta_2, \xi_2)}{f^{[2]}(\xi_1, \xi, \xi_2)} \tag{A.3.22}$$

whenever $\xi_1 < \xi < \eta_2 < \xi_2$. Next, for any $\eta_1 < \xi$ in (ξ_1, ξ_2), the left-hand side of (A.3.3) is positive if $\eta_2 \in (\xi, \xi_2)$ is sufficiently close to ξ. Letting $\eta_2 \searrow \xi$ in (A.3.21) we have

$$f^{[2]}(\eta_1, \xi, \xi) \leq \frac{f^{[2]}(\xi_1, \xi, \eta_1) f^{[2]}(\xi_2, \xi, \xi)}{f^{[2]}(\xi_1, \xi, \xi_2)}$$

so that

$$f^{[2]}(\eta_1, \xi, \xi) - f^{[2]}(\xi, \xi, \xi_2) \leq f^{[2]}(\xi, \xi, \xi_2) \frac{f^{[2]}(\xi_1, \xi, \eta_1) - f^{[2]}(\xi_1, \xi, \xi_2)}{f^{[2]}(\xi_1, \xi, \xi_2)}.$$

Dividing this by $\eta_1 - \xi_2 < 0$ yields

$$f^{[3]}(\eta_1, \xi, \xi, \xi_2) \geq f^{[2]}(\xi, \xi, \xi_2) \frac{f^{[3]}(\xi_1, \xi, \eta_1, \xi_2)}{f^{[2]}(\xi_1, \xi, \xi_2)} \tag{A.3.23}$$

whenever $\xi_1 < \eta_1 < \xi < \xi_2$.

For each $\xi_0 \in (a,b)$, the proof of Lemma A.3.4 implies that there exist ξ_1'', ξ_2'' such that $\xi_1'' < \xi_0 < \xi_1''$ and

$$\sup\{f^{[2]}(x, \xi, y) : \xi_1'' \leq x < \xi < y \leq \xi_2''\} < +\infty \tag{A.3.24}$$

(i.e., ξ_1'', ξ_2'' are ξ_1', ξ_2' in the proof of Lemma A.3.4). Choose ξ_1', ξ_2' with $\xi_1'' < \xi_1' < \xi_0 < \xi_2' < \xi_2''$. Apply (A.3.22) to show that $f^{[3]}(\xi_1, \xi, \xi, \eta_2)$ is bounded from above when $\xi_1 < \xi < \eta_2$ run over $[\xi_1', \xi_2']$. Fix $\xi_2 := \xi_2''$. Since $(\xi_1, \xi) \mapsto f^{[2]}(\xi_1, \xi, \xi_2)$ is continuous on $[\xi_1', \xi_2']^2$, the minimum value is attained at some $(\xi_1^0, \xi^0) \in [\xi_1', \xi_2']^2$. Since f is strictly convex, $f^{[2]}(\xi_1^0, \xi^0, \xi_2) > 0$ whether $\xi_1^0 = \xi^0$ or $\xi_1^0 \neq \xi^0$. Note that

$$f^{[3]}(\xi_1, \xi, \eta_2, \xi_2) = \frac{f^{[2]}(\xi_1, \xi, \eta_2) - f^{[2]}(\xi, \eta_2, \xi_2)}{\xi_1 - \xi_2}.$$

Since $|\xi_1 - \xi_2| \geq \xi_2'' - \xi_2' > 0$, it follows from (A.3.24) that $f^{[3]}(\xi_1, \xi, \eta_2, \xi_2)$ is bounded from above. Moreover, since $f^{[2]}(\xi_1, \xi, \xi) = \lim_{y \to \xi} f^{[2]}(\xi_1, \xi, y)$, $f^{[2]}(\xi_1, \xi, \xi)$ is also bounded from above by (A.3.24). Hence the required boundedness from above of $f^{[3]}(\xi_1, \xi, \xi, \eta_2)$ is proved.

On the other hand, apply (A.3.23) to show that $f^{[3]}(\eta_1, \xi, \xi, \xi_2)$ is bounded from below when $\eta_1 < \xi < \xi_2$ run over $[\xi_1', \xi_2']$. Fix $\xi_1 := \xi_1''$. In the same way as above, $f^{[2]}(\xi_1, \xi, \xi_2)$ attains the positive minimum value, $f^{[3]}(\xi_1, \xi, \eta_1, \xi_2)$ is bounded from below, and $f^{[2]}(\xi, \xi, \xi_2) > 0$ is bounded from above. Hence the required bounded from below is proved. We thus conclude that

$$K := \sup\{|f^{[3]}(\xi_1, \xi, \xi, \xi_2)| : \xi_1' \leq \xi_1 < \xi < \xi_2 \leq \xi_2'\} < +\infty. \tag{A.3.25}$$

Since f is C^1 on (a,b), one can compute

$$\begin{aligned}
\frac{d}{d\xi} f^{[2]}(\xi_1, \xi, \xi_2) &= \frac{d}{d\xi} \left\{ \frac{\frac{f(\xi) - f(\xi_1)}{\xi - \xi_1} - \frac{f(\xi) - f(\xi_2)}{\xi - \xi_2}}{\xi_1 - \xi_2} \right\} \\
&= \frac{1}{\xi_1 - \xi_2} \left\{ \frac{f'(\xi) - \frac{f(\xi) - f(\xi_1)}{\xi - \xi_1}}{\xi - \xi_1} - \frac{f'(\xi) - \frac{f(\xi) - f(\xi_2)}{\xi - \xi_2}}{\xi - \xi_2} \right\} \\
&= \frac{f^{[2]}(\xi, \xi, \xi_1) - f^{[2]}(\xi, \xi, \xi_2)}{\xi_1 - \xi_2} = f^{[3]}(\xi_1, \xi, \xi, \xi_2).
\end{aligned}$$

Hence (A.3.25) is rephrased as

$$\left| \frac{d}{d\xi} f^{[2]}(\xi_1, \xi, \xi_2) \right| \leq K$$

for all $\xi_1 < \xi < \xi_2$ in $[\xi_1', \xi_2']$. For any $\xi_1 < \lambda_1 < \lambda_2 < \xi_2$ in $[\xi_1', \xi_2']$, we have

$$f^{[3]}(\xi_1, \lambda_1, \lambda_2, \xi_2) = \frac{f^{[2]}(\xi_1, \lambda_2, \xi_2) - f^{[2]}(\xi_1, \lambda_1, \xi_2)}{\lambda_2 - \lambda_1} = \frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} \frac{d}{d\xi} f^{[2]}(\xi_1, \xi, \xi_2)\, d\xi$$

so that

$$|f^{[3]}(\xi_1, \lambda_1, \lambda_2, \xi_2)| \leq \frac{1}{\lambda_2 - \lambda_1} \int_{\lambda_1}^{\lambda_2} \left| \frac{d}{d\xi} f^{[2]}(\xi_1, \xi, \xi_2) \right| d\xi \leq K.$$

Hence Lemma A.3.1 (ii) implies that f is C^2 on (ξ_1', ξ_2'). □

Lemma A.3.5 together with Lemma A.3.3 proves Theorem 2.4.2.

Remark A.3.6. Theorem 2.4.2 tells that if f is a non-linear 2-convex function on (a,b), then it is a strictly convex C^2 function there. Then $f^{[2]}(x,y,z) > 0$ for all $x, y, z \in (a,b)$ unless x, y, z are all identical. If f is a non-linear 3-convex function on (a,b), then Corollary 2.4.6 shows that $f^{[1]}(\lambda, \cdot)$ is a non-constant 2-monotone function for every $\lambda \in (a,b)$. Hence Theorem 2.4.1 implies that $\frac{d}{dx} f^{[1]}(\lambda, x) = f^{[2]}(\lambda, x, x) > 0$ for all $\lambda, x \in (a,b)$ and so $f''(x) > 0$ for all $x \in (a,b)$. We do not know whether there is a non-linear 2-convex function f on (a,b) with $f''(x) = 0$ for some $x \in (a,b)$.

A.4 Proof of Nevanlinna's theorem, Theorem 2.6.2

To prove Theorem 2.6.2, we utilize the Poisson integral representation for analytic and also harmonic functions in the unit disk $\mathbb{D} := \{\zeta \in \mathbb{C} : |\zeta| < 1\}$. We begin with the following lemma.

Lemma A.4.1. *If f is an analytic function in a domain containing the closed unit disk $\overline{\mathbb{D}} := \{\zeta \in \mathbb{C} : |\zeta| \le 1\}$, then*

$$f(\zeta) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\frac{1-|\zeta|^2}{|e^{it}-\zeta|^2}\,dt, \qquad \zeta \in \mathbb{D}.$$

Proof. For every $\zeta \in \mathbb{D}$, the Cauchy integral formula implies that

$$f(\zeta) = \frac{1}{2\pi i}\int_{|z|=1} \frac{f(z)}{z-\zeta}\,dz = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{it})e^{it}}{e^{it}-\zeta}\,dt. \tag{A.4.1}$$

On the other hand, if $\zeta \in \mathbb{D} \setminus \{0\}$ and so $|1/\overline{\zeta}| > 1$, then the Cauchy integral theorem implies that

$$0 = \frac{1}{2\pi i}\int_{|z|=1} \frac{f(z)}{z-(1/\overline{\zeta})}\,dz = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{it})e^{it}}{e^{it}-(1/\overline{\zeta})}\,dt = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{it})\overline{\zeta}}{\overline{\zeta}-e^{-it}}\,dt. \tag{A.4.2}$$

The above last integral is zero also when $\zeta = 0$. By (A.4.1) and (A.4.2) we have

$$f(\zeta) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\left(\frac{e^{it}}{e^{it}-\zeta} - \frac{\overline{\zeta}}{\overline{\zeta}-e^{-it}}\right)dt = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,\frac{1-|\zeta|^2}{|e^{it}-\zeta|^2}\,dt. \qquad \square$$

The kernel function in the above integral representation is the so-called *Poisson kernel* that is rewritten as

$$\frac{1-|\zeta|^2}{|e^{it}-\zeta|^2} = \frac{1-r^2}{1-2r\cos(t-\theta)+r^2}, \qquad t \in [0,2\pi], \tag{A.4.3}$$

for each $\zeta = re^{i\theta}$ with $0 \le r < 1$.

A real-valued function φ on a simply connected domain D is a *harmonic function* if and only if it is the real part of an analytic function f in D (see [27, p. 202]). Moreover, it is well known that such an analytic function f in D is unique up to a pure imaginary additive constant.

Theorem A.4.2. *If φ is a nonnegative harmonic function on $\mathbb{D}$, then there exists a positive finite Borel measure σ on $[0,2\pi]$ such that*

$$\varphi(re^{i\theta}) = \int_0^{2\pi} \frac{1-r^2}{1-2r\cos(t-\theta)+r^2}\,d\sigma(t), \qquad re^{i\theta} \in \mathbb{D}.$$

Proof. There exists an analytic function f in $\mathbb{D}$ such that $\varphi(\zeta) = \operatorname{Re} f(\zeta)$, $\zeta \in \mathbb{D}$. For each $\rho \in (0,1)$, since $f(\rho\zeta)$ is analytic in $|\zeta| < 1/\rho$, Lemma A.4.1 implies that

$$f(\rho\zeta) = \frac{1}{2\pi}\int_0^{2\pi} f(\rho e^{it})\,\frac{1-r^2}{1-2r\cos(t-\theta)+r^2}\,dt$$

so that

$$\varphi(\rho\zeta) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(\rho e^{it})\,\frac{1-r^2}{1-2r\cos(t-\theta)+r^2}\,dt$$

for all $\zeta = re^{i\theta} \in \mathbb{D}$. Define a positive Borel measure σ_ρ on $[0,2\pi]$ by $d\sigma_\rho(t) := (1/2\pi)\varphi(\rho e^{it})\,dt$. Then

$$\int_0^{2\pi} d\sigma_\rho(t) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(\rho e^{it})\,dt = \varphi(0).$$

We now consider $\{\sigma_\rho : \rho \in (0,1)\}$ as a subset of the set Σ of positive linear functionals on $C([0,2\pi])$ with norm $\varphi(0)$, where $C([0,2\pi])$ is the Banach space of complex functions on $[0,2\pi]$ with sup-norm. Since the set Σ is compact and metrizable in the weak* topology, one can choose a sequence $\rho_n \in (0,1)$ with $\rho_n \nearrow 1$ such that σ_{ρ_n} converges in the weak* topology to some $\rho \in \Sigma$ regarded as a positive finite Borel measure on $[0,2\pi]$. For every $\zeta = re^{i\theta} \in \mathbb{D}$ we then have

$$\begin{aligned}\varphi(\zeta) &= \lim_{n\to\infty} \varphi(\rho_n\zeta) = \lim_{n\to\infty} \int_0^{2\pi} \frac{1-r^2}{1-2r\cos(t-\theta)+r^2}\,d\sigma_n(t)\\ &= \int_0^{2\pi} \frac{1-r^2}{1-2r\cos(t-\theta)+r^2}\,d\sigma(t). \qquad \square\end{aligned}$$

The Poisson integral representation in Theorem A.4.2 as well as in the next theorem is sometimes called the *Helglotz theorem*.

Theorem A.4.3. *If f is an analytic function in $\mathbb{D}$ with nonnegative real part, then there exists a positive finite Borel measure σ on $[0, 2\pi]$ such that*

$$f(\zeta) = \int_0^{2\pi} \frac{e^{it} + \zeta}{e^{it} - \zeta}\, d\sigma(t) + i\, \mathrm{Im}\, f(0), \qquad \zeta \in \mathbb{D}. \tag{A.4.4}$$

Proof. Let σ be the positive finite measure on $[0, 2\pi]$ taken in Theorem A.4.2 for $\varphi = \mathrm{Re}\, f$. We write $g(\zeta)$ for the first term (the integral term) of the right-hand side of (A.4.4). For each $\zeta \in \mathbb{D}$, since

$$\frac{g(\zeta + \Delta\zeta) - g(\zeta)}{\Delta\zeta} = \int_0^{2\pi} \frac{2e^{it}\zeta}{(e^{it} - \zeta)(e^{it} - \zeta - \Delta\zeta)}\, d\sigma(t)$$

and

$$\left| \frac{e^{it}\zeta}{(e^{it} - \zeta)(e^{it} - \zeta - \Delta\zeta)} \right| \le \frac{|\zeta|}{(1 - |\zeta|)(1 - |\zeta| - |\Delta\zeta|)} \le \frac{2|\zeta|}{(1 - |\zeta|)^2}$$

for all $t \in [0, 2\pi]$ and for all $\Delta\zeta \in \mathbb{C}$ with $|\Delta\zeta| < (1 - |\zeta|)/2$, it follows from the Lebesgue dominated convergence theorem that

$$\lim_{\Delta\zeta \to 0} \frac{g(\zeta + \Delta\zeta) - g(\zeta)}{\Delta\zeta} = \int_0^{2\pi} \frac{2e^{it}\zeta}{(e^{it} - \zeta)^2}\, d\sigma(t).$$

Hence g is analytic in $\mathbb{D}$. Since

$$\mathrm{Re}\left(\frac{e^{it} + \zeta}{e^{it} - \zeta} \right) = \frac{1 - |\zeta|^2}{|e^{it} - \zeta|^2}$$

is the Poisson kernel in (A.4.3), we have $\mathrm{Re}\, f(\zeta) = \mathrm{Re}\, g(\zeta)$ so that $f(\zeta) = g(\zeta) + ib$, $\zeta \in \mathbb{D}$, for some $b \in \mathbb{R}$. Letting $\zeta = 0$ gives $b = \mathrm{Im}\, f(0)$ thanks to $g(0) \in \mathbb{R}$. Hence the desired expression is obtained. □

Now we turn to the proof of Nevanlinna's theorem (Theorem 2.6.2). Assume that f is a non-constant Pick function, i.e., an analytic function in $\mathbb{C}^+$ such that $f(\mathbb{C}^+) \subset \mathbb{C}^+$. The proof is to rephrase Theorem A.4.3 by transforming f in $\mathbb{C}^+$ to an analytic function in $\mathbb{D}$ by the *Cayley transform* that is the fractional linear transform given by

$$\zeta(z) := \frac{z - i}{z + i}.$$

Exercise A.4.4. Show that the Cayley transform $\zeta(z)$ maps $\mathbb{C}^+$ bijectively to $\mathbb{D}$ with the inverse $z(\zeta) = i(1 + \zeta)/(1 - \zeta)$, and that $\zeta(z)$ maps the real line $(-\infty, \infty)$ bijectively to the unit circle $\{e^{it} : 0 < t < 2\pi\}$ but 1.

From this exercise, it follows that $-if(z(\zeta))$ is an analytic function in $\mathbb{D}$ with nonnegative real part. Hence Theorem A.4.3 implies that there exists a positive finite measure σ on $[0, 2\pi]$ such that

$$\begin{aligned} -if(z(\zeta)) &= \int_0^{2\pi} \frac{e^{it} + \zeta}{e^{it} - \zeta}\, d\sigma(t) + i\, \mathrm{Im}(-if(z(0))) \\ &= \int_0^{2\pi} \frac{e^{it} + \zeta}{e^{it} - \zeta}\, d\sigma(t) - i\, \mathrm{Re}\, f(i), \qquad \zeta \in \mathbb{D}. \end{aligned}$$

With $\alpha := \mathrm{Re}\, f(i) \in \mathbb{R}$ and $\beta := \sigma(\{0, 2\pi\}) \ge 0$, the above integral expression is rewritten as

$$f(z(\zeta)) = \alpha + \beta i\, \frac{1 + \zeta}{1 - \zeta} + \int_{(0,2\pi)} i\, \frac{e^{it} + \zeta}{e^{it} - \zeta}\, d\sigma(t), \qquad \zeta \in \mathbb{D}. \tag{A.4.5}$$

Let ν be the positive finite Borel measure on $\mathbb{R}$ transformed from $\sigma|_{(0,2\pi)}$ via the map $t \in (0, 2\pi) \mapsto \lambda = z(e^{it}) \in (-\infty, \infty)$ or $e^{it} = \zeta(\lambda)$. Substituting $\zeta(z)$ for ζ and $\zeta(\lambda)$ for e^{it} in (A.4.5) we have

$$f(z) = \alpha + \beta z + \int_{-\infty}^{\infty} i\, \frac{\zeta(\lambda) + \zeta(z)}{\zeta(\lambda) - \zeta(z)}\, d\nu(\lambda), \qquad z \in \mathbb{C}^+.$$

Since

$$i\, \frac{\zeta(\lambda) + \zeta(z)}{\zeta(\lambda) - \zeta(z)} = i\, \frac{\frac{\lambda - i}{\lambda + i} + \frac{z - i}{z + i}}{\frac{\lambda - i}{\lambda + i} - \frac{z - i}{z + i}} = \frac{1 + \lambda z}{\lambda - z},$$

we arrive at

$$f(z) = \alpha + \beta z + \int_{-\infty}^{\infty} \frac{1 + \lambda z}{\lambda - z}\, d\nu(\lambda), \qquad z \in \mathbb{C}^+.$$

A.5 Proofs of Fourier transforms for hyperbolic functions

We prove the formulas (5.1.6)–(5.1.8) of Fourier transforms for certain ratios of hyperbolic functions, which play crucial roles in Chapter 5. The proofs below are based on the residue theorem in complex function theory. A similar computation for (5.1.6) is found in [31].

Proof of (5.1.6). We may assume that $0 < a < b$ and $t > 0$. In fact, the case $t < 0$ holds by symmetry and the case $t = 0$ by continuity from the case $t \neq 0$. Define a complex function

$$f(z) := \frac{e^{itz}}{\cosh(\frac{\pi}{b} z) + \cos(\frac{\pi a}{b})},$$

which is holomorphic in $\mathbb{C}$ except the points z where the denominator vanishes. It is clear that those exceptional points are poles of $f(z)$. The equation $\cosh(\frac{\pi}{b} z) + \cos(\frac{\pi a}{b}) = 0$ means that

$$e^{\frac{2\pi}{b} z} + 2\cos\left(\frac{\pi a}{b}\right) e^{\frac{\pi}{b} z} + 1 = 0,$$

which is solved as

$$e^{\frac{\pi}{b} z} = -\cos\left(\frac{\pi a}{b}\right) \pm i \sin\left(\frac{\pi a}{b}\right) = -e^{\mp i \frac{\pi a}{b}}$$

so that $e^{\frac{\pi}{b}(z \pm ia)} = -1$. Therefore, the poles of $f(z)$ are

$$z_n^- := i((2n-1)b - a), \quad z_n^+ := i((2n-1)b + a), \qquad n \in \mathbb{Z}.$$

The residues of $f(z)$ at $z_n^{\mp}$ are computed as

$$\begin{aligned}
\operatorname{Res}(z_n^{\mp}; f) &= \frac{e^{itz_n^{\mp}}}{\frac{d}{dz}\cosh(\frac{\pi}{b} z)|_{z=z_n^{\mp}}} = \frac{e^{-t((2n-1)b \mp a)}}{\frac{2\pi}{b}\,(e^{\frac{\pi}{b} i((2n-1)b \mp a)} - e^{-\frac{\pi}{b} i((2n-1)b \mp a)})} \\
&= \frac{e^{\pm at} e^{-(2n-1)bt}}{\frac{2\pi}{b}\,(-e^{\mp i \frac{\pi a}{b}} + e^{\pm i \frac{\pi a}{b}})} = \frac{b}{\pi i} \cdot \frac{\pm e^{\pm at} e^{-(2n-1)bt}}{\sin(\frac{\pi a}{b})}.
\end{aligned}$$

Now, for $R > 0$ we take the contour

$$\Gamma_R : \begin{cases} z = x, & -R \le x \le R, \\ z = R + iy, & 0 \le y \le R, \\ z = x + iR, & R \ge x \ge -R, \\ z = -R + iy, & R \ge y \ge 0. \end{cases}$$

When $(2n-1)b - a < R < (2n-1)b + a$, the residue theorem implies that

$$\int_{\Gamma_R} f(z)\,dz = 2\pi i \left\{ \sum_{k=1}^{n} \operatorname{Res}(z_k^-; f) + \sum_{k=1}^{n-1} \operatorname{Res}(z_k^+; f) \right\}$$

so that

$$\begin{aligned}
&\int_{-R}^{R} f(x)\,dx + \int_0^R f(R+iy) i\,dy - \int_{-R}^{R} f(x+iR)\,dx - \int_0^R f(-R+iy) i\,dy \\
&\quad = \frac{2b}{\sin(\frac{\pi a}{b})} \left\{ e^{at} \sum_{k=1}^{n} e^{-(2k-1)bt} - e^{-at} \sum_{k=1}^{n-1} e^{-(2k-1)bt} \right\}.
\end{aligned} \tag{A.5.1}$$

Since

$$\begin{aligned}
|f(\pm R + iy)| &= \frac{e^{-ty}}{|\cosh(\frac{\pi R}{b}) \cos(\frac{\pi}{b} y) \pm i \sinh(\frac{\pi R}{b}) \sin(\frac{\pi}{b} y) + \cos(\frac{\pi a}{b})|} \\
&\le \frac{e^{-ty}}{\sqrt{\cosh^2(\frac{\pi R}{b}) \cos^2(\frac{\pi}{b} y) + \sinh^2(\frac{\pi R}{b}) \sin^2(\frac{\pi}{b} y)} - 1} \\
&\le \frac{e^{-ty}}{\sinh(\frac{\pi R}{b}) - 1},
\end{aligned}$$

we have

$$\int_0^R |f(\pm R+iy)|\,dy \le \frac{1}{t(\sinh(\frac{\pi R}{b})-1)} \longrightarrow 0 \quad \text{as } R\to\infty. \tag{A.5.2}$$

Moreover, for the particular choice $R=(2n-1)b$ we have

$$\begin{aligned}|f(x+iR)| &= \frac{e^{-tR}}{|\cosh(\frac{\pi}{b}x)\cos(\frac{\pi R}{b})+i\sinh(\frac{\pi}{b}x)\sin(\frac{\pi R}{b})+\cos(\frac{\pi a}{b})|}\\ &= \frac{e^{-tR}}{|-\cosh(\frac{\pi}{b}x)+\cos(\frac{\pi a}{b})|} \le \frac{e^{-tR}}{1-\cos(\frac{\pi a}{b})}\end{aligned}$$

so that

$$\int_{-R}^R |f(x+iR)|\,dx \le \frac{2Re^{-tR}}{1-\cos(\frac{\pi a}{b})} \longrightarrow 0 \quad \text{as } R=(2n-1)b\to\infty. \tag{A.5.3}$$

Hence, letting $R=(2n-1)b\to\infty$ in (A.5.1) and using (A.5.2) and (A.5.3), we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\,dx &= 2b\,\frac{e^{at}-e^{-at}}{\sin(\frac{\pi a}{b})}\sum_{k=1}^{\infty} e^{-(2k-1)bt}\\ &= \frac{2b}{\sin(\frac{\pi a}{b})}\cdot\frac{e^{at}-e^{-at}}{e^{bt}-e^{-bt}} = \frac{2b}{\sin(\frac{\pi a}{b})}\cdot\frac{\sinh(at)}{\sinh(bt)},\end{aligned}$$

which becomes (5.1.6). □

Proof of (5.1.7). As in the proof of (5.1.6) we may assume that $0<a<b$ and $t>0$. Define

$$g(z) := \frac{e^{itz}\cosh(\frac{\pi}{2b}z)}{\cosh(\frac{\pi}{b}z)+\cos(\frac{\pi a}{b})} = f(z)\cosh\left(\frac{\pi}{2b}z\right),$$

which is holomorphic in $\mathbb{C}$ except the poles $z_n^{\mp}$, $n\in\mathbb{Z}$, as above. The residues of $g(z)$ at $z_n^{\mp}$ are computed as

$$\begin{aligned}\mathrm{Res}(z_n^{\mp};g) &= \mathrm{Res}(z_n^{\mp};f)\cosh\left(\frac{\pi}{2b}z_n^{\mp}\right) = \frac{b}{\pi i}\cdot\frac{\pm e^{\pm at}e^{-(2n-1)bt}\cos(\pi(n-\frac12)\mp\frac{\pi a}{2b})}{\sin(\frac{\pi a}{b})}\\ &= \frac{b}{\pi i}\cdot\frac{\pm e^{\pm at}e^{-(2n-1)bt}(\pm(-1)^{n-1}\sin(\frac{\pi a}{2b}))}{\sin(\frac{\pi a}{b})} = \frac{b}{2\pi i}\cdot\frac{(-1)^{n-1}e^{\pm at}e^{-(2n-1)bt}}{\cos(\frac{\pi a}{2b})}.\end{aligned}$$

For the same contour Γ_R as above with $R\in((2n-1)b-a,(2n-1)b+a)$, the residue theorem implies that

$$\begin{aligned}&\int_{-R}^R g(x)\,dx + \int_0^R g(R+iy)i\,dy - \int_{-R}^R g(x+iR)\,dx - \int_0^R g(-R+iy)i\,dy\\ &= \frac{b}{\cos(\frac{\pi a}{2b})}\left\{e^{at}\sum_{k=1}^{n}(-1)^{k-1}e^{-(2k-1)bt} + e^{-at}\sum_{k=1}^{n-1}(-1)^{k-1}e^{-(2k-1)bt}\right\}.\end{aligned} \tag{A.5.4}$$

Since

$$|g(\pm R+iy)| = \left|f(\pm R+iy)\cosh\left(\frac{\pi}{2b}(\pm R+iy)\right)\right| \le \frac{2e^{-ty}\cosh(\frac{\pi R}{2b})}{\sinh(\frac{\pi R}{b})-1},$$

we have

$$\int_0^R |g(\pm R+iy)|\,dy \le \frac{2\cosh(\frac{\pi R}{2b})}{t(\sinh(\frac{\pi R}{b})-1)} \longrightarrow 0 \quad \text{as } R\to\infty. \tag{A.5.5}$$

Moreover, for the particular choice $R=(2n-1)b$ we have

$$|g(x+iR)| = \left|f(x+iR)\cosh\left(\frac{\pi}{2b}(x+iR)\right)\right| = \frac{e^{-tR}\sinh(\frac{\pi}{2b}x)}{\cosh(\frac{\pi}{b}x)-\cos(\frac{\pi a}{b})} \le Ce^{-tR}$$

for all $x\in\mathbb{R}$ with some constant $C>0$, and hence

$$\int_{-R}^R |g(x+iR)|\,dx \le 2CRe^{-tR} \longrightarrow 0 \quad \text{as } R=(2n-1)b\to\infty. \tag{A.5.6}$$

Combining (A.5.4)–(A.5.6) implies that

$$\int_{-\infty}^{\infty} g(x)\,dx = b\,\frac{e^{at}+e^{-at}}{\cos(\frac{\pi a}{2b})}\sum_{k=1}^{\infty}(-1)^{n-1}e^{-(2k-1)bt}$$
$$= \frac{b}{\cos(\frac{\pi a}{2b})}\cdot\frac{e^{at}+e^{-at}}{e^{bt}+e^{-bt}} = \frac{b}{\cos(\frac{\pi a}{2b})}\cdot\frac{\cosh(at)}{\cosh(bt)},$$

which is nothing but (5.1.7). □

Proof of (5.1.8). Considering the inverse Fourier transform of (5.1.8) and exchanging $-t$ with t, we may prove that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ist}\,\frac{t}{\sinh(\frac{t}{2})}\,dt = \frac{\pi}{\cosh^2(\pi s)}. \tag{A.5.7}$$

Define

$$h(z) := \frac{ze^{isz}}{\sinh(\frac{z}{2})}, \quad \text{where } h(0) = 2.$$

The poles of $h(z)$ are $i2n\pi$, $n \in \mathbb{Z}$. The residue of $h(z)$ at $i2n\pi$ is

$$\mathrm{Res}(i2n\pi; h) = \frac{i2n\pi e^{-2n\pi s}}{\frac{d}{dz}\sinh(\frac{z}{2})|_{z=i2n\pi}} = i(-1)^n 4n\pi e^{-2n\pi s}.$$

For the same contour Γ_R as above with $R = (2n+1)\pi$, the residue theorem implies that

$$\int_{-R}^{R} h(x)\,dx + \int_0^R h(R+iy)i\,dy - \int_{-R}^{R} h(x+iR)\,dx - \int_0^R h(-R+iy)i\,dy$$
$$= 2\pi i\sum_{k=1}^{n} i(-1)^k 4k\pi e^{-2k\pi s} = 4\pi\sum_{k=1}^{n}(-1)^{k-1}2k\pi e^{-2k\pi s}.$$

As in the proofs of (5.1.6) and (5.1.7), it is easy to verify that all the integrals but the first in the above left-hand side converge to 0 as $R = (2n+1)\pi \to \infty$. Hence we obtain

$$\int_{-\infty}^{\infty} h(x)\,dx = 4\pi\sum_{k=1}^{\infty}(-1)^{k-1}2k\pi e^{-2k\pi s} = 4\pi\frac{d}{ds}\left\{\sum_{k=1}^{\infty}(-1)^k e^{-2nk\pi s}\right\}$$
$$= 4\pi\frac{d}{ds}\left(\frac{-e^{-\pi s}}{e^{\pi s}+e^{-\pi s}}\right) = \frac{2\pi^2}{\cosh^2(\pi s)},$$

which is nothing but (A.5.7). □

REFERENCES

[1] Abraham, R., and Robbin, J., *Transversal Mappings and Flows*, Benjamin, New York (1967).
[2] Anderson, W. N., and Duffin, R. J., "Series and parallel addition of matrices," *J. Math. Anal. Appl.*, **26**: 576–594 (1969).
[3] Ando, T., *Topics on Operator Inequalities*, Lecture notes (mimeographed), Hokkaido Univ., Sapporo (1978).
[4] Ando, T., "Comparison of norms $|||f(A) - f(B)|||$ and $|||f(|A - B|)|||$," *Math. Z.*, **197**: 403–409 (1988).
[5] Ando, T., "Majorization, doubly stochastic matrices and comparison of eigenvalues," *Linear Algebra Appl.*, **118**: 163–248 (1989).
[6] Ando, T., "Majorization and inequalities in matrix theory," *Linear Algebra Appl.*, **199**: 17–67 (1994).
[7] Ando, T., *Operator-Theoretic Methods for Matrix Inequalities*, Survey notes, Hokusei Gakuen Univ., Sapporo (1998).
[8] Ando, T., and Hiai, F., "Log majorization and complementary Golden–Thompson type inequalities," *Linear Algebra Appl.*, **197/198**: 113–131 (1994).
[9] Ando, T., and Zhan, X., "Norm inequalities related to operator monotone functions," *Math. Ann.*, **315**: 771–780 (1999).
[10] Araki, H., "On an inequality of Lieb and Thirring," *Lett. Math. Phys.*, **19**: 167–170 (1990).
[11] Audenaert, K. M. R., and Aujla, J. S., On Ando's inequalities for convex and concave functions, Preprint (2007), arXiv:0704.0099.
[12] Bendat, J., and Sherman, S., "Monotone and convex operator functions," *Trans. Amer. Math. Soc.*, **79**: 58–71 (1955).
[13] Bhatia, R., *Matrix Analysis*, Springer-Verlag, New York (1996).
[14] Bhatia, R., *Positive Definite Matrices*, Princeton Univ. Press, Princeton (2007).
[15] Bhatia, R., "Linear algebra to quantum cohomology: The story of Alfred Horn's inequalities," *Amer. Math. Monthly*, **108**: 289–318 (2001).
[16] Bhatia, R., and Davis, C., "More matrix forms of the arithmetic-geometric mean inequality," *SIAM J. Matrix Anal. Appl.*, **14**: 132–136 (1993).
[17] Bhatia, R., and Davis, C., "A Cauchy–Schwarz inequality for operators with applications," *Linear Algebra Appl.*, **223/224**: 119–129 (1995).
[18] Bhatia, R., and Kittaneh, F., "Norm inequalities for positive operators," *Lett. Math. Phys.*, **43**: 225–231 (1998).

[19] Bhatia, R., and Parthasarathy, K. R., “Positive definite functions and operator inequalities,” *Bull. London Math. Soc.*, **32**: 214–228 (2000).

[20] Bhatia, R., and Sano, T., “Loewner matrices and operator convexity,” *Math. Ann.*, **344**: 703–716 (2009).

[21] Bhatia, R., and Sano, T., “Positivity and conditional positivity of Loewner matrices,” *Positivity*, to appear, 2009.

[22] Birkhoff, G., “Tres observaciones sobre el algebra lineal,” *Univ. Nac. Tucuman Rev. Ser. A*, **5**: 147–151 (1946).

[23] Birman, M. Sh., Koplienko, L. S., and Solomyak, M. Z., “Estimates for the spectrum of the difference between fractional powers of two self-adjoint operators,” *Soviet Math. (Iz. VUZ)*, **19**: 1–6 (1975).

[24] Bourin, J.-C., “A matrix subadditivity inequality for symmetric norms,” *Proc. Amer. Math. Soc.*, **138**: 495–504 (2010).

[25] Bourin, J.-C., and Uchiyama, M., “A matrix subadditivity inequality for $f(A+B)$ and $f(A)+f(B)$,” *Linear Algebra Appl.*, **423**: 512–518 (2007).

[26] Choi, M.-D., “A Schwarz inequality for positive linear maps on C^*-algebras,” *Illinois J. Math.*, **18**: 565–574 (1974).

[27] Conway, J. B., *Functions of One Complex Variable I*, Second edition, Springer-Verlag, New York-Berlin (1978).

[28] Daleckii, Ju. L., and Krein, S. G., “Integration and differentiation of functions of Hermitian operators and applications to the theory of perturbations,” *Amer. Math. Soc. Transl., Ser. 2*, **47**: 1–30 (1965).

[29] Davis, C., Notions generalizing convexity for functions defined on spaces of matrices, in *Convexity*, Proceedings of Symposia in Pure Mathematics, Vol. 7, Amer. Math. Soc., Providence, R.I., 1963, pp. 187–201.

[30] Donoghue, W. F., Jr., *Monotone Matrix Functions and Analytic Continuation*, Springer-Verlag, Berlin-Heidelberg-New York (1974).

[31] Drissi, D., “Sharp inequalities for some operator means,” *SIAM J. Matrix Anal. Appl.*, **28**: 822–828 (2006).

[32] Fulton, W., “Eigenvalues, invariant factors, highest weights, and Schubert calculus,” *Bull. Amer. Math. Soc.*, **37**: 209–249 (2000).

[33] Furuta, T., “Simple proof of the concavity of operator entropy $f(A) = -A\log A$,” *Math. Ineq. Appl.*, **3**: 305–306 (2000).

[34] Golden, S., “Lower bounds for Helmholtz function,” *Phys. Rev.*, **137**: B1127–B1128 (1965).

[35] Hansen, F., and Pedersen, G. K., “Jensen’s inequality for operators and Löwner’s theorem,” *Math. Ann.*, **258**: 229–241 (1982).

[36] Heinz, E., “Beiträge zur Störungstheorie der Spektralzerlegung,” *Math. Ann.*, **123**: 415–438 (1951).

[37] Hiai, F., Log-majorizations and norm inequalities for exponential operators, in *Linear Operators*, Janas, J., Szafraniec, F. H., and Zemánek, J., (eds.), Banach Center Publications, Vol. 38, 1997, pp. 119–181.

[38] Hiai, F., and Kosaki, H., “Comparison of various means for operators,” *J. Funct. Anal.*, **163**: 300–323 (1999).

[39] Hiai, F., and Kosaki, H., “Means for matrices and comparison of their norms,” *Indiana Univ. Math. J.*, **48**: 899–936 (1999).

[40] Hiai, F., and Kosaki, H., *Means of Hilbert Space Operators*, Lecture Notes in Math. **1820**, Springer-Verlag (2003).

[41] Hiai, F., and Nakamura, Y., “Majorization for generalized s-numbers in semifinite von Neumann algebras,” *Math. Z.*, **195**: 17–27 (1987).

[42] Hiai, F., and Petz, D., “The Golden–Thompson trace inequality is complemented,” *Linear Algebra Appl.*, **181**: 153–185 (1993).

[43] Hiai, F., and Petz, D., “Riemannian metrics on positive definite matrices related to means,” *Linear Algebra Appl.*, **430**: 3105–3130 (2009).

[44] Hiai, F., and Sano, T., Loewner matrices of matrix convex and monotone functions, Preprint (2010), arXiv:1007.2478.

[45] Horn, R. A., “Schlicht mappings and infinitely divisible kernels,” *Pacific J. Math.*, **38**: 423–430 (1971).

[46] Horn, R. A., “Norm bounds for Hadamard products and the arithmetic-geometric mean inequality for unitarily invariant norms,” *Linear Algebra Appl.*, **223/224**: 355–361 (1995).

[47] Horn, A., “Eigenvalues of sums of Hemitian matrices,” *Pacific J. Math.*, **12**: 225–241 (1962).

[48] Kato, T., Trotter’s product formula for an arbitrary pair of self-adjoint contraction semigroups, in *Topics in Functional Analysis*, Gohberg, I., and Kac, M., (eds.), Ad. Math. Suppl. Studies 3, Academic Press, New York (1978) 185–195.

[49] Klyachko, A. A., “Stable bundles, representation theory and Hermitian operators,” *Selecta Math.*, **4**: 419–445 (1998).

[50] Knuston, A., and Tao, T., “The honeycomb model of $GL_n(\mathbb{C})$ tensor products I: Proof of the saturation conjecture,” *J. Amer. Math. Soc.*, **12**: 1055–1090 (1999).

[51] Korányi, A., “On a theorem of Löwner and its connection with resolvents of transformations,” *Acta Sci. Math. (Szeged)*, **17**: 63–70 (1956).

[52] Kosaki, H., “Arithmetic-geometric mean and related inequalities for operators,” *J. Funct. Anal.*, **156**: 429–451 (1998).

[53] Kosaki, H., Strong monotonicity for various means, Preprint (2009).

[54] Kosaki, H., Positive definiteness of functions with applications to operator norm inequalities, III, Preprint (2010).

[55] Kosaki, H., *Positive Definiteness of Functions with Applications to Operator Norm Inequalities*, *Memoirs Amer. Math. Soc.*, Providence, R.I., to appear.

[56] Kosem, T., Inequalities between $\|f(A+B)\|$ and $\|f(A)+f(B)\|$, *Linear Algebra Appl.*, **418**: 153–160 (2006).

[57] Kraus, F., Über konvexe Matrixfunktionen, *Math. Z.*, **41**: 18–42 (1936); also English version translated by T. Ando.

[58] Kubo, F., and Ando, T., Means of positive linear operators, *Math. Ann.*, **246**: 205–224 (1980).

[59] Lenard, A., “Generalization of the Golden–Thompson inequality $\mathrm{Tr}(e^A e^B) \geq \mathrm{Tr}\, e^{A+B}$,” *Indiana Univ. Math. J.*, **21**: 457–467 (1971).

[60] Li, C.-K., and Mathias, R., “The LIdskii–Mirsky–Wielandt theorem—additive and multiplicative versions,” *Numer. Math.*, **81**: 377–413 (1999).

[61] Lin, T.-P., “The power mean and the logarithmic mean,” *Amer. Math. Monthly*, **81**: 879–883 (1974).

[62] Löwner, K., “Über monotone Matrixfunctionen,” *Math. Z.*, **38**: 177–216 (1934).

[63] Marshall, A. W., and Olkin, I., *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York (1979).

[64] Mathias, R., “An arithmetic-geometric-harmonic mean inequality involving Hadamard products,” *Linear Algebra Appl.*, **184**: 71–78 (1993).

[65] Petz, D., "Monotone metrics on matrix spaces," *Linear Algebra Appl.*, **244**: 81–96 (1996).
[66] Petz, D., *Quantum Information Theory and Quantum Statistics*, Springer-Verlag, Berlin-Heidelberg (2008).
[67] Phelps, R. R., *Lectures on Choquet's Theorem*, Second edition, Lecture Notes in Math. **1757**, Springer-Verlag (2001).
[68] Powers, R. T., and Størmer, E., "Free states of the canonical anticommutation relations," *Comm. Math. Phys.*, **16**: 1–33 (1970).
[69] Pusz, W., and Woronowicz, S. L., "Functional calculus for sesquilinear forms and the purification map," *Rep. Math. Phys.*, **8**: 159–170 (1975).
[70] Reed, M., and Simon, B., *Methods of Modern Mathematical Physics II*, Academic Press, New York (1975).
[71] Suzuki, M., "Quantum statistical Monte Carlo methods and applications to spin systems," *J. Stat. Phys.*, **43**: 883–909 (1986).
[72] Symanzik, K., "Proof and refinements of an inequality of Feynman," *J. Math. Phys.*, **6**: 1155–1156 (1965).
[73] Thompson, C. J., "Inequality with applications in statistical mechanics," *J. Math. Phys.*, **6**: 1812–1813 (1965).
[74] Thompson, C. J., "Inequalities and partial orders on matrix spaces," *Indiana Univ. Math. J.*, **21**: 469–480 (1971).
[75] Uchiyama, M., "Subadditivity of eigenvalue sums," *Proc. Amer. Math. Soc.*, **134**: 1405–1412 (2006).
[76] Uchiyama, M., "Operator monotone functions, positive definite kernels and majorization," *Proc. Amer. Math. Soc.*, to appear.
[77] von Neumann, J., "Some matrix-inequalities and metrization of matrix-space," *Tomsk Univ. Rev.*, **1**: 286–299 (1937) (Collected Works, Vol. IV, pp. 205–218).
[78] Wang, B.-Y., and Gong, M.-P., "Some eigenvalue inequalities for positive semidefinite matrix power products," *Linear Algebra Appl.*, **184**: 249–260 (1993).
[79] Wielandt, H., "An extremum property of sums of eigenvalues," *Proc. Amer. Math. Soc.*, **6**: 106–110 (1955).
[80] Zhan, X., "Inequalities for unitarily invariant norms," *SIAM J. Matrix Anal. Appl.*, **20**: 466–470 (1998).

Index