LOG-MAJORIZATIONS AND NORM INEQUALITIES
FOR EXPONENTIAL OPERATORS

FUMIO HIAI

Department of Mathematics, Ibaraki University
Mito, Ibaraki 310, Japan
E-mail: f37551@sinet.ad.jp

Abstract. Concise but self-contained reviews are given on theories of majorization and symmetrically normed ideals, including the proofs of the Lidskii–Wielandt and the Gelfand–Naimark theorems. Based on these reviews, we discuss logarithmic majorizations and norm inequalities of Golden–Thompson type and its complementary type for exponential operators on a Hilbert space. Furthermore, we obtain norm convergences for the exponential product formula as well as for that involving operator means.

1. Introduction. Since the notion of majorization was introduced by Hardy, Little-wood, and Pólya, it has been discussed by many mathematicians in various circumstances with various applications. First let us recall the notion of (weak) majorization in the simplest case of real vectors. For real vectors $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$, the weak majorization $a \prec_w b$ means that $\sum_{i=1}^{k} a_{[i]} \leq \sum_{i=1}^{k} b_{[i]}$ holds for $1 \leq k \leq n$, where $(a_{[1]}, \ldots, a_{[n]})$ is the decreasing rearrangement of a. The majorization $a \prec b$ means that $a \prec_w b$ and $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$. When a and b are nonnegative, the multiplicative or logarithmic (weak) majorization can be also defined by taking product $\prod$ in place of sum $\sum$ in the above, which we referred to in [6] as the log-majorization.

Several (weak) majorizations are known for the eigenvalues and the singular values of matrices and compact operators, as was fully clarified in Marshall and Olkin's monograph [62] and also in [4, 61]. These majorizations give rise to powerful devices in deriving various norm inequalities (in particular, perturbation norm inequalities) as well as trace or determinant inequalities for matrices or operators (see e.g. [15]). Among other things, the Lidskii–Wielandt majorization theorem is especially famous and important. A crucial reason why the (weak) majorization is useful in operator norm inequalities is the following fact: For bounded Hilbert space operators A and B, the weak majorization $\mu(A) \prec_w$

$\mu(B)$ holds if and only if $\|A\| \leq \|B\|$ for any symmetric (or unitarily invariant) norm $\|\cdot\|$, where $\mu(A) = (\mu_1(A), \mu_2(A), \ldots)$ are the (generalized) singular values of A with multiplicities. General theory of symmetric norms and symmetrically normed ideals was extensively developed in Gohberg and Krein's monumental monograph [32] (also [77, 79]). The majorization technique sometimes plays an important role in the study of symmetrically normed ideals. For instance, the Hölder type inequality for symmetric norms is a simple consequence of Horn's majorization of multiplicative type.

The celebrated Golden–Thompson trace inequality, independently proved by Golden [33], Symanzik [83], and Thompson [84], is $\operatorname{tr} e^{H+K} \leq \operatorname{tr} e^H e^K$ for self-adjoint operators (particularly for Hermitian matrices) H and K. So far, there have been many extended or related inequalities around the Golden–Thompson inequality. For example, when H and K are Hermitian matrices, this inequality was extended in [58, 85] to the weak majorization $\mu(e^{H+K}) \prec_w \mu(e^{H/2} e^K e^{H/2})$ or equivalently $\|e^{H+K}\| \leq \|e^{H/2} e^K e^{H/2}\|$ for any unitarily invariant norm, and in [81] to $\operatorname{tr} e^{H+K} \leq \operatorname{tr}(e^{H/n} e^{K/n})^n$ for all $n \in \mathbb{N}$. Also, the Araki–Lieb–Thirring inequality [12] (also [88]) is regarded as a strengthened Golden–Thompson inequality and is reformulated in terms of log-majorization (see [6, 40]). On the other hand, a complementary counterpart of the Golden–Thompson trace inequality was discovered in [44] in the course of study on lower and upper bounds for the relative entropy, and was strengthened in [6] to the form of log-majorization by using the technique of antisymmetric tensor powers. Restricted to the matrix case, the above log-majorizations of Golden–Thompson type and its complementary type yield the following norm inequalities: If H and K are Hermitian matrices and $0 < \alpha < 1$, then

$$\|(e^{rH/(1-\alpha)} \#_\alpha e^{rK/\alpha})^{1/r}\| \leq \|e^{H+K}\| \leq \|(e^{rH/2} e^{rK} e^{rH/2})^{1/r}\|, \quad r > 0,$$

for any unitarily invariant norm $\|\cdot\|$, where $\#_\alpha$ denotes the α-power mean, i.e. the operator mean corresponding to the operator monotone function x^α. Moreover, the above left-hand (resp. right-hand) side increases (resp. decreases) to $\|e^{H+K}\|$ as $r \downarrow 0$.

The Golden–Thompson trace inequality was originally motivated by quantum statistical mechanics. When a Hamitonian K is given as a self-adjoint operator (assumed here to be lower-bounded) on a Hilbert space, the partition function $\operatorname{tr} e^{-\beta K}$ and the free energy $\log \operatorname{tr} e^{-\beta K}$ where β is an inverse temperature constant are basically important from the quantum statistical mechanical viewpoint. When K receives a lower-bounded perturbation by H, physicists sometimes approximate $\operatorname{tr} e^{-\beta(H+K)}$ by $\operatorname{tr}(e^{-\beta H/n} e^{-\beta K/n})^n$ via the Trotter product formula. Although the convergence $\operatorname{tr}(e^{-\beta H/n} e^{-\beta K/n})^n \to \operatorname{tr} e^{-\beta(H+K)}$ might have been strongly believed by physicists, there was no rigorous proof up to [40]. Indeed, it was more strongly proved in [40] that if H and K are lower-bounded self-adjoint operators such that e^{-K} is of trace class and $H + K$ is essentially self-adjoint, then the following trace norm convergence holds:

$$\lim_{r \downarrow 0} \|(e^{-rH/2} e^{-rK} e^{-rH/2})^{1/r} - e^{-(H+K)}\|_1 = 0.$$

(Another $\|\cdot\|_1$-convergence under a rather strong assumption was given in [65]. Other recent developments on the Trotter–Kato product formula in the operator norm and trace norm are found in [37, 46, 47, 75].)

The present paper enjoys both aspects of a review paper and of a research paper. It is organized in five sections which are divided into several subsections. Our main aim is to present log-majorizations and norm inequalities for infinite-dimensional exponential operators. For this sake, in Sections 1 and 2 we concisely review the majorization theory and theory of symmetrically normed ideals. Although several distinguished monographs, as cited above, are available, we intend to make the exposition completely self-contained, so that Sections 1 and 2 may independently serve as a concise text on these subjects. The main part of Section 1 is the proofs of the Lidskii–Wielandt and the Gelfand–Naimark theorems for (generalized) singular values of matrices (operators). Our proofs are rather new and based on the real interpolation method (or the K-functional method). In Section 2 we stress the majorization technique in the theory of symmetrically normed ideals.

Section 3 is taken from [40] and is not new, but we sometimes give more detailed accounts for the convenience of the reader. We investigate log-majorizations and norm inequalities of Golden–Thompson type for exponential operators. For instance, it is shown that if H and K are lower-bounded self-adjoint operators, then

$$\|e^{-(H\hat{+}K)}\| \leq \|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|, \quad r > 0,$$

for any symmetric norm $\|\cdot\|$, where $H\hat{+}K$ is the form sum of H and K. Preliminaries on antisymmetric tensor powers and the Trotter–Kato exponential product formula are included, which are quite beneficial in proving our results. Also in Subsection 3.5 we discuss the trace norm convergence of exponential product formula together with some technical preliminaries in the framework of von Neumann algebras.

Section 4 is considered as a complementary counterpart of Section 3. Extending the matrix case [6, 44] (also [71]), we investigate log-majorizations and norm inequalities involving operator means (in particular, the α-power mean), which are opposite to those in Section 3 and considered as complementary Golden–Thompson type. If H is bounded self-adjoint and K is lower-bounded self-adjoint, then the following is proved for any symmetric norm $\|\cdot\|$ and $0 < \alpha < 1$:

$$\|(e^{-rH/(1-\alpha)} \#_\alpha e^{-rK/\alpha})^{1/r}\| \leq \|e^{-(H+K)}\|, \quad r > 0.$$

The most important ingredient in the extension from the matrix case to the infinite-dimensional case is an exponential product formula for operator means established in Subsection 4.3. Finally in Section 5 we obtain further log-majorization results, for example, the log-majorization equivalent to the Furuta inequality [28], generalized log-majorizations of Horn's type and of Golden–Thompson type, etc. Some determinant inequalities are also included. Most results of Section 4 and many of Section 5 are new.

Although we confine ourselves to the setting of Hilbert space operators (in other words, the setup of $\mathcal{B}(\mathcal{H})$) in this paper, it should be mentioned that many subjects treated here extend to the von Neumann algebra setup. In fact, based on the noncommutative integration theory ([20, 66, 78]), we can discuss the majorization theory in semifinite von Neumann algebras (see e.g. [21, 38, 39, 41, 42, 43, 48, 49, 64]) by using the notion of generalized s-numbers introduced in [25, 26] for measurable operators. Noncommutative Banach function spaces (i.e. generalized symmetrically normed ideals) associated with semifinite von Neumann algebras have been discussed in [21, 22, 53, 89, 90] for instance.

In particular, theory of noncommutative L^p-spaces over arbitrary von Neumann algebras was developed (e.g. the Haagerup L^p-spaces [35]). Kosaki [55] extended the Araki–Lieb–Thirring inequality (i.e. a log-majorization result in Subsection 3.2) to the von Neumann algebra case. Furthermore, many authors have worked on the Golden–Thompson inequality in von Neumann algebras in several ways ([8, 25, 45, 70, 76]). However, at present, the von Neumann algebra versions for the norm convergence in Subsection 3.5 and for the study of complementary Golden–Thompson type in Section 4 are not yet investigated. When we want to extend the study of Section 4 to the von Neumann algebra setup, the antisymmetric tensor technique is no longer available, so that we would have to exploit a new method.

The contents of the paper are as follows:

1. Majorization and log-majorization. The purpose of this section is to give a concise but self-contained review on the majorization theory for (generalized) singular values of matrices and operators. Complete expositions on the subject are found in [4, 62]. Also see [5] for recent developments.

1.1. *Majorization for vectors.* Let us start with the majorization for real vectors, which was introduced by Hardy, Littlewood, and Pólya. For two vectors $a = (a_1, \ldots, a_n)$

and $b = (b_1, \ldots, b_n)$ in $\mathbb{R}^n$, the *weak majorization* $a \prec_w b$ means that

$$(1.1) \qquad \sum_{i=1}^{k} a_{[i]} \leq \sum_{i=1}^{k} b_{[i]}, \qquad 1 \leq k \leq n,$$

where $(a_{[1]}, \ldots, a_{[n]})$ is the *decreasing rearrangement* of a, i.e. $a_{[1]} \geq \ldots \geq a_{[n]}$ are the components of a in decreasing order. The *majorization* $a \prec b$ means that $a \prec_w b$ and equality holds for $k = n$ in (1.1). The following characterizations of majorization and weak majorization are fundamental.

PROPOSITION 1.1. *The following conditions for $a, b \in \mathbb{R}^n$ are equivalent:*

(i) $a \prec b$;

(ii) $\sum_{i=1}^{n} |a_i - r| \leq \sum_{i=1}^{n} |b_i - r|$ *for all* $r \in \mathbb{R}$;

(iii) $\sum_{i=1}^{n} f(a_i) \leq \sum_{i=1}^{n} f(b_i)$ *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. $D = [d_{ij}]$ with $d_{ij} \geq 0$, $\sum_{j=1}^{n} d_{ij} = 1$ for $1 \leq i \leq n$, and $\sum_{i=1}^{n} d_{ij} = 1$ for $1 \leq j \leq n$.*

P r o o f. (i)$\Rightarrow$(iv). We show that there exist a finite number of matrices $D_1, \ldots, D_N$ of the form $\lambda I + (1 - \lambda)\Pi$ where $0 \leq \lambda \leq 1$ and Π is a permutation matrix interchanging two coordinates only such that $a = D_N \ldots D_1 b$. Then (iv) follows because $D_N \ldots D_1$ becomes a convex combination of permutation matrices. We may assume that $a_1 \geq \ldots \geq a_n$ and $b_1 \geq \ldots \geq b_n$. Suppose $a \neq b$ and choose the largest j such that $a_j < b_j$. Then there exists k with $k > j$ such that $a_k > b_k$. Choose the smallest such k. Let $1 - \lambda_1 = \min\{b_j - a_j, a_k - b_k\}/(b_j - b_k)$ and Π_1 be the permutation matrix interchanging the jth and kth coordinates. Then $0 < \lambda_1 < 1$ because $b_j > a_j \geq a_k > b_k$. 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 $b_1^{(1)} \geq \ldots \geq b_n^{(1)}$. Moreover the jth or kth coordinates of a and $b^{(1)}$ are equal. When $a \neq b^{(1)}$, we can apply the above argument to a and $b^{(1)}$. Repeating 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 get

$$\sum_{i=1}^{n} |a_i - r| = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} d_{ij}(b_j - r) \right| \leq \sum_{i,j=1}^{n} d_{ij} |b_j - r| = \sum_{j=1}^{n} |b_j - r|.$$

(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 get

$$(1.2) \qquad \sum_{i=1}^{n} (a_i - r)_+ \leq \sum_{i=1}^{n} (b_i - r)_+, \qquad r \in \mathbb{R}.$$

Now prove that (1.2) implies $a \prec_w b$. When $b_{[k]} \geq r \geq b_{[k+1]}$, $\sum_{i=1}^{k} a_{[i]} \leq \sum_{i-1}^{k} b_{[i]}$ follows

because

$$\sum_{i=1}^{n}(a_i - r)_+ \geq \sum_{i=1}^{k}(a_{[i]} - r)_+ \geq \sum_{i=1}^{k} a_{[i]} - kr,$$

$$\sum_{i=1}^{n}(b_i - r)_+ = \sum_{i=1}^{k} b_{[i]} - kr.$$

(iv)$\Rightarrow$(iii). Suppose that $a_i = \sum_{k=1}^{N} \lambda_k b_{\pi_k(i)}$, $1 \leq i \leq n$, where $\lambda_k > 0$, $\sum_{k=1}^{N} \lambda_k = 1$, and π_k are permutations on $\{1, \ldots, n\}$. Then the convexity of f implies that

$$\sum_{i=1}^{n} f(a_i) \leq \sum_{i=1}^{n}\sum_{k=1}^{N} \lambda_k f(b_{\pi_k(i)}) = \sum_{i=1}^{n} f(b_i).$$

(iii)$\Rightarrow$(v) is trivial because $f(x) = |x - r|$ is convex. ∎

Note that (v)$\Rightarrow$(iv) is seen directly from the well-known theorem of Birkhoff [17] saying that every doubly stochastic matrix is a convex combination of permutation matrices.

PROPOSITION 1.2. *The following conditions* (i)–(iv) *for* $a, b \in \mathbb{R}^n$ *are equivalent*:

(i) $a \prec_w b$;

(ii) *there exists* $c \in \mathbb{R}^n$ *such that* $a \leq c \prec b$, *where* $a \leq c$ *means that* $a_i \leq c_i$, $1 \leq i \leq n$;

(iii) $\sum_{i=1}^{n}(a_i - r)_+ \leq \sum_{i=1}^{n}(b_i - r)_+$ *for all* $r \in \mathbb{R}$;

(iv) $\sum_{i=1}^{n} f(a_i) \leq \sum_{i=1}^{n} f(b_i)$ *for any increasing convex function* f *on an interval containing all* a_i, b_i.

Moreover, if $a, b \geq 0$, *then the above conditions are equivalent to the following*:

(v) $a = Sb$ *for some doubly substochastic* $n \times n$ *matrix* S, *i.e.* $S = [s_{ij}]$ *with* $s_{ij} \geq 0$, $\sum_{j=1}^{n} s_{ij} \leq 1$ *for* $1 \leq i \leq n$, *and* $\sum_{i=1}^{n} s_{ij} \leq 1$ *for* $1 \leq j \leq n$.

P r o o f. (i)$\Rightarrow$(ii). By induction on n. We may assume that $a_1 \geq \ldots \geq a_n$ and $b_1 \geq \ldots \geq b_n$. Let $\alpha = \min_{1 \leq k \leq n}(\sum_{i=1}^{k} b_i - \sum_{i=1}^{k} a_i)$ and define $\tilde{a} = (a_1 + \alpha, a_2, \ldots, a_n)$. Then $a \leq \tilde{a} \prec_w b$ and $\sum_{i=1}^{k} \tilde{a}_i = \sum_{i=1}^{k} b_i$ for some $1 \leq k \leq n$. When $k = n$, $a \leq \tilde{a} \prec b$. When $k < n$, we get $(\tilde{a}_1, \ldots, \tilde{a}_k) \prec (b_1, \ldots, b_k)$ and $(\tilde{a}_{k+1}, \ldots, \tilde{a}_n) \prec_w (b_{k+1}, \ldots, b_n)$. Hence the induction assumption implies that $(\tilde{a}_{k+1}, \ldots, \tilde{a}_n) \leq (c_{k+1}, \ldots, c_n) \prec (b_{k+1}, \ldots, b_n)$ for some $(c_{k+1}, \ldots, c_n) \in \mathbb{R}^{n-k}$. Then $a \leq (\tilde{a}_1, \ldots, \tilde{a}_k, c_{k+1}, \ldots, c_n) \prec b$ is immediate from $\tilde{a}_k \geq b_k \geq b_{k+1} \geq c_{k+1}$.

(ii)$\Rightarrow$(iv). Let $a \leq c \prec b$. If f is increasing and convex on an interval $[\alpha, \beta]$ containing a_i, b_i, then $c_i \in [\alpha, \beta]$ and

$$\sum_{i=1}^{n} f(a_i) \leq \sum_{i=1}^{n} f(c_i) \leq \sum_{i=1}^{n} f(b_i)$$

by Proposition 1.1.

(iv)$\Rightarrow$(iii) is trivial and (iii)$\Rightarrow$(i) was already shown in the proof (ii)$\Rightarrow$(i) of Proposition 1.1.

Now assume $a, b \geq 0$ and prove (ii)$\Leftrightarrow$(v). If $a \leq c \prec b$, then we have, by Proposition 1.1, $c = Db$ for some doubly stochastic matrix D and $a_i = \alpha_i c_i$ for some $0 \leq \alpha_i \leq 1$.

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 and hence $a \le Db \prec b$. ∎

Let $a, b \in \mathbb{R}^n$ and $a, b \ge 0$. We define the *weak log-majorization* $a \prec_{w(\log)} b$ when

$$(1.3) \qquad \prod_{i=1}^{k} a_{[i]} \le \prod_{i=1}^{k} b_{[i]}, \qquad 1 \le k \le n,$$

and the *log-majorization* $a \prec_{(\log)} b$ when $a \prec_{w(\log)} b$ and equality holds for $k = n$ in (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 $\log a = (\log a_1, \ldots, \log a_n)$.

The notions of weak majorization and weak log-majorization are similarly defined for positive bounded infinite sequences. But, to avoid discussing the decreasing rearrangement of an infinite sequence (this is considered as a special case of generalized singular values of a bounded operator introduced later in this section), we here confine ourselves to infinite sequences $a = (a_1, a_2, \ldots)$ and $b = (b_1, b_2, \ldots)$ such that $a_1 \ge a_2 \ge \ldots \ge 0$ and $b_1 \ge b_2 \ge \ldots \ge 0$. For such a, b we define $a \prec_w b$ and $a \prec_{w(\log)} b$ when $\sum_{i=1}^{k} a_i \le \sum_{i=1}^{k} b_i$ and $\prod_{i=1}^{k} a_i \le \prod_{i=1}^{k} b_i$, respectively, for all $k \in \mathbb{N}$.

PROPOSITION 1.3. *Let $a, b \in \mathbb{R}^n$ with $a, b \ge 0$ and suppose $a \prec_{w(\log)} b$. If f is a continuous increasing 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 $a \prec_w b$. Moreover, the same assertions hold also for infinite sequences a, b with $a_1 \ge a_2 \ge \ldots \ge 0$ and $b_1 \ge b_2 \ge \ldots \ge 0$, whenever $f(0) \ge 0$ is additionally assumed.*

Proof. First assume that $a, b \in \mathbb{R}^n$ are strictly positive and $a \prec_{w(\log)} b$, so that $\log a \prec_w \log b$. Since $g \circ h$ is convex when g and h are convex with g increasing, the function $(f(e^x) - r)_+$ is increasing and convex for any $r \in \mathbb{R}$. Hence we get, by Proposition 1.2,

$$\sum_{i=1}^{n} (f(a_i) - r)_+ \le \sum_{i=1}^{n} (f(b_i) - r)_+,$$

which implies $f(a) \prec_w f(b)$ by Proposition 1.2 again. 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 $f(a) \prec_w f(b)$.

The case of infinite sequences is immediate from the above case. In fact, $a \prec_{w(\log)} b$ implies that $(a_1, \ldots, a_n) \prec_{w(\log)} (b_1, \ldots, b_n)$ for every $n \in \mathbb{N}$. Hence $(f(a_1), \ldots, f(a_n)) \prec_w (f(b_1), \ldots, f(b_n))$, $n \in \mathbb{N}$, so that $f(a) \prec_w f(b)$. ∎

1.2. *Generalized singular values.* In the sequel of this section, we discuss the majorization theory for singular values of matrices and operators. Our goal is to prove the Lidskii–Wielandt and the Gelfand–Naimark theorems for generalized singular values of bounded operators. Indeed, these theorems were proved by using the real interpolation method in the setting of von Neumann algebras in [41, 64] (also [21]). In Subsection 1.3 we first prove the theorems for matrices by using this new method. After that, in Subsection 1.4 we extend them from matrices to operators in a rather simple way.

For any $n \times n$ matrix A let $\mu(A) = (\mu_1(A), \ldots, \mu_n(A))$ be the vector of singular values of A in decreasing order, i.e. $\mu_1(A) \geq \ldots \geq \mu_n(A)$ are the eigenvalues of $|A| = (A^*A)^{1/2}$ with multiplicities. When A is Hermitian, the vector of eigenvalues of A in decreasing order is denoted by $\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A))$. The notion of singular values is generalized to infinite-dimensional operators. Let $\mathcal{H}$ be a Hilbert space (always assumed to be separable) and $\mathcal{B}(\mathcal{H})$ the algebra of all bounded operators on $\mathcal{H}$. For any $A \in \mathcal{B}(\mathcal{H})$ we define the *generalized singular values* $\mu_1(A) \geq \mu_2(A) \geq \ldots$ of A by

$$\mu_n(A) = \inf\{\lambda \geq 0 : \operatorname{rank}(I - E_{|A|}(\lambda)) < n\}, \quad n \in \mathbb{N},$$

where $|A| = \int_0^\infty \lambda dE_{|A|}(\lambda)$ is the spectral decomposition of $|A|$ so that $I - E_{|A|}(\lambda)$ is the spectral projection of $|A|$ corresponding to the interval (λ, ∞). The above definition of $\mu_n(A)$ is a special case of the generalized s-numbers of measurable operators in the setting of von Neumann algebras introduced in [25, 26] (also [69]). If A is a compact operator (in particular, a matrix), then $\mu_n(A)$ are the usual singular values of A in decreasing order with multiplicities.

Let $\mu_\infty(A) = \lim_{n \to \infty} \mu_n(A)$. Then it is easy to see that $\mu_\infty(A) = \|A\|_e$, the essential norm of A; namely $\mu_\infty(A)$ is equal to the largest $\alpha \in \mathbb{R}$ such that $E_{|A|}(\alpha + \varepsilon) - E_{|A|}(\alpha - \varepsilon)$ is of infinite rank for every $\varepsilon > 0$. Note that $\mu_\infty(A) = 0$ if and only if A is compact. Furthermore, if $\mu_n(A) > \mu_\infty(A)$, then $\mu_n(A)$ is an eigenvalue of $|A|$ with finite multiplicity.

The basic properties of $\mu_n(A)$ are summarized as follows. See [26] for the proof in the von Neumann algebra setting.

PROPOSITION 1.4. *Let* $A, B, X, Y \in \mathcal{B}(\mathcal{H})$ *and* $n, m \in \mathbb{N}$.

(1) *Mini-max expression*:

$$(1.4) \qquad \mu_n(A) = \inf\{\|A(I - P)\|_\infty : P \text{ is a projection, } \operatorname{rank} P = n - 1\},$$

where $\|\cdot\|_\infty$ *denotes the operator norm and a projection means always an orthogonal one. Furthermore, if* $A \geq 0$ *then*

$$(1.5) \qquad \mu_n(A) = \inf\{\sup_{\xi \in \mathcal{M}^\perp, \, \|\xi\|=1} \langle A\xi, \xi \rangle : \mathcal{M} \text{ is a subspace of } \mathcal{H}, \dim \mathcal{M} = n - 1\}.$$

(2) *Approximation number expression*:

$$(1.6) \qquad \mu_n(A) = \inf\{\|A - X\|_\infty : X \in \mathcal{B}(\mathcal{H}), \operatorname{rank} X < n\}.$$

(3) $\mu_1(A) = \|A\|_\infty$.
(4) $\mu_n(\alpha A) = |\alpha| \mu_n(A)$ *for* $\alpha \in \mathbb{C}$.
(5) $\mu_n(A) = \mu_n(A^*)$.
(6) *If* $0 \leq A \leq B$ *then* $\mu_n(A) \leq \mu_n(B)$.
(7) $\mu_n(XAY) \leq \|X\|_\infty \|Y\|_\infty \mu_n(A)$.
(8) $\mu_{n+m-1}(A + B) \leq \mu_n(A) + \mu_m(B)$.
(9) $\mu_{n+m-1}(AB) \leq \mu_n(A)\mu_m(B)$.
(10) $|\mu_n(A) - \mu_n(B)| \leq \|A - B\|_\infty$.
(11) $\mu_n(f(A)) = f(\mu_n(A))$ *if* $A \geq 0$ *and* f *is a continuous increasing function on* $[0, \infty)$ *with* $f(0) \geq 0$.

P r o o f. (1) Let α_n be the right-hand side of (1.4). First note that this does not change when rank $P = n - 1$ in (1.4) is replaced by rank $P < n$. So, if rank$(I - E_{|A|}(\lambda)) < n$, then

$$\alpha_n \leq \|AE_{|A|}(\lambda)\|_\infty = \| |A|E_{|A|}(\lambda)\|_\infty \leq \lambda.$$

Hence $\alpha_n \leq \mu_n(A)$. Conversely, for any $\varepsilon > 0$ choose a projection P with rank $P = n-1$ such that $\|A(I - P)\|_\infty < \alpha_n + \varepsilon$. Suppose rank$(I - E_{|A|}(\alpha_n + \varepsilon)) \geq n$. Then there exists $\xi \in \mathcal{H}$ with $\|\xi\| = 1$ such that $(I - E_{|A|}(\alpha_n + \varepsilon))\xi = \xi$ but $P\xi = 0$. This implies that

$$\alpha_n + \varepsilon \leq \| |A|\xi\| = \|A(I - P)\xi\| < \alpha_n + \varepsilon,$$

a contradiction. Hence rank$(I - E_{|A|}(\alpha_n + \varepsilon)) < n$ and $\mu_n(A) \leq \alpha_n + \varepsilon$, implying $\mu_n(A) \leq \alpha_n$.

When $A \geq 0$, since $E_{A^{1/2}}(\lambda) = E_A(\lambda^2)$, we get $\mu_n(A) = \mu_n(A^{1/2})^2$. So (1.5) follows from (1.4), because the right-hand side of (1.5) is written as

$$\inf\{\|A^{1/2}(I - P)\|_\infty^2 : P \text{ is a projection, } \operatorname{rank} P = n - 1\}.$$

(2) Let β_n be the right-hand side of (1.6). If rank$(I - E_{|A|}(\lambda)) < n$, then rank$(A(I - E_{|A|}(\lambda))) < n$ and $\beta_n \leq \|AE_{|A|}(\lambda)\|_\infty \leq \lambda$. Hence $\beta_n \leq \mu_n(A)$. Conversely, if rank $X < n$, then the support projection P of $|X|$ has rank $< n$. Since $X(I - P) = 0$, we get by (1.4)

$$\mu_n(A) \leq \|A(I - P)\|_\infty = \|(A - X)(I - P)\|_\infty \leq \|A - X\|_\infty,$$

implying $\mu_n(A) \leq \beta_n$.

(3) is (1.4) for $n = 1$. (4) and (5) follow from (1.4) and (1.6), respectively. (6) is a consequence of (1.5). It is immediate from (1.4) that $\mu_n(XA) \leq \|X\|_\infty \mu_n(A)$. Also $\mu_n(AY) = \mu_n(Y^*A^*) \leq \|Y\|_\infty \mu_n(A)$ by (5). Hence (7) holds.

Next we show (8)–(10). By (1.6), for every $\varepsilon > 0$, there exist $X, Y \in \mathcal{B}(\mathcal{H})$ with rank $X < n$, rank $Y < m$ such that $\|A - X\|_\infty < \mu_n(A) + \varepsilon$ and $\|B - Y\|_\infty < \mu_m(B) + \varepsilon$. Since rank$(X + Y) < n + m - 1$, we have

$$\mu_{n+m-1}(A + B) \leq \|(A + B) - (X + Y)\|_\infty < \mu_n(A) + \mu_m(B) + 2\varepsilon,$$

implying (8). For $Z = XB + (A - X)Y$ we get

$$\operatorname{rank} Z \leq \operatorname{rank} X + \operatorname{rank} Y < n + m - 1,$$

$$\|AB - Z\|_\infty = \|(A - X)(B - Y)\|_\infty < (\mu_n(A) + \varepsilon)(\mu_m(B) + \varepsilon).$$

These imply (9). Letting $m = 1$ and replacing B by $B - A$ in (8) we get

$$\mu_n(B) \leq \mu_n(A) + \|B - A\|_\infty,$$

which shows (10).

(11) Assume $f(0) = 0$ and let $f_\varepsilon(t) = f(t) + \varepsilon t$ for $\varepsilon > 0$. Then $\mu_n(f_\varepsilon(A)) = f_\varepsilon(\mu_n(A))$ is easily checked from $E_{f_\varepsilon(A)}(\lambda) = E_A(f_\varepsilon^{-1}(\lambda))$. Since by (10)

$$|\mu_n(f_\varepsilon(A)) - \mu_n(f(A))| \leq \|f_\varepsilon(A) - f(A)\|_\infty \to 0$$

as $\varepsilon \downarrow 0$, we have $\mu_n(f(A)) = \lim_{\varepsilon \downarrow 0} f_\varepsilon(\mu_n(A)) = f(\mu_n(A))$. When $\alpha = f(0) > 0$, let $g(t) = f(t) - \alpha$. Then the assertion follows as

$$\mu_n(f(A)) = \mu_n(g(A) + \alpha I) = \mu_n(g(A)) + \alpha = g(\mu_n(A)) + \alpha = f(\mu_n(A)). \quad \blacksquare$$

1.3. *Majorization for matrices*: *Lidskii–Wielandt and Gelfand–Naimark theorems.*
The following majorization results are the celebrated Lidskii–Wielandt theorem for singular values of matrices as well as for eigenvalues of Hermitian matrices.

THEOREM 1.5. *For any $n \times n$ matrices A and B,*

$$|\mu(A) - \mu(B)| \prec_w \mu(A - B),$$

that is,

$$\sum_{j=1}^{k} |\mu_{i_j}(A) - \mu_{i_j}(B)| \leq \sum_{j=1}^{k} \mu_j(A - B)$$

for every $1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

THEOREM 1.6. *For any Hermitian $n \times n$ matrices A and B,*

$$\lambda(A) - \lambda(B) \prec \lambda(A - B),$$

or equivalently

$$(\lambda_i(A) + \lambda_{n-i+1}(B)) \prec \lambda(A + B).$$

The following results due to Ky Fan are consequences of the above theorems. The direct proofs of these weakened majorizations are much easier than those of Theorems 1.5 and 1.6.

COROLLARY 1.7. (1) *For any matrices A and B,*

$$\mu(A + B) \prec_w \mu(A) + \mu(B).$$

(2) *For any Hermitian matrices A and B,*

$$\lambda(A + B) \prec \lambda(A) + \lambda(B).$$

The proof of Theorem 1.5 is based on the real interpolation method. We need several lemmas. In the following, tr X denotes the usual trace of an $n \times n$ matrix X and $\|X\|_1$ the trace norm of X, i.e. $\|X\|_1 = \mathrm{tr}\,|X| = \sum_{j=1}^{n} \mu_j(X)$.

LEMMA 1.8. *If A and B are Hermitian, then*

$$\sum_{j=1}^{n} |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|_1.$$

P r o o f. Take the Jordan decomposition $A - B = (A - B)_+ - (A - B)_-$ and let $C = A + (A - B)_- = B + (A - B)_+$. Then we have

$$\|A - B\|_1 = \mathrm{tr}(A - B)_+ + \mathrm{tr}(A - B)_-$$
$$= 2\,\mathrm{tr}\,C - \mathrm{tr}\,A - \mathrm{tr}\,B$$
$$= \sum_{j=1}^{n} \{2\lambda_j(C) - \lambda_j(A) - \lambda_j(B)\}.$$

Since $A \leq C$ and $B \leq C$, it follows from Proposition 1.4(6) that $\lambda_j(A) \leq \lambda_j(C)$ and $\lambda_j(B) \leq \lambda_j(C)$. Hence

$$|\lambda_j(A) - \lambda_j(B)| \leq 2\lambda_j(C) - \lambda_j(A) - \lambda_j(B),$$

which gives the conclusion. ∎

The following simple observation is due to Wielandt (see [61]).

LEMMA 1.9. *For any $n \times n$ matrix X let $\hat{X} = \begin{bmatrix} 0 & X^* \\ X & 0 \end{bmatrix}$, a $2n \times 2n$ Hermitian matrix. Then*

$$\lambda_j(\hat{X}) = \begin{cases} \mu_j(X), & 1 \leq j \leq n, \\ -\mu_{2n-j+1}(X), & n+1 \leq j \leq 2n. \end{cases}$$

P r o o f. We get

$$|\hat{X}| = \begin{bmatrix} |X| & 0 \\ 0 & |X^*| \end{bmatrix},$$

$$\hat{X}_+ = \frac{1}{2}(|\hat{X}| + \hat{X}) = \frac{1}{2}\begin{bmatrix} |X| & X^* \\ X & |X^*| \end{bmatrix},$$

$$\hat{X}_- = \frac{1}{2}(|\hat{X}| - \hat{X}) = \frac{1}{2}\begin{bmatrix} |X| & -X^* \\ -X & |X^*| \end{bmatrix}.$$

Since $\lambda(|X|) = \lambda(|X^*|) = \mu(X)$, we have

$$\lambda_{2j-1}(|\hat{X}|) = \lambda_{2j}(|\hat{X}|) = \mu_j(X), \quad 1 \leq j \leq n.$$

But since

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \hat{X}_+ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \hat{X}_-,$$

it follows that

$$\lambda_j(\hat{X}_+) = \lambda_j(\hat{X}_-) = \lambda_{2j}(|\hat{X}|), \quad 1 \leq j \leq n.$$

These give the assertion. ∎

LEMMA 1.10. *For any A and B,*

$$\sum_{j=1}^{n} |\mu_j(A) - \mu_j(B)| \leq \|A - B\|_1.$$

P r o o f. Let $\hat{A} = \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$ and $\hat{B} = \begin{bmatrix} 0 & B^* \\ B & 0 \end{bmatrix}$. Then

$$\|\hat{A} - \hat{B}\|_1 = \operatorname{tr} \begin{bmatrix} |A - B| & 0 \\ 0 & |(A-B)^*| \end{bmatrix} = 2\|A - B\|_1.$$

By Lemmas 1.8 and 1.9 we have

$$\|\hat{A} - \hat{B}\|_1 \geq \sum_{j=1}^{2n} |\lambda_j(\hat{A}) - \lambda_j(\hat{B})| = 2\sum_{j=1}^{n} |\mu_j(A) - \mu_j(B)|. \quad \blacksquare$$

The right-hand side of (1.7) below is known as the K-functional in the real interpolation theory.

LEMMA 1.11. *For any A and $1 \leq k \leq n$,*

$$(1.7) \qquad \sum_{j=1}^{k} \mu_j(A) = \min\{\|X\|_1 + k\|Y\|_\infty : A = X + Y\}.$$

P r o o f. For any decomposition $A = X + Y$, since $\mu_j(A) \leq \mu_j(X) + \|Y\|_\infty$ by Proposition 1.4(10), we get

$$\sum_{j=1}^{k} \mu_j(A) \leq \sum_{j=1}^{k} \mu_j(X) + k\|Y\|_\infty \leq \|X\|_1 + k\|Y\|_\infty.$$

For the converse, take the polar decomposition $A = V|A|$ and the spectral decompositon $|A| = \sum_{j=1}^{n} \mu_j(A)P_j$ with mutually orthogonal projections P_j of rank 1. Define

$$X = V \sum_{j=1}^{k} \{\mu_j(A) - \mu_k(A)\}P_j, \quad Y = V\Big\{\mu_k(A) \sum_{j=1}^{k} P_j + \sum_{j=k+1}^{n} \mu_j(A)P_j\Big\}.$$

Then $X + Y = A$ and

$$\|X\|_1 = \sum_{j=1}^{k} \mu_j(A) - k\mu_k(A), \quad \|Y\|_\infty = \mu_k(A),$$

so that

$$\|X\|_1 + k\|Y\|_\infty = \sum_{j=1}^{k} \mu_j(A). \quad \blacksquare$$

P r o o f   o f   T h e o r e m   1.5. Fix k and choose, by Lemma 1.11, X and Y such that $A - B = X + Y$ and

$$\sum_{j=1}^{k} \mu_j(A - B) = \|X\|_1 + k\|Y\|_\infty.$$

Let $a_j = \mu_j(X + B) - \mu_j(B)$ and $b_j = \mu_j(A) - \mu_j(X + B)$ for $1 \leq j \leq n$. Since $a_j + b_j = \mu_j(A) - \mu_j(B)$, considering the diagonal matrices we have by Lemma 1.11,

$$\sum_{j=1}^{k} |\mu(A) - \mu(B)|_{[j]} = \sum_{j=1}^{k} \mu_j(\mathrm{Diag}(\mu_1(A) - \mu_1(B), \ldots, \mu_n(A) - \mu_n(B)))$$

$$\leq \|\mathrm{Diag}(a_1, \ldots, a_n)\|_1 + k\|\mathrm{Diag}(b_1, \ldots, b_n)\|_\infty$$

$$= \sum_{j=1}^{n} |a_j| + k \max_{1 \leq j \leq n} |b_j|.$$

But Lemma 1.10 gives

$$\sum_{j=1}^{n} |a_j| = \sum_{j=1}^{n} |\mu_j(X + B) - \mu_j(B)| \leq \|X\|_1,$$

while by Proposition 1.4(10)

$$\max_{1 \leq j \leq n} |b_j| = \max_{1 \leq j \leq n} |\mu_j(A) - \mu_j(X + B)| \leq \|A - (X + B)\|_\infty = \|Y\|_\infty.$$

Therefore

$$\sum_{j=1}^{k} |\mu(A) - \mu(B)|_{[j]} \leq \|X\|_1 + k\|Y\|_\infty = \sum_{j=1}^{k} \mu_j(A - B),$$

as desired. $\blacksquare$

Proof of Theorem 1.6. Let A and B be Hermitian. Since

$$\sum_{j=1}^{n} \{\lambda_j(A) - \lambda_j(B)\} = \operatorname{tr}(A - B) = \sum_{j=1}^{n} \lambda_j(A - B),$$

it suffices to show that $\lambda(A) - \lambda(B) \prec_w \lambda(A - B)$. Choose $a, b \in \mathbb{R}$ such that $A + aI \geq B + bI \geq 0$. Then Theorem 1.5 gives

$$\lambda(A + aI) - \lambda(B + bI) \prec_w \lambda(A - B + (a - b)I).$$

So the assertion follows from $\lambda(A + aI) = \lambda(A) + a$ and analogous equalities for $\lambda(B + bI)$ and $\lambda(A - B + (a - b)I)$. ∎

Another important majorization for singular values of matrices is the Gelfand–Naimark theorem as follows.

THEOREM 1.12. *For any $n \times n$ matrices A and B,*

(1.8) $$(\mu_i(A)\mu_{n-i+1}(B)) \prec_{(\log)} \mu(AB),$$

or equivalently

(1.9) $$\prod_{j=1}^{k} \mu_{i_j}(AB) \leq \prod_{j=1}^{k} \{\mu_j(A)\mu_{i_j}(B)\}$$

for all $1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

An immediate corollary of this theorem is the majorization result due to Horn.

COROLLARY 1.13. *For any matrices A and B,*

$$\mu(AB) \prec_{(\log)} \mu(A)\mu(B),$$

where $\mu(A)\mu(B) = (\mu_i(A)\mu_i(B))$.

The following proof of Theorem 1.12 is a modification of [64] where (1.9) was proved in the von Neumann algebra setting.

LEMMA 1.14. *Assume that A is invertible. Then for any $1 \leq k \leq n$, there exist X and Y such that $X \geq I$, $A = YX$, and*

$$\prod_{j=1}^{k} \mu_j(A) = \det X \cdot \|Y\|_{\infty}^{k},$$

where $\det X$ denotes the determinant of X.

Proof. Let $A = V|A|$ be the polar decomposition and $|A| = \sum_{j=1}^{n} \mu_j(A)P_j$ the spectral decomposition. Define

$$X = \sum_{j=1}^{k} \frac{\mu_j(A)}{\mu_k(A)} P_j + \sum_{j=k+1}^{n} P_j, \quad Y = V\Big\{ \mu_k(A) \sum_{j=1}^{k} P_j + \sum_{j=k+1}^{n} \mu_j(A)P_j \Big\}.$$

Then we have $X \geq I$ and $YX = A$. Moreover,

$$\det X \cdot \|Y\|_{\infty}^{k} = \prod_{j=1}^{k} \frac{\mu_j(A)}{\mu_k(A)} \cdot \mu_k(A)^k = \prod_{j=1}^{k} \mu_j(A). \quad \blacksquare$$

Proof of Theorem 1.12. To prove (1.8) and (1.9), we may assume by Proposition 1.14(10) that A and B are invertible. Note that

$$\prod_{i=1}^{n} \{\mu_i(A)\mu_{n-i+1}(B)\} = |\det A \cdot \det B| = \prod_{i=1}^{n} \mu_i(AB).$$

Hence (1.8) is equivalent to (1.9) if we replace A and B by AB and B^{-1}, respectively, in (1.8). Now let $1 \le i_1 < \ldots < i_k \le n$ and X, Y be as in Lemma 1.14. Since

$$\mu_i(AB) = \mu_i(YXB) \le \|Y\|_{\infty} \mu_i(XB), \quad 1 \le i \le n,$$

we get

$$\prod_{j=1}^{k} \mu_{i_j}(AB) \le \|Y\|_{\infty}^{k} \cdot \prod_{j=1}^{k} \mu_{i_j}(XB).$$

On the other hand, since by Proposition 1.4,

$$\mu_i(XB) = \mu_i(B^*X^2B)^{1/2} \ge \mu_i(B^*B)^{1/2} = \mu_i(B), \quad 1 \le i \le n,$$

we get

$$\prod_{j=1}^{k} \frac{\mu_{i_j}(XB)}{\mu_{i_j}(B)} \le \prod_{i=1}^{n} \frac{\mu_i(XB)}{\mu_i(B)} = \frac{\det|XB|}{\det|B|} = \det X.$$

Therefore

$$\prod_{j=1}^{k} \mu_{i_j}(AB) \le \det X \cdot \|Y\|_{\infty}^{k} \cdot \prod_{j=1}^{k} \mu_{i_j}(B) = \prod_{j=1}^{k} \{\mu_j(A)\mu_{i_j}(B)\},$$

showing (1.9). ∎

Remark 1.15. The above proof is a multiplicative or logarithmic counterpart of the real interpolation method in the proof of Theorem 1.5. As shown in [64], for $1 \le k \le n$ we have

$$(1.10) \qquad \log \prod_{j=1}^{k} \mu_j(A) = \min\{\|\log|X|\|_1 + k \log \|Y\|_{\infty} : A = YX\}$$

whenever $\mu_k(A) > 0$. In fact, if $A = YX$ then

$$\log \prod_{j=1}^{k} \mu_j(A) \le \sum_{j=1}^{k} \log\{\|Y\|_{\infty}\mu_j(X)\} \le \sum_{j=1}^{k} \lambda_j(\log|X|) + k \log \|Y\|_{\infty}$$

$$\le \|\log|X|\|_1 + k \log \|Y\|_{\infty}.$$

This together with the proof of Lemma 1.14 shows (1.10).

1.4. *Majorization for operators.* In this subsection let us extend Theorems 1.5 and 1.12 to generalized singular values $\mu(A) = (\mu_1(A), \mu_2(A), \ldots)$ of infinite-dimensional operators. The next lemma is useful for this sake.

LEMMA 1.16. *Let $A \in \mathcal{B}(\mathcal{H})$, $n \in \mathbb{N}$, and $\varepsilon > 0$. Then there exists a projection P of finite rank such that*

$$\mu_i(PAP) \ge (\mu_i(A) - \varepsilon)_+, \quad 1 \le i \le n.$$

Furthermore, there exists an orthonormal set $\{\xi_1,\ldots,\xi_n\}$ such that

$$\| \, |A|\xi_i - \mu_i(A)\xi_i\| \le \varepsilon, \quad 1 \le i \le n.$$

P r o o f. When $\mu_m(A) > 0 = \mu_{m+1}(A)$ for some $1 \le m < n$, we may consider $(\mu_1(A),\ldots,\mu_m(A))$ instead of $(\mu_1(A),\ldots,\mu_n(A))$. So it suffices to assume that $\mu_n(A) > \varepsilon > 0$. Let $\alpha = \mu_\infty(A)$ $(= \lim_{n\to\infty}\mu_n(A))$. Suppose that $\mu_n(A) = \alpha$ and $\mu_{k-1}(A) > \alpha = \mu_k(A) = \ldots = \mu_n(A)$ for some $1 \le k \le n$. Then all $\mu_i(A)$, $1 \le i < k$ (if any), are eigenvalues of $|A|$ and $E_{|A|}(\alpha) - E_{|A|}(\alpha - \varepsilon)$ is of infinite rank. So we can choose an orthonormal set $\{\xi_1,\ldots,\xi_n\}$ such that $|A|\xi_i = \mu_i(A)\xi_i$ for $1 \le i < k$ and ξ_i, $k \le i \le n$, are in the range of $E_{|A|}(\alpha) - E_{|A|}(\alpha - \varepsilon)$. Let P_0 and P be the projections onto the subspaces spanned by $\{\xi_1,\ldots,\xi_n\}$ and $\{\xi_1,\ldots,\xi_n,A\xi_1,\ldots,A\xi_n\}$, respectively. For $1 \le i \le n$ we get

$$\mu_i(PAP) \ge \mu_i((PAP)P_0) = \mu_i(AP_0) = \mu_i(P_0|A|^2 P_0)^{1/2} \ge \alpha - \varepsilon.$$

For $1 \le i < k$, since $\mu_i(P_0|A|^2 P_0)^{1/2} = \mu_i(A)$, we get $\mu_i(PAP) = \mu_i(A)$. Hence the required condition is fulfilled. When $\mu_n(A) > \alpha$, all $\mu_i(A)$, $1 \le i \le n$, are eigenvalues of $|A|$ and the proof is done as above. Also the second assertion follows from the above proof. ∎

THEOREM 1.17. *If $A, B \in \mathcal{B}(\mathcal{H})$, then*

$$(1.11) \qquad \sum_{j=1}^{k} |\mu_{i_j}(A) - \mu_{i_j}(B)| \le \sum_{j=1}^{k} \mu_j(A - B)$$

for all $1 \le i_1 < i_2 < \ldots < i_k$. In particular,

$$(1.12) \qquad \mu(A + B) \prec_w \mu(A) + \mu(B), \quad A, B \in \mathcal{B}(\mathcal{H}).$$

P r o o f. Let $1 \le i_1 < \ldots < i_k = n$. For any $\varepsilon > 0$ there exists, by Lemma 1.16, projections P, Q of finite rank such that $\mu_i(PAP) \ge (\mu_i(A) - \varepsilon)_+$ and $\mu_i(QBQ) \ge (\mu_i(B) - \varepsilon)_+$ for $1 \le i \le n$. Let $E = P \vee Q$, a projection of finite rank. Then for $1 \le i \le n$,

$$\mu_i(A) \ge \mu_i(EAE) \ge \mu_i(PAP) \ge (\mu_i(A) - \varepsilon)_+$$

and $\mu_i(B) \ge \mu_i(EBE) \ge (\mu_i(B) - \varepsilon)_+$. Applying Theorem 1.5 to $EAE, EBE \in \mathcal{B}(E\mathcal{H})$ (considered as matrices), we have

$$\sum_{j=1}^{k} |\mu_{i_j}(EAE) - \mu_{i_j}(EBE)| \le \sum_{j=1}^{k} \mu_j(E(A - B)E) \le \sum_{j=1}^{k} \mu_j(A - B).$$

Letting $\varepsilon \downarrow 0$ we obtain (1.11), which implies (1.12) by letting $i_j = j$ and replacing A by $A + B$. ∎

THEOREM 1.18. *If $A, B \in \mathcal{B}(\mathcal{H})$, then*

$$(1.13) \qquad \prod_{j=1}^{k} \mu_{i_j}(AB) \le \prod_{j=1}^{k} \{\mu_j(A)\mu_{i_j}(B)\}$$

for all $1 \le i_1 < i_2 < \ldots < i_k$. In particular,

$$(1.14) \qquad \mu(AB) \prec_{w(\log)} \mu(A)\mu(B), \quad A, B \in \mathcal{B}(\mathcal{H}).$$

Proof. Let $1 \leq i_1 < \ldots < i_k = n$ and $\varepsilon > 0$. By Lemma 1.16 there exists a projection P of finite rank such that $\mu_i(PABP) \geq (\mu_i(AB) - \varepsilon)_+$ for $1 \leq i \leq n$. Let E be the projection onto the subspace spanned by $P\mathcal{H} \cup BP\mathcal{H}$, which is of finite rank. Then $PABP = (PAE)(EBP)$. Applying (1.9) to $PAE, EBP \in \mathcal{B}(E\mathcal{H})$, we have

$$\sum_{j=1}^{k} \mu_{i_j}(PABP) \leq \prod_{j=1}^{k} \{\mu_j(PAE)\mu_{i_j}(EBP)\} \leq \prod \{\mu_j(A)\mu_{i_j}(B)\},$$

which shows (1.13) as $\varepsilon \downarrow 0$. ∎

2. Symmetric norms and symmetrically normed ideals. This section is a self-contained review on symmetric norms and symmetrically normed ideals. Our exposition is somewhat restricted to the material which will be necessary in the subsequent sections. See [32, 79] (also [77]) for full theory on the subject.

2.1. *Symmetric gauge functions and symmetric norms.* Let s_{fin} denote the linear space of all infinite sequences of real numbers having only finitely many nonzero terms. A norm Φ on s_{fin} is called to be *symmetric* if Φ satisfies

(2.1) $$\Phi(a_1, a_2, \ldots) = \Phi(\varepsilon_1 a_{\pi(1)}, \varepsilon_2 a_{\pi(2)}, \ldots)$$

for any permutation π on $\mathbb{N}$ and $\varepsilon_i = \pm 1$. This condition is equivalently written as

$$\Phi(a_1, a_2, \ldots) = \Phi(a_1^*, a_2^*, \ldots)$$

where $(a_1^*, a_2^*, \ldots)$ is the decreasing rearrangement of $(|a_1|, |a_2|, \ldots)$. A symmetric norm on s_{fin} is called a *symmetric gauge function*.

LEMMA 2.1. *Let Φ be a symmetric gauge function.*

(1) *If $(a_i), (b_i) \in s_{\text{fin}}$ and $|a_i| \leq |b_i|$ for $i \in \mathbb{N}$, then*

$$\Phi(a_1, a_2, \ldots) \leq \Phi(b_1, b_2, \ldots).$$

(2) *Under the normalization $\Phi(1, 0, 0, \ldots) = 1$,*

$$\sup_i |a_i| \leq \Phi(a_1, a_2, \ldots) \leq \sum_i |a_i|, \quad (a_i) \in s_{\text{fin}},$$

that is, the ℓ_∞-norm (resp. ℓ_1-norm) is the least (resp. greatest) symmetric gauge function.

Proof. (1) In view of (2.1) we may show that for $0 \leq \alpha \leq 1$,

$$\Phi(\alpha a_1, a_2, a_3, \ldots) \leq \Phi(a_1, a_2, a_3, \ldots).$$

This is seen as follows:

$\Phi(\alpha a_1, a_2, a_3, \ldots)$

$$= \Phi\left(\frac{1+\alpha}{2}a_1 + \frac{1-\alpha}{2}(-a_1), \frac{1+\alpha}{2}a_2 + \frac{1-\alpha}{2}a_2, \frac{1+\alpha}{2}a_3 + \frac{1-\alpha}{2}a_3, \ldots\right)$$

$$\leq \frac{1+\alpha}{2}\Phi(a_1, a_2, a_3, \ldots) + \frac{1-\alpha}{2}\Phi(-a_1, a_2, a_3, \ldots) = \Phi(a_1, a_2, a_3, \ldots).$$

(2) Since by (2.1) and (1),

$$|a_i| = \Phi(a_i, 0, 0, \ldots) \leq \Phi(a_1, \ldots, a_i, \ldots),$$

the first inequality holds. The second follows from

$$\Phi(a_1,\ldots,a_n,0,0,\ldots) \leq \sum_{i=1}^{n} \Phi(a_i,0,0,\ldots) = \sum_{i=1}^{n} |a_i|. \quad \blacksquare$$

LEMMA 2.2. *If* $(a_1,\ldots,a_n),(b_1,\ldots,b_n) \in \mathbb{R}^n$ *and* $(|a_1|,\ldots,|a_n|) \prec_w (|b_1|,\ldots,|b_n|)$, *then*

$$\Phi(a_1,\ldots,a_n,0,0,\ldots) \leq \Phi(b_1,\ldots,b_n,0,0,\ldots).$$

P r o o f. By Proposition 1.2 there exists $(c_1,\ldots,c_n) \in \mathbb{R}^n$ such that

$$(|a_1|,\ldots,|a_n|) \leq (c_1,\ldots,c_n) \prec (|b_1|,\ldots,|b_n|).$$

Proposition 1.1 says that $(c_1,\ldots,c_n)$ is a convex combination of coordinate permutations of $(|b_1|,\ldots,|b_n|)$. This implies by Lemma 2.1(1) and (2.1) that

$$\Phi(a_1,\ldots,a_n,0,0,\ldots) \leq \Phi(c_1,\ldots,c_n,0,0,\ldots) \leq \Phi(b_1,\ldots,b_n,\ldots). \quad \blacksquare$$

Let $\mathcal{C}(\mathcal{H})$ be the algebra of all compact operators on a separable Hilbert space $\mathcal{H}$. We denote by $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ the set of all operators of finite rank on $\mathcal{H}$. A norm $\|\cdot\|$ on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is called to be *unitarily invariant* if

$$\|UAV\| = \|A\|$$

for every $A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ and any unitaries U,V on $\mathcal{H}$. A unitarily invariant norm on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is also called a *symmetric norm*. The following fundamental theorem is due to von Neumann [67] (also [63, 77]).

THEOREM 2.3. *There is a bijective correspondence between symmetric gauge functions* Φ *and unitarily invariant norms* $\|\cdot\|$ *on* $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ *which is determined by the formula*

$$(2.2) \qquad \|A\| = \Phi(\mu_1(A),\mu_2(A),\ldots), \qquad A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H}).$$

Furthermore, a norm $\|\cdot\|$ *on* $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ *is unitarily invariant if and only if*

$$(2.3) \qquad \|XAY\| \leq \|X\|_\infty \|Y\|_\infty \|A\|$$

for every $A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ *and* $X,Y \in \mathcal{B}(\mathcal{H})$.

P r o o f. Suppose that Φ is a symmetric gauge function. Define $\|\cdot\|$ on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ by the formula (2.2). Let $A,B \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$. Since $\mu(A+B),\mu(A)+\mu(B) \in s_{\mathrm{fin}}$ and $\mu(A+B) \prec_w \mu(A)+\mu(B)$ by (1.12) (or Corollary 1.7(1)), we get, by Lemma 2.2,

$$\|A+B\| \leq \Phi(\mu_1(A)+\mu_1(B),\mu_2(A)+\mu_2(B),\ldots)$$
$$\leq \Phi(\mu_1(A),\mu_2(A),\ldots) + \Phi(\mu_1(B),\mu_2(B),\ldots) = \|A\| + \|B\|.$$

Also it is clear that $\|A\| = 0$ if and only if $\mu_1(A) = 0$ or $A = 0$. For $\alpha \in \mathbb{C}$ we get by Proposition 1.4(4)

$$\|\alpha A\| = \Phi(|\alpha|\mu_1(A),|\alpha|\mu_2(A),\ldots) = |\alpha| \|A\|.$$

Hence $\|\cdot\|$ is a norm on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$, which is unitarily invariant because $\mu_n(UAV) = \mu_n(A)$ for all unitaries U,V.

Conversely, suppose that $\|\cdot\|$ is a unitarily invariant norm on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$. Choose an orthonormal basis $\{\varphi_i\}$ of $\mathcal{H}$ and define $\Phi : s_{\mathrm{fin}} \to \mathbb{R}$ by

$$\Phi(a_1, a_2, \ldots) = \Big\| \sum_i a_i \varphi_i \otimes \varphi_i \Big\|, \quad (a_i) \in s_{\mathrm{fin}},$$

where $\varphi \otimes \psi$ denotes the Schatten form, i.e. $(\varphi \otimes \psi)\xi = \langle \xi, \psi \rangle \varphi$ for $\varphi, \psi, \xi \in \mathcal{H}$. Then it is immediate that Φ is a norm on s_{fin}. For any permutation π on $\mathbb{N}$ and $\varepsilon_i = \pm 1$, we can define unitaries U, V on $\mathcal{H}$ by $U\varphi_{\pi(i)} = \varepsilon_i \varphi_i$ and $V\varphi_{\pi(i)} = \varphi_i$ for $i \in \mathbb{N}$, so that

$$\Phi(a_1, a_2, \ldots) = \Big\| U\Big(\sum_i a_{\pi(i)}\varphi_{\pi(i)} \otimes \varphi_{\pi(i)}\Big)V^* \Big\| = \Big\| \sum_i a_{\pi(i)} U\varphi_{\pi(i)} \otimes V\varphi_{\pi(i)} \Big\|$$

$$= \Big\| \sum_i \varepsilon_i a_{\pi(i)} \varphi_i \otimes \varphi_i \Big\| = \Phi(\varepsilon_1 a_{\pi(i)}, \varepsilon_2 a_{\pi(2)}, \ldots).$$

Hence Φ is a symmetric gauge function. Any $A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is written as $A = \sum_{i=1}^n \mu_i(A)\xi_i \otimes \eta_i$ for some orthonormal sets $\{\xi_1, \ldots, \xi_n\}$ and $\{\eta_1, \ldots, \eta_n\}$. Take unitaries U, V such that $U\varphi_i = \xi_i$, $V\varphi_i = \eta_i$ for $1 \le i \le n$. Then we get

$$\Phi(\mu_1(A), \mu_2(A), \ldots) = \Big\| \sum_{i=1}^n \mu_i(A)\varphi_i \otimes \varphi_i \Big\| = \Big\| U\Big(\sum_{i=1}^n \mu_i(A)\varphi_i \otimes \varphi_i\Big)V^* \Big\| = \|A\|,$$

so that (2.2) holds. Therefore the first assertion is proved.

It is clear that a norm $\|\cdot\|$ on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is unitarily invariant if it satisfies (2.3). Conversely, if $\|\cdot\|$ is a unitarily invariant norm with the corresponding gauge function Φ, then by Proposition 1.4(7),

$$\|XAY\| \le \Phi(\|X\|_\infty\|Y\|_\infty \mu_1(A), \|X\|_\infty\|Y\|_\infty \mu_2(A), \ldots) = \|X\|_\infty\|Y\|_\infty\|A\|$$

for $A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ and $X, Y \in \mathcal{B}(\mathcal{H})$. ∎

2.2. *Symmetrically normed ideals.* Let Φ be a symmetric gauge function. When $a = (a_1, a_2, \ldots)$ is a bounded real sequence, we define

$$\Phi(a) = \sup_n \Phi(a_1, \ldots, a_n, 0, 0, \ldots) \in [0, \infty].$$

Let s_Φ be the set of all bounded real sequences a with $\Phi(a) < \infty$. Moreover, extending $\|\cdot\|$ on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$, we define for any $A \in \mathcal{B}(\mathcal{H})$,

(2.4) $$\|A\| = \sup_n \Phi(\mu_1(A), \ldots, \mu_n(A), 0, 0, \ldots) \in [0, \infty].$$

Let $\mathcal{C}_\Phi(\mathcal{H})$ denote the set of all $A \in \mathcal{B}(\mathcal{H})$ with $\|A\| < \infty$, i.e.

$$\mu(A) = (\mu_1(A), \mu_2(A), \ldots) \in s_\Phi.$$

In this way, a symmetric norm $\|\cdot\|$ on $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ can extend to all operators in $\mathcal{B}(\mathcal{H})$ permitting ∞. Then we have:

PROPOSITION 2.4. *Let $A, B, X, Y \in \mathcal{B}(\mathcal{H})$ and $\|\cdot\|$ be a symmetric norm.*

(1) $\|A\| = \|A^*\|$.
(2) $\|XAY\| \le \|X\|_\infty\|Y\|_\infty\|A\|$.
(3) *If $\mu(A) \prec_w \mu(B)$ (in particular, if $|A| \le |B|$), then $\|A\| \le \|B\|$.*
(4) *If $\mu(A) \prec_w \mu(B)$ and $B \in \mathcal{C}(\mathcal{H})$, then $A \in \mathcal{C}(\mathcal{H})$.*

(5) *Under the normalization $\Phi(1, 0, 0, \ldots) = 1$ (or $\|P\| = 1$ for a projection of rank one), $\|A\|_\infty \le \|A\| \le \|A\|_1$.*

P r o o f. (1) and (2) immediately follow from definition (2.4), Lemma 2.1(1), and the corresponding properties of Proposition 1.4. If $\mu(A) \prec_w \mu(B)$ then Lemma 2.2 gives

$$\Phi(\mu_1(A), \ldots, \mu_n(A), 0, 0, \ldots) \le \Phi(\mu_1(B), \ldots, \mu_n(B), 0, 0, \ldots), \quad n \in \mathbb{N},$$

and so $\|A\| \le \|B\|$. Hence (3) holds. (4) is seen from the fact that $A \in \mathcal{C}(\mathcal{H})$ if and only if $\mu_n(A) \to 0$, which is equivalent to $n^{-1} \sum_{i=1}^n \mu_i(A) \to 0$. (5) follows from Lemma 2.1(2). ∎

THEOREM 2.5. *Let Φ be a symmetric gauge function.*

(1) $\mathcal{C}_\Phi(\mathcal{H})$ *is a Banach space in the norm* (2.4) *and is a two-sided ideal of* $\mathcal{B}(\mathcal{H})$.
(2) *If Φ is inequivalent to the ℓ_∞-norm, then $\mathcal{C}_\Phi(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$.*

P r o o f. (1) Let $A, B \in \mathcal{C}_\Phi(\mathcal{H})$. For every $n \in \mathbb{N}$, since (1.12) gives

$$(\mu_1(A + B), \ldots, \mu_n(A + B)) \prec_w (\mu_1(A) + \mu_1(B), \ldots, \mu_n(A) + \mu_n(B)),$$

we have by Lemma 2.2,

$$\Phi(\mu_1(A + B), \ldots, \mu_n(A + B), 0, 0, \ldots)$$
$$\le \Phi(\mu_1(A), \ldots, \mu_n(A), 0, 0, \ldots) + \Phi(\mu_1(B), \ldots, \mu_n(B), 0, 0, \ldots) \le \|A\| + \|B\|.$$

This implies that $A + B \in \mathcal{C}_\Phi(\mathcal{H})$ and $\|A + B\| \le \|A\| + \|B\|$. Now it is easily verified that $\mathcal{C}_\Phi(\mathcal{H})$ is a linear space and $\|\cdot\|$ is a norm on $\mathcal{C}_\Phi(\mathcal{H})$. To prove the completeness, let $\{A_k\}$ be a Cauchy sequence in $\mathcal{C}_\Phi(\mathcal{H})$. Since $\{A_k\}$ is $\|\cdot\|_\infty$-Cauchy by Proposition 2.4(5), there exists $A \in \mathcal{B}(\mathcal{H})$ such that $\|A_k - A\|_\infty \to 0$. For every $n \in \mathbb{N}$, since

$$\Phi(\mu_1(A_k - A_m), \ldots, \mu_n(A_k - A_m), 0, 0, \ldots) \le \|A_k - A_m\|$$

and $\mu_i(A_k - A_m) \to \mu_i(A_k - A)$ as $m \to \infty$ thanks to Proposition 1.4(10), we get

$$\Phi(\mu_1(A_k - A), \ldots, \mu_n(A_k - A), 0, 0, \ldots) \le \lim_{m \to \infty} \|A_k - A_m\|.$$

Therefore $\|A_k - A\| \le \lim_m \|A_k - A_m\|$, which implies that $A \in \mathcal{C}_\Phi(\mathcal{H})$ and $\|A_k - A\| \to 0$ as $k \to \infty$. It is immediate from Proposition 2.4(2) that $\mathcal{C}_\Phi(\mathcal{H})$ is a two-sided ideal of $\mathcal{B}(\mathcal{H})$.

(2) Suppose that $\mathcal{C}_\Phi(\mathcal{H})$ contains a non-compact operator A. Then we have $\alpha = \inf_n \mu_n(A) > 0$ and

$$\alpha \Phi(1, 1, 1, \ldots) = \Phi(\alpha, \alpha, \alpha, \ldots) \le \Phi(\mu_1(A), \mu_2(A), \ldots) = \|A\| < \infty,$$

so that $\Phi(a_1, a_2, \ldots) \le \alpha^{-1} \|A\| \sup_n |a_n|$. This together with Lemma 2.1(2) implies that Φ is equivalent to the ℓ_∞-norm. ∎

The Banach space $\mathcal{C}_\Phi(\mathcal{H})$ as well as $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ defined below is called a *symmetrically normed ideal* associated with a symmetric gauge function Φ. It is seen as in Theorem 2.5 that s_Φ is a Banach space in the norm Φ. In fact, if $\{\varphi_i\}$ is an orthonormal basis of $\mathcal{H}$, then it is not difficult to show that $\Phi(a) = \|\sum_i a_i \varphi_i \otimes \varphi_i\|$ for any bounded real sequence $a = (a_i)$ and s_Φ is isometrically imbedded in $\mathcal{C}_\Phi(\mathcal{H})$ by $a \in s_\Phi \mapsto \sum_i a_i \varphi_i \otimes \varphi_i$ whose range is a closed (real) subspace of $\mathcal{C}_\Phi(\mathcal{H})$. Let $s_\Phi^{(0)}$ be the closure of s_{fin} in s_Φ.

We denote by $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ the set of all $A \in \mathcal{B}(\mathcal{H})$ such that $\mu(A) \in s_\Phi^{(0)}$. When $s_\Phi = s_\Phi^{(0)}$ or $\mathcal{C}_\Phi(\mathcal{H}) = \mathcal{C}_\Phi^{(0)}(\mathcal{H})$, Φ is called *regular*.

THEOREM 2.6. *Let* Φ, $\mathcal{C}_\Phi(\mathcal{H})$, *and* $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ *be as above.*

(1) $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ *is the closure of* $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ *in* $\mathcal{C}_\Phi(\mathcal{H})$ *and* $\mathcal{C}_\Phi^{(0)}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$.

(2) $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ *is a Banach space in the norm* $\|\cdot\|$ *of* $\mathcal{C}_\Phi(\mathcal{H})$ *and is a two-sided ideal of* $\mathcal{B}(\mathcal{H})$.

P r o o f. (1) It is clear that $\mathcal{C}_{\mathrm{fin}}(\mathcal{H}) \subset \mathcal{C}_\Phi^{(0)}(\mathcal{H})$. It follows from Lemma 2.1(2) that if $a = (a_1, a_2, \ldots) \in s_\Phi^{(0)}$, then a belongs to the ℓ_∞-norm closure of s_{fin} so that $a_n \to 0$. This shows that $\mathcal{C}_\Phi^{(0)}(\mathcal{H}) \subset \mathcal{C}(\mathcal{H})$. So each $A \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$ is written as $A = \sum_i \mu_i(A)\xi_i \otimes \eta_i$ with orthonormal sequences $\{\xi_i\}$ and $\{\eta_i\}$. Then

$$\left\| A - \sum_{i=1}^n \mu_i(A)\xi_i \otimes \eta_i \right\| = \left\| \sum_{i=n+1}^\infty \mu_i(A)\xi_i \otimes \eta_i \right\| = \Phi(\mu_{n+1}(A), \mu_{n+2}(A), \ldots).$$

But it is easily seen from $\mu(A) \in s_\Phi^{(0)}$ that

$$\lim_{n\to\infty} \Phi(\mu_{n+1}(A), \mu_{n+2}(A), \ldots) = 0.$$

Hence A is in the closure of $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$. Conversely, let A be in the closure of $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$. For any $\varepsilon > 0$ there exists $B \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ such that $\|A - B\| < \varepsilon$. Since (1.11) says that for every $n \in \mathbb{N}$,

$$(|\mu_1(A) - \mu_1(B)|, \ldots, |\mu_n(A) - \mu_n(B)|) \prec_w (\mu_1(A - B), \ldots, \mu_n(A - B)),$$

it follows from Lemma 2.2 that

$$\Phi(\mu_1(A) - \mu_1(B), \mu_2(A) - \mu_2(B), \ldots) \leq \|A - B\| < \varepsilon.$$

Since $\mu(B) \in s_{\mathrm{fin}}$, we have $\mu(A) \in s_\Phi^{(0)}$ and so $A \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$.

(2) The first assertion follows from (1). The second is easily verified by using Proposition 2.4(2). ∎

For instance, let Φ_p be the ℓ_p-norm and $\|\cdot\|_p$ the corresponding symmetric norm where $1 \leq p \leq \infty$. When $1 \leq p < \infty$ we have $\mathcal{C}_{\Phi_p}(\mathcal{H}) = \mathcal{C}_p(\mathcal{H})$, the *Schatten p-class*. More generally, for $0 < p < \infty$ we can define the *p-class* $\mathcal{C}_p(\mathcal{H})$ as the space of all $A \in \mathcal{C}(\mathcal{H})$ such that

$$\|A\|_p = (\mathrm{tr}\,|A|^p)^{1/p} = \left\{ \sum_i \mu_i(A)^p \right\}^{1/p} < \infty.$$

But when $0 < p < 1$, $\|\cdot\|_p$ is not a norm but a quasi-norm. In particular, $\mathcal{C}_1(\mathcal{H})$ is the *trace class* and $\mathcal{C}_2(\mathcal{H})$ is the *Hilbert-Schmidt class*. When $p = \infty$ we have $\mathcal{C}_{\Phi_\infty}(\mathcal{H}) = \mathcal{B}(\mathcal{H})$ and $\mathcal{C}_{\Phi_\infty}^{(0)}(\mathcal{H}) = \mathcal{C}(\mathcal{H})$. Note that $\|A\|_p \geq \|A\|_q$ and $\mathcal{C}_p(\mathcal{H}) \subset \mathcal{C}_q(\mathcal{H})$ if $0 < p < q \leq \infty$. Another important class of symmetric norms is the Ky Fan norms $\|\cdot\|_{(k)}$ defined by

$$\|A\|_{(k)} = \sum_{i=1}^k \mu_i(A), \quad k \in \mathbb{N}.$$

Obviously, all $\|\cdot\|_{(k)}$ are equivalent to $\|\cdot\|_\infty$ and $\|\cdot\|_{(1)} = \|\cdot\|_\infty$. See [32] for more delicate examples of symmetrically normed ideals such as Macaev ideals.

Here we show the Hölder inequality to illustrate the usefulness of the majorization technique.

PROPOSITION 2.7. *Let* $0 < p, p_1, p_2 \leq \infty$ *and* $1/p = 1/p_1 + 1/p_2$. *Then*

$$\|AB\|_p \leq \|A\|_{p_1} \|B\|_{p_2}, \quad A, B \in \mathcal{B}(\mathcal{H}).$$

Hence $AB \in \mathcal{C}_p(\mathcal{H})$ *if* $A \in \mathcal{C}_{p_1}(\mathcal{H})$ *and* $B \in \mathcal{C}_{p_2}(\mathcal{H})$.

P r o o f. Suppose that $0 < p_1, p_2 < \infty$, because the result is obvious by Proposition 1.4(7) when $p_1 = \infty$ or $p_2 = \infty$. Since by (1.14)

$$(\mu_i(AB)^p) \prec_{w(\log)} (\mu_i(A)^p \mu_i(B)^p),$$

it follows from Proposition 1.3 that

$$(\mu_i(AB)^p) \prec_w (\mu_i(A)^p \mu_i(B)^p).$$

Since $(p_1/p)^{-1} + (p_2/p)^{-1} = 1$, the usual Hölder inequality for vectors shows that for every $n \in \mathbb{N}$,

$$\left\{ \sum_{i=1}^n \mu_i(AB)^p \right\}^{1/p} \leq \left\{ \sum_{i=1}^n \mu_i(A)^p \mu_i(B)^p \right\}^{1/p}$$

$$\leq \left\{ \sum_{i=1}^n \mu_i(A)^{p_1} \right\}^{1/p_1} \left\{ \sum_{i=1}^n \mu_i(B)^{p_2} \right\}^{1/p_2} \leq \|A\|_{p_1} \|B\|_{p_2}.$$

This yields the conclusion. ∎

Corresponding to each symmetric gauge function Φ, let us define $\Phi' : s_{\mathrm{fin}} \to \mathbb{R}$ by

$$\Phi'(b_1, b_2, \ldots) = \sup \left\{ \sum_i a_i b_i : a \in s_{\mathrm{fin}}, \ \Phi(a) \leq 1 \right\}.$$

Then it is an easy task to check that Φ' is again a symmmetric norm on s_{fin}. The symmetric gauge function Φ' is said to be *conjugate* to Φ. Note that $\Phi'' = \Phi$. For example, when $1 \leq p \leq \infty$ and $1/p + 1/q = 1$, the ℓ_p-norm is conjugate to the ℓ_q-norm.

The following generalized Hölder inequality can be shown as Proposition 2.7.

LEMMA 2.8. *Let* Φ, Φ_1, Φ_2 *be symmetric gauge functions with the corresponding norms* $\| \cdot \|, \|\| \cdot \|\|_1, \|\| \cdot \|\|_2$. *If*

$$\Phi(a_1 b_1, a_2 b_2, \ldots) \leq \Phi_1(a)\Phi_2(b), \quad a, b \in s_{\mathrm{fin}},$$

then

$$\|AB\| \leq \|\|A\|\|_1 \|\|B\|\|_2, \quad A, B \in \mathcal{B}(\mathcal{H}).$$

In particular, if $\| \cdot \|'$ *is the symmetric norm corresponding to* Φ' *conjugate to* Φ, *then* $\|AB\|_1 \leq \|A\| \|B\|'$ *for every* $A, B \in \mathcal{B}(\mathcal{H})$. *Hence* $AB \in \mathcal{C}_1(\mathcal{H})$ *if* $A \in \mathcal{C}_\Phi(\mathcal{H})$ *and* $B \in \mathcal{C}_{\Phi'}(\mathcal{H})$.

P r o o f. By (1.14), Proposition 1.3, and Lemma 2.2, we have for every $n \in \mathbb{N}$

$\Phi(\mu_1(AB), \ldots, \mu_n(AB), 0, 0, \ldots)$

$\quad \leq \Phi(\mu_1(A)\mu_1(B), \ldots, \mu_n(A)\mu_n(B), 0, 0, \ldots)$

$\quad \leq \Phi_1(\mu_1(A), \ldots, \mu_n(A), 0, 0, \ldots)\Phi_2(\mu_1(B), \ldots, \mu_n(B), 0, 0, \ldots) \leq \|\|A\|\|_1 \|\|B\|\|_1,$

showing the first assertion. For the second part, note by definition of Φ' that

$$\sum_i |a_i b_i| \leq \Phi(a)\Phi'(b), \quad a, b \in s_{\mathrm{fin}}. \quad \blacksquare$$

THEOREM 2.9. *If Φ and Φ' are conjugate symmetric gauge functions, then the dual Banach space $\mathcal{C}_\Phi^{(0)}(\mathcal{H})^*$ of $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$ is isometrically isomorphic to $\mathcal{C}_{\Phi'}(\mathcal{H})$ by the duality $(A, B) \mapsto \mathrm{tr}(AB)$ for $A \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$ and $B \in \mathcal{C}_{\Phi'}(\mathcal{H})$, where* tr *denotes the usual trace on $\mathcal{C}_1(\mathcal{H})$.*

P r o o f. For any $B \in \mathcal{C}_{\Phi'}(\mathcal{H})$ we can define, by Lemma 2.8, a linear functional $f_B : \mathcal{C}_\Phi^{(0)} \to \mathbb{C}$ by $f_B(A) = \mathrm{tr}(AB)$. Since

$$|\mathrm{tr}(AB)| \leq \|AB\|_1 \leq \|A\| \, \|B\|', \quad A \in \mathcal{C}_\Phi^{(0)}(\mathcal{H}),$$

it follows that $f_B \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})^*$ and $\|f_B\| \leq \|B\|'$. Now let $f \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})^*$ and consider a sesqui-linear form $(\xi, \eta) \mapsto f(\xi \otimes \eta)$, $\xi, \eta \in \mathcal{H}$. Since

$$|f(\xi \otimes \eta)| \leq \|f\| \, \|\xi \otimes \eta\| = \|f\|\Phi(1, 0, 0, \ldots)\|\xi\| \, \|\eta\|,$$

there exists $B \in \mathcal{B}(\mathcal{H})$ such that $\langle B\xi, \eta \rangle = f(\xi \otimes \eta)$ for all $\xi, \eta \in \mathcal{H}$. For each $n \in \mathbb{N}$ and $\varepsilon > 0$, by Lemma 1.16 there exists a projection P of finite rank such that $\mu_i(PBP) \geq (\mu_i(B) - \varepsilon)_+$ for $1 \leq i \leq n$ and also $m = \mathrm{rank}\, P \geq n$. We write PBP as $PBP = \sum_{i=1}^m \mu_i(PBP)\xi_i \otimes \eta_i$ with orthonormal bases $\{\xi_i\}$ and $\{\eta_i\}$ of $P\mathcal{H}$. Since

$$\mu_i(PBP) = \langle PBP\eta_i, \xi_i \rangle = \langle B\eta_i, \xi_i \rangle = f(\eta_i \otimes \xi_i),$$

if $a \in s_{\mathrm{fin}}$ and $\Phi(a) \leq 1$, then

$$\sum_{i=1}^n a_i \mu_i(PBP) = f\Big(\sum_{i=1}^n a_i \eta_i \otimes \xi_i\Big) \leq \|f\| \, \Big\|\sum_{i=1}^n a_i \eta_i \otimes \xi_i\Big\|$$

$$= \|f\|\Phi(a_1, \ldots, a_n, 0, 0, \ldots) \leq \|f\|.$$

This shows that

$$\Phi'(\mu_1(PBP), \ldots, \mu_n(PBP), 0, 0, \ldots) \leq \|f\|.$$

Letting $\varepsilon \downarrow 0$ we get

$$\Phi'(\mu_1(B), \ldots, \mu_n(B), 0, 0, \ldots) \leq \|f\|, \quad n \in \mathbb{N},$$

so that $B \in \mathcal{C}_{\Phi'}(\mathcal{H})$ and $\|B\|' \leq \|f\|$. Since $f(A) = f_B(A)$ for $A \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ and $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is dense in $\mathcal{C}_\Phi^{(0)}(\mathcal{H})$, we have $f = f_B$ and hence $\|f_B\| = \|B\|'$. Thus $B \in \mathcal{C}_{\Phi'}(\mathcal{H}) \mapsto f_B \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})^*$ is a surjective isometry. $\blacksquare$

As special cases we have $\mathcal{C}_1(\mathcal{H})^* \cong \mathcal{B}(\mathcal{H})$ and $\mathcal{C}_p(\mathcal{H})^* \cong \mathcal{C}_q(\mathcal{H})$ when $1 < p < \infty$, $1/p + 1/q = 1$. The above theorem shows that $\mathcal{C}_\Phi(\mathcal{H})^* \cong \mathcal{C}_{\Phi'}(\mathcal{H})$ if Φ is regular and that $\mathcal{C}_\Phi(\mathcal{H})$ is reflexive if and only if both Φ and Φ' are regular.

2.3. *Further properties of symmetric norms.* In this subsection let us present, for later use, some further results concerning symmetric norms. The close relation between the (log-)majorization and the symmetric norm inequalities is summarized in the following proposition.

PROPOSITION 2.10. *Consider the following conditions for $A, B \in \mathcal{B}(\mathcal{H})$. Then*:

$$(\mathrm{i}) \Leftrightarrow (\mathrm{ii}) \Rightarrow (\mathrm{iii}) \Leftrightarrow (\mathrm{iv}) \Leftrightarrow (\mathrm{v}) \Leftrightarrow (\mathrm{vi}).$$

(i) $\mu(A) \prec_{w(\log)} \mu(B)$;

(ii) $\|f(|A|)\| \leq \|f(|B|)\|$ *for every symmetric norm* $\| \cdot \|$ *and every continuous increasing function f on $[0, \infty)$ such that $f(0) \geq 0$ and $f(e^x)$ is convex*;

(iii) $\mu(A) \prec_w \mu(B)$;

(iv) $\|A\|_{(k)} \leq \|B\|_{(k)}$ *for every $k \in \mathbb{N}$*;

(v) $\|A\| \leq \|B\|$ *for every symmetric norm $\| \cdot \|$*;

(vi) $\|f(|A|)\| \leq \|f(|B|)\|$ *for every symmetric norm $\| \cdot \|$ and every increasing convex function f on $[0, \infty)$ such that $f(0) \geq 0$*.

P r o o f. (i)$\Rightarrow$(ii). Let f be as in (ii). By Propositions 1.3 and 1.4(11) we have

$$(2.5) \qquad \mu(f(|A|)) = f(\mu(A)) \prec_w f(\mu(B)) = \mu(f(|B|)).$$

This implies by Proposition 2.4(3) that $\|f(|A|)\| \leq \|f(|B|)\|$ for any symmetric norm.

(ii)$\Rightarrow$(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

$$\mu_i(f(|A|)) = f(\mu_i(A)) = \log(\varepsilon + \mu_i(A)) - \log \varepsilon,$$

$\|f(|A|)\|_{(k)} \leq \|f(|B|)\|_{(k)}$ means that

$$\prod_{i=1}^{k}(\varepsilon + \mu_i(A)) \leq \prod_{i=1}^{k}(\varepsilon + \mu_i(B)).$$

Letting $\varepsilon \downarrow 0$ we get $\prod_{i=1}^{k} \mu_i(A) \leq \prod_{i=1}^{k} \mu_i(B)$ and hence (i) follows.

(iii) $\Leftrightarrow$ (iv) is trivial by definition of $\| \cdot \|_{(k)}$ and (vi)$\Rightarrow$(v)$\Rightarrow$(iv) is clear. Finally assume (iii) and let f be as in (vi). Proposition 1.2 yields (2.5) again, so that (vi) follows. Hence (iii)$\Rightarrow$(vi) holds. ∎

The following is a noncommutative analogue of Fatou's lemma.

PROPOSITION 2.11. *Any symmetric norm $\| \cdot \|$ given by (2.4) is lower-semicontinuous in WOT (i.e. the weak operator topology) on $\mathcal{B}(\mathcal{H})$.*

P r o o f. Let Φ be the symmetric gauge function for $\| \cdot \|$. Using (2.4), Lemma 1.16, and Theorem 2.9, we have for every $A \in \mathcal{B}(\mathcal{H})$,

$\|A\| = \sup\{\|PAP\| : P \text{ is a projection of finite rank}\}$

$\qquad = \sup\{|\mathrm{tr}(XPAP)| : X \in \mathcal{C}_{\Phi'}^{(0)}(\mathcal{H}), \ \|X\|' \leq 1, \ P \text{ is a projection of finite rank}\}.$

Hence the assertion follows because $A \mapsto \mathrm{tr}(XPAP)$ is continuous in WOT whenever P is of finite rank. ∎

Let $\mathcal{B}(\mathcal{H})_+$ denote the set of positive operators in $\mathcal{B}(\mathcal{H})$. The next proposition extending [79, Theorem 2.16] is a noncommutative variant of the dominated convergence theorem.

PROPOSITION 2.12. *Let $A_j, A \in \mathcal{B}(\mathcal{H})$ and $B \in \mathcal{B}(\mathcal{H})_+$. Let Φ be a symmetric gauge function with the corresponding norm $\| \cdot \|$. Assume that $|A_j| \leq B$ and $|A_j^*| \leq B$ for*

all j as well as $|A| \leq B$ and $|A^*| \leq B$. If $A_j \to A$ in WOT and $B \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$, then $\|A_j - A\| \to 0$.

For the proof we need:

LEMMA 2.13. *Let Φ and $\|\cdot\|$ be as above.*

(1) *If $1 < p < \infty$ and $1/p + 1/q = 1$, then*

$$\Phi(a_1 b_1, a_2 b_2, \ldots) \leq \Phi(|a_1|^p, |a_2|^p, \ldots)^{1/p} \Phi(|b_1|^q, |b_2|^q, \ldots)^{1/q}, \quad a, b \in s_{\mathrm{fin}}.$$

(2) *For every $1 < p < \infty$ define*

$$\Phi^{(p)}(a_1, a_2, \ldots) = \Phi(|a_1|^p, |a_2|^p, \ldots)^{1/p}, \quad a \in s_{\mathrm{fin}}.$$

Then $\Phi^{(p)}$ is a symmetric gauge function and the corresponding symmetric norm is $\| \, |\cdot|^p \, \|^{1/p}$.

P r o o f. (1) We may assume that $a, b \geq 0$. From the Young inequality $a_i b_i \leq p^{-1} a_i^p + q^{-1} b_i^q$, Lemma 2.1(1) gives

$$\Phi(a_1 b_1, a_2 b_2, \ldots) \leq \frac{1}{p} \Phi(a_1^p, a_2^p, \ldots) + \frac{1}{q} \Phi(b_1^q, b_2^q, \ldots).$$

Replacing a, b by $\alpha a, \alpha^{-1} b$ with $\alpha > 0$, we get

$$\Phi(a_1 b_1, a_2 b_2, \ldots) \leq \frac{\alpha^p}{p} \Phi(a_1^p, a_2^p, \ldots) + \frac{\alpha^{-q}}{q} \Phi(b_1^q, b_2^q, \ldots).$$

Then the desired inequality follows by minimizing the above right-hand side in $\alpha > 0$.

(2) Let $1/p + 1/q = 1$ and $a, b \in s_{\mathrm{fin}}$. We have

$$\begin{aligned}
\Phi(|a_1 + b_1|^p, |a_2 + b_2|^p, \ldots) &\leq \Phi(|a_1| \, |a_1 + b_1|^{p-1}, |a_2| \, |a_2 + b_2|^{p-1}, \ldots) \\
&\quad + \Phi(|b_1| \, |a_1 + b_1|^{p-1}, |b_2| \, |a_2 + b_2|^{p-1}, \ldots) \\
&\leq \{\Phi(|a_1|^p, |a_2|^p, \ldots)^{1/p} + \Phi(|b_1|^p, |b_2|^p, \ldots)^{1/p}\} \\
&\quad \times \Phi(|a_1 + b_1|^p, |a_2 + b_2|^p, \ldots)^{1/q}
\end{aligned}$$

thanks to (1). Hence $\Phi^{(p)}(a + b) \leq \Phi^{(p)}(a) + \Phi^{(p)}(b)$. The remaining properties of $\Phi^{(p)}$, being a symmetric gauge function as well as the last assertion are obvious. ∎

P r o o f   o f   P r o p o s i t i o n   2.12. For any $\varepsilon > 0$, since $B \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$, there exists a projection P of finite rank such that $\|BP^\perp\| < \varepsilon$ where $P^\perp = I - P$. In view of Lemma 2.13(2) let $\|\cdot\|^{(2)}$ denote the symmetric norm corresponding to $\Phi^{(2)}$. Then we get by assumption

$$\| \, |A_j|^{1/2} P^\perp \|^{(2)} = \|P^\perp |A_j| P^\perp\|^{1/2} \leq \|P^\perp B P^\perp\|^{1/2} < \varepsilon^{1/2}$$

and similarly $\| \, |A_j^*|^{1/2} P^\perp \|^{(2)} < \varepsilon^{1/2}$. Since $\Phi(ab) \leq \Phi^{(2)}(a) \Phi^{(2)}(b)$ for $a, b \in s_{\mathrm{fin}}$ by Lemma 2.13(1), the generalized Hölder inequality (Lemma 2.8) gives

$$\|A_j P^\perp\| = \| \, |A_j| P^\perp \| \leq \| \, |A_j|^{1/2} \|^{(2)} \| \, |A_j|^{1/2} P^\perp \|^{(2)} \leq \|B\|^{1/2} \varepsilon^{1/2},$$

and similarly $\|P^\perp A_j P\| \leq \|A_j^* P^\perp\| \leq \|B\|^{1/2} \varepsilon^{1/2}$. The same inequalities hold also for A. Therefore

$$\begin{aligned}
\|A_j - A\| &\leq \|(A_j - A)P^\perp\| + \|P^\perp(A_j - A)P\| + \|P(A_j - A)P\| \\
&\leq 4\|B\|^{1/2} \varepsilon^{1/2} + \|P(A_j - A)P\|.
\end{aligned}$$

Since P is of finite rank and $A_j \to A$ in WOT, it immediately follows that $\|P(A_j - A)P\| \to 0$. Since ε is arbitrary, we have $\|A_j - A\| \to 0$. ∎

For instance, if $A_j, B \in \mathcal{B}(\mathcal{H})_+$, $A_j \leq B \in \mathcal{C}(\mathcal{H})$, and $A_j \to A \in \mathcal{B}(\mathcal{H})$ in WOT (hence $A \leq B$), then $\|A_j - A\|_\infty \to 0$. As is remarked in [79, p. 39], the condition $|A_j^*| \leq B$ cannot be dropped in the above proposition.

2.4. *Ando's extension of Birman–Koplienko–Solomyak majorization result.* The final result of this section is taken from [3]. (See [42, Appendix] by H. Kosaki for its extension to the von Neumann algebra case.) Before stating the theorem, let us recall the notion of operator monotone functions. A continuous real function f on $[0, \infty)$ is said to be *operator monotone* if $A \leq B$ implies $f(A) \leq f(B)$ for any $A, B \in \mathcal{B}(\mathcal{H})_+$. Such a function is characterized by the following integral representation:

$$(2.6) \qquad f(x) = a + bx + \int_0^\infty \frac{xt}{x+t} d\nu(t), \qquad x \geq 0,$$

where $a \in \mathbb{R}$, $b \geq 0$, and ν is a positive measure on $(0, \infty)$ with $\int_0^\infty t(1+t)^{-1} d\nu(t) < \infty$. This type of integral representation is a central result in the Löwner theory of operator monotone functions (see [1, 24, 36] for details). The operator monotonicity of x^θ on $[0, \infty)$ with $0 < \theta < 1$ is due to the Löwner–Heinz inequality. Also it is well known that $\log(1+x)$ is operator monotone.

THEOREM 2.14. *Let $A, B \in \mathcal{B}(\mathcal{H})_+$ and $\|\cdot\|$ be any symmetric norm.*

(1) *If f is an operator monotone function on $[0, \infty)$ with $f(0) \geq 0$, then*

$$(2.7) \qquad \mu(f(A) - f(B)) \prec_w \mu(f(|A - B|)).$$

(2) *If g is a continuous strictly increasing function on $[0, \infty)$ such that $g(0) = 0$ and $\lim_{t \to \infty} g(t) = \infty$ and whose inverse function is operator monotone, then*

$$\mu(g(|A - B|)) \prec_w \mu(g(A) - g(B)).$$

When $f(x) = x^\theta$ with $0 < \theta < 1$, the weak majorization (2.7) above was formerly proved by Birman, Koplienko, and Solomyak [18], which gives the generalized Powers–Størmer inquality: If $0 < \theta < 1$ and $\theta \leq p \leq \infty$, then

$$(2.8) \qquad \|A^\theta - B^\theta\|_{p/\theta} \leq \|A - B\|_p^\theta, \qquad A, B \in \mathcal{B}(\mathcal{H})_+.$$

The case when $\theta = 1/2$ and $p = 1$ is known as the Powers–Størmer inequality [72], whose extension to the von Neumann algebra setting was obtained in [9, 34].

We first prepare simple facts to prove the theorem.

LEMMA 2.15. *Let $X, Y \in \mathcal{B}(\mathcal{H})$ be self-adjoint and $X = X_+ - X_-$, $Y = Y_+ - Y_-$ be the Jordan decompositions.*

(1) *If $X \leq Y$ then $\mu_i(X_+) \leq \mu_i(Y_+)$ for all i.*
(2) *If $\mu(X_+) \prec_w \mu(Y_+)$ and $\mu(X_-) \prec_w \mu(Y_-)$, then $\mu(X) \prec_w \mu(Y)$.*

Proof. (1) Let Q be the support projection of X_+. Since

$$X_+ = QXQ \leq QYQ \leq QY_+Q,$$

we have $\mu_i(X_+) \leq \mu_i(QY_+Q) \leq \mu_i(Y_+)$ by Proposition 1.4.

(2) It is rather easy to see that $\mu(X)$ is the decreasing rearrangement of the combination of $\mu(X_+)$ and $\mu(X_-)$. So for each $k \in \mathbb{N}$ we can choose $0 \le m \le k$ so that

$$\sum_{i=1}^{k} \mu_i(X) = \sum_{i=1}^{m} \mu_i(X_+) + \sum_{i=1}^{k-m} \mu_i(X_-).$$

Hence

$$\sum_{i=1}^{k} \mu_i(X) \le \sum_{i=1}^{m} \mu_i(Y_+) + \sum_{i=1}^{k-m} \mu_i(Y_-) \le \sum_{i=1}^{k} \mu_i(Y),$$

as desired. ∎

Proof of Theorem 2.14. (1) First assume that $A \ge B \ge 0$ and let $C = A - B \ge 0$. In view of Proposition 2.10 it suffices to prove that

(2.9) $$\|f(B+C) - f(B)\|_{(k)} \le \|f(C)\|_{(k)}, \quad k \in \mathbb{N}.$$

For $t > 0$ let

$$h_t(x) = \frac{x}{x+t} = 1 - \frac{t}{x+t},$$

which is increasing on $[0, \infty)$ with $h_t(0) = 0$. According to the integral representation (2.6) with $a, b \ge 0$, we have

$$\mu_i(f(C)) = f(\mu_i(C)) = a + b\mu_i(C) + \int_0^\infty \frac{\mu_i(C)t}{\mu_i(C)+t} d\nu(t) = a + b\mu_i(C) + \int_0^\infty t\mu_i(h_t(C)) d\nu(t),$$

so that

(2.10) $$\|f(C)\|_{(k)} \ge b\|C\|_{(k)} + \int_0^\infty t\|h_t(C)\|_{(k)} d\nu(t).$$

On the other hand, since

$$f(B+C) = aI + b(B+C) + \int_0^\infty th_t(B+C) d\nu(t)$$

as well as the analogous expression for $f(B)$, we have

$$f(B+C) - f(B) = bC + \int_0^\infty t\{h_t(B+C) - h_t(B)\} d\nu(t),$$

so that

(2.11) $$\|f(B+C) - f(B)\|_{(k)} \le b\|C\|_{(k)} + \int_0^\infty t\|h_t(B+C) - h_t(B)\|_{(k)} d\nu(t).$$

By (2.10) and (2.11) it suffices for (2.9) to show that

$$\|h_t(B+C) - h_t(B)\|_{(k)} \le \|h_t(C)\|_{(k)}, \quad t > 0, \ k \in \mathbb{N}.$$

Since $h_t(x) = h_1(x/t)$, the case $t = 1$ is enough because we may replace B and C by $t^{-1}B$ and $t^{-1}C$, respectively. So what remains to prove is the following:

(2.12) $$\|(B+I)^{-1} - (B+C+I)^{-1}\|_{(k)} \le \|I - (C+I)^{-1}\|_{(k)}, \quad k \in \mathbb{N}.$$

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 $\|(B + I)^{-1/2}\|_\infty \leq 1$, we obtain

$$\begin{aligned}
\mu_i((B + I)^{-1} - (B + C + I)^{-1}) &\leq \mu_i(h_1((B + I)^{-1/2} C(B + I)^{-1/2})) \\
&= h_1(\mu_i((B + I)^{-1/2} C(B + I)^{-1/2})) \\
&\leq h_1(\mu_i(C)) = \mu_i(I - (C + I)^{-1})
\end{aligned}$$

by repeated use of Proposition 1.4(7). Therefore (2.12) is proved.

Next let us prove the general case $A, B \geq 0$. Since $0 \leq A \leq B + (A - B)_+$, it follows that

$$f(A) - f(B) \leq f(B + (A - B)_+) - f(B),$$

which implies by Lemma 2.15(1) that

$$\|(f(A) - f(B))_+\|_{(k)} \leq \|f(B + (A - B)_+) - f(B)\|_{(k)}.$$

Applying the first case to $B + (A - B)_+$ and B, we get

$$\|f(B + (A - B)_+) - f(B)\|_{(k)} \leq \|f((A - B)_+)\|_{(k)}.$$

Therefore

(2.13) $$\mu((f(A) - f(B))_+) \prec_w \mu(f((A - B)_+)).$$

Exchanging the role of A, B gives

(2.14) $$\mu((f(A) - f(B))_-) \prec_w \mu(f((A - B)_-)).$$

Here we may assume $f(0) = 0$ because f can be replaced by $f - f(0)$. Then it is immediate that $f((A - B)_+)f((A - B)_-) = 0$ and $f((A - B)_+) + f((A - B)_-) = f(|A - B|)$. So $\mu(f(A) - f(B)) \prec_w \mu(f(|A - B|))$ follows from (2.13) and (2.14) thanks to Lemma 2.15(2).

(2) Let f be the inverse of g. Since f satisfies the condition of (1), we have

$$\sum_{i=1}^k \mu_i(f(A) - f(B)) \leq \sum_{i=1}^k f(\mu_i(A - B)), \quad k \in \mathbb{N}.$$

Here replace A and B by $g(A)$ and $g(B)$, respectively. Then

$$\sum_{i=1}^k \mu_i(A - B) \leq \sum_{i=1}^k f(\mu_i(g(A) - g(B))), \quad k \in \mathbb{N},$$

which means that $\mu(A - B) \prec_w f(\mu(g(A) - g(B)))$. As is well known, the operator monotonicity of f implies the concavity of f, so that g is convex. Hence $\mu(g(|A - B|)) = g(\mu(A - B)) \prec_w \mu(g(A) - g(B))$ by Proposition 1.2. ∎

PROBLEM 2.16. Let $A, B \in \mathcal{B}(\mathcal{H})_+$ and f be an operator monotone function on $[0, \infty)$ with $f(0) = 0$. Then it is natural to ask whether or not the following variant of (2.7) holds:

$$\mu(f(A + B)) \prec_w \mu(f(A) + f(B)).$$

But it seems that the method in proving (2.7) does not work in this case. Theorem 2.14(1) and (1.12) show that

$$\mu(f(A+B)) \prec_w \mu(f(A)) + \mu(f(B)),$$

while a stronger result is found in [4, Theorem 6.9].

3. Inequalites of Golden–Thompson type. This section is mostly taken from [40]. We obtain log-majorization results and norm inequalities of Golden–Thompson type for exponential operators. Furthermore, we discuss the norm convergence of exponential product formula, which is the main result of [40]. But accounts more detailed than those in [40] are supplied concerning technical parts: Lemmas 3.1, 3.2, 3.8, and theory of state perturbation in von Neumann algebras.

3.1. *Antisymmetric tensor powers.* First let us establish a machinery of antisymmetric tensors, which is quite useful in deriving log-majorization results. Let $\mathcal{H}$ be a separable Hilbert space as before. For each $n \in \mathbb{N}$ let $\otimes^n \mathcal{H}$ denote the n-fold tensor product of $\mathcal{H}$ with itself, which is the completed Hilbert space of the n-fold algebraic tensor product with respect to the inner product defined by

$$\langle \xi_1 \otimes \ldots \otimes \xi_n, \eta_1 \otimes \ldots \otimes \eta_n \rangle = \prod_{i=1}^n \langle \xi_i, \eta_i \rangle.$$

For $\xi_1, \ldots, \xi_n \in \mathcal{H}$ define $\xi_1 \wedge \ldots \wedge \xi_n \in \otimes^n \mathcal{H}$ by

$$(3.1) \qquad \xi_1 \wedge \ldots \wedge \xi_n = \frac{1}{\sqrt{n!}} \sum_\pi (\text{sign}\,\pi) \xi_{\pi(1)} \otimes \ldots \otimes \xi_{\pi(n)},$$

where π runs over all permutations on $\{1, \ldots, n\}$ and $\text{sign}\,\pi = \pm 1$ according as π is even or odd. The closed subspace of $\otimes^n \mathcal{H}$ spanned by $\{\xi_1 \wedge \ldots \wedge \xi_n : \xi_i \in \mathcal{H}\}$ is called the n-fold *antisymmetric tensor product* of $\mathcal{H}$ and denoted by $\Lambda^n \mathcal{H}$. In fact, the linear extension of the map $\xi_1 \otimes \ldots \otimes \xi_n \mapsto \frac{1}{\sqrt{n!}} \xi_1 \wedge \ldots \wedge \xi_n$ is the projection of $\otimes^n \mathcal{H}$ onto $\Lambda^n \mathcal{H}$. A straightforward computation from (3.1) shows that

$$(3.2) \qquad \langle \xi_1 \wedge \ldots \wedge \xi_n, \eta_1 \wedge \ldots \wedge \eta_n \rangle = \det[\langle \xi_i, \eta_j \rangle]_{1 \le i,j \le n}.$$

Note that $\xi_1 \wedge \ldots \wedge \xi_n \ne 0$ if and only if $\{\xi_1, \ldots, \xi_n\}$ is linearly independent. Moreover, if $\{\varphi_i\}$ is an orthonormal basis of $\mathcal{H}$, then $\{\varphi_{i_1} \wedge \ldots \wedge \varphi_{i_n} : i_1 < \ldots < i_n\}$ is an orthonormal basis of $\Lambda^n \mathcal{H}$.

For every $A \in \mathcal{B}(\mathcal{H})$ the n-fold tensor product $\otimes^n A \in \mathcal{B}(\otimes^n \mathcal{H})$ is given as

$$(\otimes^n A)(\xi_1 \otimes \ldots \otimes \xi_n) = A\xi_1 \otimes \ldots \otimes A\xi_n.$$

Since $\Lambda^n \mathcal{H}$ is invariant for $\otimes^n A$, the *antisymmetric tensor power* $\Lambda^n A$ of A can be defined as $\Lambda^n A = \otimes^n A|_{\Lambda^n \mathcal{H}}$; in fact,

$$(3.3) \qquad (\Lambda^n A)(\xi_1 \wedge \ldots \wedge \xi_n) = A\xi_1 \wedge \ldots \wedge A\xi_n.$$

When $\dim \mathcal{H} = N < \infty$, $\Lambda^N \mathcal{H} = \mathbb{C}$, $\Lambda^N A = \det A$, and $\Lambda^n \mathcal{H} = \{0\}$ for $n > N$. In the matrix theory $\Lambda^n A$ is usually called the nth *compound* of A.

The following are elementary properties of antisymmetric tensor powers.

LEMMA 3.1. *Let $X, X_j, Y, A \in \mathcal{B}(\mathcal{H})$ and $n \in \mathbb{N}$.*

(1) $\Lambda^n(X^*) = (\Lambda^n X)^*$.

(2) $\Lambda^n(XY) = (\Lambda^n X)(\Lambda^n Y)$.

(3) *If $\|X_j - X\|_\infty \to 0$, then $\|\Lambda^n X_j - \Lambda^n X\|_\infty \to 0$.*

(4) *If $X_j \to X$ in SOT (resp. WOT) and $\sup_j \|X_j\|_\infty < \infty$, then $\Lambda^n X_j \to \Lambda^n X$ in SOT (resp. WOT).*

(5) *If $A \geq 0$, then $\Lambda^n A \geq 0$ and $\Lambda^n(A^p) = (\Lambda^n A)^p$ for all $p > 0$.*

(6) $\Lambda^n(|X|) = |\Lambda^n X|$.

Proof. (1) and (2) are the restrictions of the corresponding formulas $\otimes^n(X^*) = (\otimes^n X)^*$ and $\otimes^n(XY) = (\otimes^n X)(\otimes^n Y)$ to $\Lambda^n \mathcal{H}$. For (3) and (4) it suffices to show the corresponding convergences for $\otimes^n$, which are readily verified. If $A \geq 0$ then $\Lambda^n A = (\Lambda^n(A^{1/2}))^*(\Lambda^n(A^{1/2})) \geq 0$ by (1) and (2). When p is rational, the second assertion of (5) is immediate from (2). Then (3) implies the case of general $p > 0$. Finally (6) follows from (1), (2), and (5). ∎

A useful technique in the majorization theory for operators is supplied by the following lemma. For matrices this is a consequence of the so-called Binet–Cauchy theorem ([62, pp. 503–504]).

LEMMA 3.2. *For every $A \in \mathcal{B}(\mathcal{H})$ and $n \in \mathbb{N}$,*

$$(3.4) \qquad \prod_{i=1}^n \mu_i(A) = \mu_1(\Lambda^n A) \ (= \|\Lambda^n A\|_\infty).$$

Proof. We may assume by Lemma 3.1(6) that $A \geq 0$. First suppose that A is compact. Then there exists an orthonormal basis $\{\varphi_i\}$ of $\mathcal{H}$ such that $A\varphi_i = \mu_i(A)\varphi_i$ for all i. Since $\{\varphi_{i_1} \wedge \ldots \wedge \varphi_{i_n} : i_1 < \ldots < i_n\}$ is a complete set of eigenvectors of $\Lambda^n A$ with

$$(\Lambda^n A)(\varphi_{i_1} \wedge \ldots \wedge \varphi_{i_n}) = \Big\{ \prod_{j=1}^n \mu_{i_j}(A) \Big\} \varphi_{i_1} \wedge \ldots \wedge \varphi_{i_n},$$

we have (3.4). Next let $A \geq 0$ be general. Let P_k be projections of finite rank such that $P_k \uparrow I$. Since $P_k A P_k \to A$ in SOT, Lemma 3.1(4) implies that $\Lambda^n(P_k A P_k) \to \Lambda^n A$ in SOT. Since $\| \cdot \|_\infty$ is lower-semicontinuous in WOT, we have

$$\|\Lambda^n A\|_\infty \leq \liminf_{k \to \infty} \|\Lambda^n(P_k A P_k)\|_\infty = \liminf_{k \to \infty} \prod_{i=1}^n \mu_i(P_k A P_k) \leq \prod_{i=1}^n \mu_i(A)$$

by the first case and Proposition 1.4(7). On the other hand, by Lemma 1.16 we can choose sequences of orthonormal sets $\{\xi_1^{(k)}, \ldots, \xi_n^{(k)}\}$ such that $\lim_k \|A\xi_i^{(k)} - \mu_i(A)\xi_i^{(k)}\| = 0$ for $1 \leq i \leq n$. Hence it follows from (3.2) and (3.3) that

$$\|\Lambda^n A\|_\infty \geq \langle (\Lambda^n A)(\xi_1^{(k)} \wedge \ldots \wedge \xi_n^{(k)}), \xi_1^{(k)} \wedge \ldots \wedge \xi_n^{(k)} \rangle = \det[\langle A\xi_i^{(k)}, \xi_j^{(k)} \rangle]_{1 \leq i,j \leq n} \to \prod_{i=1}^n \mu_i(A)$$

as $k \to \infty$, completing the proof. ∎

Before going into the main part of this section, let us prove the Weyl majorization theorem as warming-up practice in the antisymmetric tensor technique.

THEOREM 3.3. *Let $A \in \mathcal{C}(\mathcal{H})$ and $\lambda_1(A), \lambda_2(A), \ldots$ be the eigenvalues of A arranged as $|\lambda_1(A)| \geq |\lambda_2(A)| \geq \ldots$ with algebraic multiplicities counted. Then*

$$\prod_{i=1}^{n} |\lambda_i(A)| \leq \prod_{i=1}^{n} \mu_i(A), \quad n \in \mathbb{N}.$$

P r o o f. If λ is an eigenvalue of A with algebraic multiplicity m, then there exists a set $\{\eta_1, \ldots, \eta_m\}$ of independent vectors such that

$$A\eta_j - \lambda\eta_j \in \text{span}\{\eta_1, \ldots, \eta_{j-1}\}, \quad 1 \leq j \leq m.$$

Hence for each $n \in \mathbb{N}$ we can choose independent vectors $\xi_1, \ldots, \xi_n$ such that $A\xi_i = \lambda_i(A)\xi_i + \zeta_i$ with $\zeta_i \in \text{span}\{\xi_1, \ldots, \xi_{i-1}\}$ for $1 \leq i \leq n$. Then it is readily checked that

$$(\Lambda^n A)(\xi_1 \wedge \ldots \wedge \xi_n) = A\xi_1 \wedge \ldots \wedge A\xi_n = \left\{\prod_{i=1}^{n} \lambda_i(A)\right\} \xi_1 \wedge \ldots \wedge \xi_n$$

and $\xi_1 \wedge \ldots \wedge \xi_n \neq 0$, so that $\prod_{i=1}^{n} \lambda_i(A)$ is an eigenvalue of $\Lambda^n A$. Hence Lemma 3.2 gives

$$\left| \prod_{i=1}^{n} \lambda_i(A) \right| \leq \|\Lambda^n A\|_\infty = \prod_{i=1}^{n} \mu_i(A). \quad \blacksquare$$

3.2. *Araki's log-majorization result.* Extending a trace inequality of Lieb and Thirring [60], Araki [12] showed the following log-majorization result by essentially the same method as below. Also, the same result for matrices was shown in [88] in the same way. Note that this was further generalized in [55] to the von Neumann algebra case, while the antisymmetric tensor technique can be no longer used. A generalization in another direction will be given in Proposition 5.8.

THEOREM 3.4. *For every $A, B \in \mathcal{B}(\mathcal{H})_+$,*

(3.5) $\mu((A^{1/2}BA^{1/2})^r) \prec_{w(\log)} \mu(A^{r/2}B^r A^{r/2}), \quad r \geq 1,$

or equivalently

(3.6) $\mu((A^{p/2}B^p A^{p/2})^{1/p}) \prec_{w(\log)} \mu((A^{q/2}B^q A^{q/2})^{1/q}), \quad 0 < p \leq q.$

P r o o f. We can pass to the limit from $A + \varepsilon I$ and $B + \varepsilon I$ as $\varepsilon \downarrow 0$ by Proposition 1.4(10). So we may assume that A and B are invertible. First let us show that

(3.7) $\|(A^{1/2}BA^{1/2})^r\|_\infty \leq \|A^{r/2}B^r A^{r/2}\|_\infty, \quad r \geq 1.$

To do so, it suffices to show that $A^{r/2}B^r A^{r/2} \leq I$ implies $A^{1/2}BA^{1/2} \leq I$, equivalently $B^r \leq A^{-r}$ implies $B \leq A^{-1}$. But this is just the Löwner–Heinz inequality. For every $n \in \mathbb{N}$, since by Lemma 3.1,

$$\Lambda^n((A^{1/2}BA^{1/2})^r) = ((\Lambda^n A)^{1/2}(\Lambda^n B)(\Lambda^n A)^{1/2})^r,$$

$$\Lambda^n(A^{r/2}B^r A^{r/2}) = (\Lambda^n A)^{r/2}(\Lambda^n B)^r (\Lambda^n A)^{r/2},$$

it follows from (3.7) with $\Lambda^n A, \Lambda^n B$ instead of A, B that

$$\|\Lambda^n((A^{1/2}BA^{1/2})^r)\|_\infty \leq \|\Lambda^n(A^{r/2}B^r A^{r/2})\|_\infty.$$

This means thanks to Lemma 3.2 that

$$\prod_{i=1}^{n} \mu_i((A^{1/2}BA^{1/2})^r) \le \prod_{i=1}^{n} \mu_i(A^{r/2}B^rA^{r/2}).$$

Hence (3.5) is proved. If we replace A, B by A^p, B^p and take $r = q/p$, then

$$\mu((A^{p/2}B^pA^{p/2})^{q/p}) \prec_{w(\log)} \mu(A^{q/2}B^qA^{q/2}),$$

which implies (3.6) by Proposition 1.4(11). ∎

Theorem 3.4 and Proposition 2.10 give:

COROLLARY 3.5. *Let $A, B \in \mathcal{B}(\mathcal{H})_+$ and $\|\cdot\|$ be any symmetric norm. If f is a continuous increasing function on $[0, \infty)$ such that $f(0) \ge 0$ and $f(e^x)$ is convex, then*

$$\|f((A^{1/2}BA^{1/2})^r)\| \le \|f(A^{r/2}B^rA^{r/2})\|, \quad r \ge 1.$$

In particular,

$$\|(A^{1/2}BA^{1/2})^r\| \le \|A^{r/2}B^rA^{r/2}\|, \quad r \ge 1.$$

3.3. *Trotter–Kato exponential product formula.* To obtain norm inequalities of Golden–Thompson type, let us recall the form sum of positive self-adjoint operators. Let H and K be positive self-adjoint (not necessarily bounded) operators on $\mathcal{H}$. Let $\mathcal{D}_0 = \mathcal{D}(H^{1/2}) \cap \mathcal{D}(K^{1/2})$, $\mathcal{D}(H^{1/2})$ being the domain of $H^{1/2}$, $\mathcal{H}_0$ the closure of $\mathcal{D}_0$, and P_0 the projection onto $\mathcal{H}_0$. We denote by $H \hat{+} K$ the *form sum* of H and K, which is the positive self-adjoint operator on $\mathcal{H}_0$ associated with the closed positive quadratic form $\xi \in \mathcal{D}_0 \mapsto \|H^{1/2}\xi\|^2 + \|K^{1/2}\xi\|^2$ so that

$$\|(H\hat{+}K)^{1/2}\xi\|^2 = \|H^{1/2}\xi\|^2 + \|K^{1/2}\xi\|^2, \quad \xi \in \mathcal{D}_0.$$

The form sum $H\hat{+}K$ can be defined also for lower-bounded self-adjoint H, K as follows:

$$H\hat{+}K = (H_+\hat{+}K_+) - P_0(H_- + K_-)P_0,$$

where $H = H_+ - H_-$ is the Jordan decomposition; or equivalently

$$H\hat{+}K = ((H + aI)\hat{+}(K + bI)) - (a + b)P_0$$

by taking $a, b \in \mathbb{R}$ such that $H + aI, K + bI \ge 0$. Note that if $H + K$ is essentially self-adjoint, then $H\hat{+}K$ coincides with the closure of $H + K$.

The following (1) is the so-called Trotter–Kato product formula, which was shown in [51] in a more general form (see also [19]). The symmetric and continuous parameter version (2) was given in [40]. It is a simpler fact that if $X, Y \in \mathcal{B}(\mathcal{H})$ then $\lim_{n\to\infty}(e^{X/n}e^{Y/n})^n = e^{X+Y}$ in the norm $\|\cdot\|_\infty$. But a quite general Trotter-like convergence for exponential products of bounded operators was established in [82].

THEOREM 3.6. (1) *If H and K are lower-bounded self-adjoint operators on $\mathcal{H}$, then*

$$\operatorname*{s-lim}_{n\to\infty}(e^{-tH/n}e^{-tK/n})^n = e^{-t(H\hat{+}K)}, \quad t > 0,$$

the convergence being uniform in $t \in [a, b]$ for any $0 < a < b$. Here s-lim *means the convergence in SOT and $e^{-t(H\hat{+}K)}$ is written for $e^{-t(H\hat{+}K)}P_0$ for brevity.*

(2) *With the same assumption and notation,*

$$\text{s-}\lim_{r\downarrow 0}(e^{-rtH/2}e^{-rtK}e^{-rtH/2})^{1/r} = e^{-t(H\hat{+}K)}, \quad t > 0,$$

the convergence being uniform in $t \in [a, b]$ for any $0 < a < b$.

P r o o f. Let us only show (2), which is an easy consequence of (1). We may assume $H, K \geq 0$ by taking $H + aI, K + aI$. For $0 < r < 1$ write $1/r = n + s$ where $n = n(r) \in \mathbb{N}$ and $s = s(r) \in [0, 1)$. Then

$$(e^{-rtH/2}e^{-rtK}e^{-rtH/2})^{1/r}$$
$$= (e^{-rtH/2}e^{-rtK}e^{-rtH/2})^s e^{-rtH/2}e^{-rtK}(e^{-rtH}e^{-rtK})^{n-1}e^{-rtH/2}.$$

When $r \downarrow 0$ (hence $n \to \infty$), we have $e^{-rtH/2} \to I$ and $e^{-rtK} \to I$ in SOT uniformly in $t \in [0, b]$. Since

$$e^{-rtH/2}e^{-rtK}e^{-rtH/2} \leq (e^{-rtH/2}e^{-rtK}e^{-rtH/2})^s \leq I,$$

we get

$$(3.8) \qquad\qquad \text{s-}\lim_{r\downarrow 0}(e^{-rtH/2}e^{-rtK}e^{-rtH/2})^s = I$$

uniformly in $t \in [0, b]$. Since $r(n-1)t \to t$, the uniform convergence of (1) implies that

$$(3.9) \qquad\qquad \text{s-}\lim_{r\downarrow 0}\{(e^{-rtH}e^{-rtK})^{n-1} - e^{-r(n-1)t(H\hat{+}K)}\} = 0$$

uniformly in $t \in [a, b]$. Finally, it is immediate that

$$(3.10) \qquad\qquad \text{s-}\lim_{r\downarrow 0}\{e^{-r(n-1)t(H\hat{+}K)} - e^{-t(H\hat{+}K)}\} = 0$$

uniformly in $t \in [0, b]$. The above uniform convergences (3.8)–(3.10) altogether yield the conclusion. ∎

3.4. *Log-majorization and norm inequalities of Golden–Thompson type.* The next theorem is the Golden–Thompson inequality strengthened to the form of log-majorization.

THEOREM 3.7. *If H and K are lower-bounded self-adjoint operator on $\mathcal{H}$, then*

$$\mu(e^{-(H\hat{+}K)}) \prec_{w(\log)} \mu((e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}), \quad r > 0.$$

To prove the theorem, we give the following infinite-dimensional extension of Ky Fan's multiplicative formula.

LEMMA 3.8. *For every $A \in \mathcal{B}(\mathcal{H})$ and $n \in \mathbb{N}$,*

$$\prod_{i=1}^{n}\mu_i(A) = \sup\{\det(P|A|P|_{P\mathcal{H}}) : P \text{ is a projection of rank } n\}$$
$$= \sup\{\operatorname{Re}\det(PUAP|_{P\mathcal{H}}) : U \text{ is a unitary and } P \text{ is a projection of rank } n\}.$$

Hence the function $A \mapsto \prod_{i=1}^{n}\mu_i(A)$ is lower-semicontinuous in WOT on $\mathcal{B}(\mathcal{H})$.

P r o o f. If P is a projection of rank n and U is a unitary, then

$$\det(P|A|P|_{P\mathcal{H}}) = \prod_{i=1}^{n}\mu_i(P|A|P) \leq \prod_{i=1}^{n}\mu_i(A),$$

$$\operatorname{Re}\det(PUAP|_{P\mathcal{H}}) \le |\det(PUAP|_{P\mathcal{H}})| = \prod_{i=1}^{n} \mu_i(PUAP) \le \prod_{i=1}^{n} \mu_i(A).$$

To show the converse, we may assume that $\mu_n(A) > 0$. By Lemma 1.16 (and its proof), we can choose sequences of orthonormal sets $\{\xi_1^{(k)}, \ldots, \xi_n^{(k)}\}$ in the range of $|A|$ such that $\lim_k \||A|\xi_i^{(k)} - \mu_i(A)\xi_i^{(k)}\| = 0$ for $1 \le i \le n$. Let P_k be the projection onto the span of $\{\xi_1^{(k)}, \ldots, \xi_n^{(k)}\}$ and $A = W|A|$ the polar decomposition. For each k, since $\{W\xi_1^{(k)}, \ldots, W\xi_n^{(k)}\}$ is orthonormal, there exists a unitary U_k such that $U_k \xi_i^{(k)} = W\xi_i^{(k)}$ for $1 \le i \le n$. Then

$$\det(P_k|A|P_k|_{P_k\mathcal{H}}) = \det[\langle |A|\xi_i^{(k)}, \xi_j^{(k)}\rangle]_{i,j},$$

$$\det(P_k U_k^* A P_k|_{P_k\mathcal{H}}) = \det[\langle A\xi_i^{(k)}, U_k \xi_j^{(k)}\rangle] = \det[\langle A\xi_i^{(k)}, W\xi_j^{(k)}\rangle] = \det[\langle |A|\xi_i^{(k)}, \xi_j^{(k)}\rangle].$$

Since $\det[\langle |A|\xi_i^{(k)}, \xi_j^{(k)}\rangle]_{ij} \to \prod_{i=1}^{n} \mu_i(A)$ as $k \to \infty$, we get the required formulas. Furthermore, the last assertion is immediate because $A \mapsto \operatorname{Re}\det(PUAP|_{P\mathcal{H}})$ is continuous in WOT whenever P is of finite rank. ∎

By the way, the following additive formula can be shown in a similar way: For every $A \in \mathcal{B}(\mathcal{H})$ and $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} \mu_i(A) = \sup\{\operatorname{tr}(P|A|P) : P \text{ is a projection of rank } n\}$$

$$= \sup\{\operatorname{Re}\operatorname{tr}(PUAP) : U \text{ is a unitary and } P \text{ is a projection of rank } n\}$$

and hence $\sum_{i=1}^{n} \mu_i(A)$ is lower-semicontinuous in WOT. However, we do not know an explicit counter-example to $\mu_n(A)$ itself being lower-semicontinuous in WOT.

Indeed, Lemmas 3.1(4) and 3.2 serve for the proof of Theorem 3.7. But the above multiplicative and additive formulas are worth pointing out by themselves.

Proof of Theorem 3.7. For every $n \in \mathbb{N}$ we have

$$\prod_{i=1}^{n} \mu_i(e^{-(H \hat{+} K)}) \le \liminf_{r \downarrow 0} \prod_{i=1}^{n} \mu_i((e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r})$$

by Theorem 3.6(2) and Lemma 3.8. Since $\prod_{i=1}^{n} \mu_i((e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r})$ decreases as $r \downarrow 0$ by (3.6), we conclude that

$$\prod_{i=1}^{n} \mu_i(e^{-(H \hat{+} K)}) \le \prod_{i=1}^{n} \mu_i((e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}), \quad r > 0,$$

as desired. ∎

COROLLARY 3.9. *If H and K are lower-bounded self-adjoint operators on $\mathcal{H}$ and $\|\cdot\|$ is a symmetric norm, then*

$$\|e^{-(H \hat{+} K)}\| \le \|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|, \quad r > 0,$$

and the above right-hand side decreases as $r \downarrow 0$. In particular,

(3.11) $$\|e^{-(H \hat{+} K)}\| \le \|e^{-H/2}e^{-K}e^{-H/2}\| \le \|e^{-H}e^{-K}\|.$$

P r o o f. The first assertion is a consequence of Theorem 3.7, (3.6), and Proposition 2.10. The second inequality of (3.11) follows because by Proposition 1.4(5)

$$\|e^{-H}e^{-K}\| = \| \, |e^{-K}e^{-H}| \, \| = \|(e^{-H}e^{-2K}e^{-H})^{1/2}\|. \quad \blacksquare$$

The specialization of (3.11) to the trace norm $\|\cdot\|_1$ is the celebrated Golden–Thompson trace inequality independently established in [33, 83, 84]. It was shown in [81] that $\operatorname{tr} e^{H+K} \le \operatorname{tr}(e^{H/n}e^{K/n})^n$ for every Hermitian matrices H, K and $n \in \mathbb{N}$. For the matrix case, (3.11) was given in [58, 85]. Also (3.11) for the norm $\|\cdot\|_\infty$ is known as Segal's inequality ([73, p. 260]). A rather trivial consequence of (3.11) is that $\|e^{-(H\hat{+}K)}\| \le \|e^{-H}\|_\infty \|e^{-K}\|$ for any symmetric norm $\|\cdot\|$. Hence, if Φ is a symmetric gauge function and $e^{-K} \in \mathcal{C}_\Phi(\mathcal{H})$, then $e^{-(H\hat{+}K)} \in \mathcal{C}_\Phi(\mathcal{H})$ for every lower-bounded H.

Concerning (quasi-)norms $\|\cdot\|_p$ we have:

COROLLARY 3.10. (1) *If H, K are as above and $0 < p \le \infty$, then*

$$\|e^{-(H\hat{+}K)}\|_p \le \|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|_p, \quad r > 0,$$

and the above the right-hand side decreases as $r \downarrow 0$.

(2) *When $0 < p, p_1, p_2 \le \infty$ and $1/p = 1/p_1 + 1/p_2$, if $e^{-H} \in \mathcal{C}_{p_1}(\mathcal{H})$ and $e^{-K} \in \mathcal{C}_{p_2}(\mathcal{H})$, then $e^{-(H\hat{+}K)} \in \mathcal{C}_p(\mathcal{H})$.*

P r o o f. (1) follows from Corollary 3.9 because for any $0 < p < \infty$,

$$\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|_p = \|(e^{-rH/2}e^{-rK}e^{-rH/2})^{p/r}\|_1^{1/p}$$

and $\|e^{-(H\hat{+}K)}\|_p = \|e^{-p(H\hat{+}K)}\|_1^{1/p}$. (2) is immediate from (3.11) and the Hölder inequality (Proposition 2.7). $\blacksquare$

The next example is given to show that $\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|$ does not converge to $\|e^{-(H\hat{+}K)}\|$ as $r \downarrow 0$ in general.

EXAMPLE 3.11. Define 2×2 matrices P_0 and Q_n, $n \in \mathbb{N}$, by

$$P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q_n = \xi_n \otimes \xi_n \quad \text{with} \quad \xi_n = \begin{bmatrix} \sqrt{1 - 1/n} \\ \sqrt{1/n} \end{bmatrix}.$$

Then P_0 and Q_n are projections with $P_0 \wedge Q_n = 0$, so that $P_0^\perp + Q_n^\perp$ is invertible. Let $0 < \delta < 1$ be arbitrary. For each n, since $\exp\{(\log \varepsilon)(P_0^\perp + Q_n^\perp)\} \to 0$ as $\varepsilon \downarrow 0$, we can choose $0 < \varepsilon_n < 1$ such that

$$\|\exp\{(\log \varepsilon_n)(P_0^\perp + Q_n^\perp)\}\|_\infty \le \delta.$$

Putting $H_n = (-\log \varepsilon_n)P_0^\perp$ and $K_n = (-\log \varepsilon_n)Q_n^\perp$, we define $H = \bigoplus_{n=1}^\infty H_n$ and $K = \bigoplus_{n=1}^\infty K_n$, which are positive self-adjoint operators on $\mathcal{H} = \bigoplus_1^\infty \mathbb{C}^2$. Then

$$\begin{aligned} e^{-rH/2}e^{-rK}e^{-rH/2} &= \bigoplus_{n=1}^\infty (e^{-rH_n/2}e^{-rK_n}e^{-rH_n/2}) \\ &= \bigoplus_{n=1}^\infty \{(P_0 + \varepsilon_n P_0^\perp)^{r/2}(Q_n + \varepsilon_n Q_n^\perp)^r (P_0 + \varepsilon_n P_0^\perp)^{r/2}\}. \end{aligned}$$

Since

$$(P_0 + \varepsilon_n P_0^\perp)^{r/2}(Q_n + \varepsilon_n Q_n^\perp)^r (P_0 + \varepsilon_n P_0^\perp)^{r/2}$$
$$\geq (P_0 + \varepsilon_n P_0^\perp)^{r/2} Q_n (P_0 + \varepsilon_n P_0^\perp)^{r/2} = (P_0 + \varepsilon_n P_0^\perp)^{r/2}\xi_n \otimes (P_0 + \varepsilon_n P_0^\perp)^{r/2}\xi_n,$$

we have

$$\|e^{-rH/2}e^{-rK}e^{-rH/2}\|_\infty \geq \sup_n \|(P_0 + \varepsilon_n P_0^\perp)^{r/2}\xi_n\|^2 = \sup_n \left(1 - \frac{1}{n} + \frac{\varepsilon_n^r}{n}\right) \geq 1,$$

so that $\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}\|_\infty \geq 1$ for all $r > 0$. On the other hand,

$$e^{-(H \hat{+} K)} = \bigoplus_{n=1}^\infty e^{-(H_n + K_n)} = \bigoplus_{n=1}^\infty \exp\{(\log \varepsilon_n)(P_0^\perp + Q_n^\perp)\}$$

and hence $\|e^{-(H \hat{+} K)}\|_\infty \leq \delta$.

3.5. *Trace norm convergence of exponential product formula.* In this subsection let H and K be lower-bounded self-adjoint operators such that $H + K$ is essentially self-adjoint. For simplicity we denote the closure of $H + K$ by the same $H + K$.

The following trace norm convergence of exponential product formula was established in [40].

THEOREM 3.12. *If $e^{-K} \in \mathcal{C}_1(\mathcal{H})$, then*

$$(3.12) \qquad \lim_{r \downarrow 0} \|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r} - e^{-(H+K)}\|_1 = 0,$$

$$(3.13) \qquad \lim_{r \downarrow 0} \|(e^{-rK/2}e^{-rH}e^{-rK/2})^{1/r} - e^{-(H+K)}\|_1 = 0,$$

$$(3.14) \qquad \lim_{n \to \infty} \|(e^{-H/n}e^{-K/n})^{n+1} - e^{-(H+K)}\|_1 = 0,$$

$$(3.15) \qquad \lim_{n \to \infty} \operatorname{tr}(e^{-H/n}e^{-K/n})^n = \operatorname{tr} e^{-(H+K)}.$$

First note that all the operators in the theorem belong to $\mathcal{C}_1(\mathcal{H})$. For instance, we have $(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r} \in \mathcal{C}_1(\mathcal{H})$ for all $r > 0$, because with $H \geq aI$,

$$\mu_k((e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r}) = \mu_k(e^{-rK/2}e^{-rH}e^{-rK/2})^{1/r}$$
$$\leq \mu_k(e^{-ra}e^{-rK})^{1/r} = e^{-a}\mu_k(e^{-K}).$$

The proof of the theorem in [40] is based on Araki's Trotter product formula [8] involving the modular operator and Donald's perturbation theory [23] for positive normal functionals on von Neumann algebras. So, for convenience, we first give a brief survey on these materials.

Let $\mathcal{M}$ be a von Neumann algebra and $(\mathcal{M}, \mathcal{H}, J, \mathcal{P})$ be a *standard form* ([9, 34]) of $\mathcal{M}$. For example, a standard form of $\mathcal{M} = \mathcal{B}(\mathcal{H})$ is given as $(\mathcal{B}(\mathcal{H}), \mathcal{C}_2(\mathcal{H}), {}^*, \mathcal{C}_2(\mathcal{H})_+)$, where $\mathcal{B}(\mathcal{H})$ is represented on $\mathcal{C}_2(\mathcal{H})$ by left multiplication and $\mathcal{C}_2(\mathcal{H})_+$ is the set of positive operators in $\mathcal{C}_2(\mathcal{H})$. Let $\mathcal{M}_*^+$ denote the positive part of the predual $\mathcal{M}_*$, and $\varphi \in \mathcal{M}_*^+$ be faithful. Then φ is written as $\varphi = \langle \cdot \Phi, \Phi \rangle$ with the cyclic and separating vector $\Phi \in \mathcal{P}$. Let Δ_Φ be the *modular operator* associated with Φ (or φ). See [80] for modular theory

of von Neumann algebras. For each $h \in \mathcal{M}_{\mathrm{sa}}$ where $\mathcal{M}_{\mathrm{sa}}$ is the self-adjoint part of $\mathcal{M}$, Araki [7] defined the *perturbed vector* Φ^h by

$$(3.16) \qquad \Phi^h = \sum_{n=0}^{\infty} (-1)^n \int_0^{1/2} dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{n-1}} dt_n \Delta_\Phi^{t_n} h \Delta_\Phi^{t_{n-1}-t_n} h \ldots \Delta_\Phi^{t_1-t_2} h \Phi,$$

where Φ is in the domain of $\Delta_\Phi^{z_1} h \Delta_\Phi^{z_2} h \ldots \Delta_\Phi^{z_n} h$ for $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$ with $\mathrm{Re}\, z \in \{(s_1, \ldots, s_n) \in \mathbb{R}^n : s_1, \ldots, s_n \geq 0, \ s_1 + \ldots + s_n \leq 1/2\}$ and the sum in (3.16) absolutely converges. Then Φ^h is also a cyclic and separating vector in $\mathcal{P}$. It is known [8] that Φ is in the domain of $\exp \frac{1}{2}(\log \Delta_\Phi - h)$ and

$$\Phi^h = \exp\left(\frac{\log \Delta_\Phi - h}{2}\right)\Phi.$$

Moreover, the *perturbed functional* φ^h is defined by

$$(3.17) \qquad\qquad\qquad \varphi^h = \langle \cdot \, \Phi^h, \Phi^h \rangle.$$

Note that Φ^h and φ^h here mean Φ^{-h} and φ^{-h} in the notation of [7, 8].

The following is a version of the Trotter product formula given in [8, Remarks 1, 2].

PROPOSITION 3.13. *Let Φ be as above. Then for every $h \in \mathcal{M}_{\mathrm{sa}}$,*

$$\Phi^h = \lim_{n \to \infty} (\Delta_\Phi^{1/2n} e^{-h/2n})^n \Phi = \lim_{n \to \infty} (e^{-h/2n} \Delta_\Phi^{1/2n})^n \Phi \quad \text{strongly.}$$

The relative entropy for positive normal functionals on a von Neumann algebra was first introduced in [87] in the semifinite case and was extended in [10, 11] to the general case. For each $\varphi, \psi \in \mathcal{M}_*^+$ let Φ and Ψ be the vector representatives in $\mathcal{P}$ so that $\varphi = \langle \cdot \, \Phi, \Phi \rangle$ and $\psi = \langle \cdot \, \Psi, \Psi \rangle$. Then Araki's *relative entropy* $S(\psi, \varphi)$ is defined by

$$S(\psi, \varphi) = -\langle (\log \Delta_{\Phi, \Psi}) \Psi, \Psi \rangle$$

if the support of ψ is dominated by that of φ, and $S(\psi, \varphi) = \infty$ otherwise, where $\Delta_{\Phi, \Psi}$ is the relative modular operator ([11]). See [86] for another definition and [54, 68] for properties of relative entropy.

Next let h be a lower-bounded self-adjoint operator affiliated with $\mathcal{M}$. So we have the spectral decomposition $h = \int_a^\infty \lambda de_h(\lambda)$, where $\{e_h(\lambda)\}_{\lambda \geq a}$ is a spectral resolution in $\mathcal{M}$ with $a \in \mathbb{R}$. For any $\psi \in \mathcal{M}_*^+$, $\psi(h)$ can be defined as $\psi(h) = \int_a^\infty \lambda d\psi(e_h(\lambda))$. For each $\varphi \in \mathcal{M}_*^+$ define $c(\varphi, h) \in [-\infty, \infty)$ by

$$c(\varphi, h) = \sup\{-\psi(h) - S(\psi, \varphi) : \psi \in \Sigma_*(\mathcal{M})\},$$

where $\Sigma_*(\mathcal{M})$ denotes the set of normal states on $\mathcal{M}$. According to [23, Theorem 3.1], if $c(\varphi, h) > -\infty$, that is, $\psi(h) + S(\psi, \varphi) < \infty$ for some $\psi \in \Sigma_*(\mathcal{M})$, then there exists a unique $\omega \in \Sigma_*(\mathcal{M})$ such that $-\omega(h) - S(\omega, \varphi) = c(\varphi, h)$. This ω is denoted by $[\varphi^h]$, while it was denoted in [23] by φ^h. More generally, h can be an extended-valued lower-bounded operator (see [23]), but a self-adjoint h is enough for our purpose.

Note [70, Proposition 1] that when $\varphi \in \mathcal{M}_*^+$ is faithful and $h \in \mathcal{M}_{\mathrm{sa}}$, $[\varphi^h]$ coincides with (3.17) up to a normalization constant; more precisely $[\varphi^h] = \varphi^h/\varphi^h(I)$ and $\varphi^h(I) = e^{c(\varphi, h)}$. So for any $\varphi \in \mathcal{M}_*^+$ and any lower-bounded self-adjoint operator h affiliated with

$\mathcal{M}$ such that $c(\varphi, h) > -\infty$, we can define the *perturbed functional* φ^h by

$$(3.18) \qquad\qquad\qquad \varphi^h = e^{c(\varphi,h)}[\varphi^h].$$

Then h is called a *relative Hamiltonian* of $\omega = \varphi^h$ relative to φ.

The following convergence result was proved in [23, Proposition 3.15], which plays a crucial role in the proof of Theorem 3.12.

PROPOSITION 3.14. *Let h_n and h be lower-bounded self-adjoint operators affiliated with $\mathcal{M}$ such that $h_n \uparrow h$ in the sense that $\psi(h_n) \uparrow \psi(h)$ for all $\psi \in \mathcal{M}_*^+$. If $\varphi \in \mathcal{M}_*^+$ satisfies $c(\varphi, h) > -\infty$, then $c(\varphi, h_n) \to c(\varphi, h)$, $\|[\varphi^{h_n}] - [\varphi^h]\| \to 0$, and hence $\|\varphi^{h_n} - \varphi^h\| \to 0$.*

Now, under the above preparations, let us sketch the proof of Theorem 3.12.

P r o o f   o f   T h e o r e m   3.12 (Sketch). The von Neumann algebra $\mathcal{M} = \mathcal{B}(\mathcal{H})$ is represented by left multiplication on the standard Hilbert space $\mathcal{C}_2(\mathcal{H})$. Set $\Phi = e^{-K/2} \in \mathcal{C}_2(\mathcal{H})_+$, which is a cyclic and separating vector for $\mathcal{M}$ and defines a faithful $\varphi \in \mathcal{M}_*^+$ as $\varphi = \langle \cdot \Phi, \Phi \rangle = \mathrm{tr}(\cdot e^{-K})$. We write $\varphi = e^{-K}$ under the usual identification $\mathcal{B}(\mathcal{H})_* = \mathcal{C}_1(\mathcal{H})$. Note that

$$(3.19) \qquad\qquad \Delta_\Phi^{it} X \Delta_\Phi^{-it} = e^{-itK} X e^{itK}, \qquad t \in \mathbb{R}, \ X \in \mathcal{B}(\mathcal{H}),$$

because both sides represent the modular automorphism group associated with Φ.

We may assume that $H, K \geq 0$. First assume that H is bounded. An argument of analytic continuation using (3.19) yields

$$(3.20) \qquad\qquad (e^{-H/2n} \Delta_\Phi^{1/2n})^n \Phi = (e^{-H/2n} e^{-K/2n})^n, \qquad n \in \mathbb{N}.$$

Since $(e^{-H/2n} e^{-K/2n})^n \to e^{-(H+K)/2}$ in SOT as $n \to \infty$, Proposition 3.13 and (3.20) show that $\Phi^H = e^{-(H+K)/2}$ and

$$\lim_{n \to \infty} \|(e^{-H/2n} e^{-K/2n})^n - e^{-(H+K)/2}\|_2 = 0,$$

which yields

$$\lim_{n \to \infty} \mathrm{tr}(e^{-H/4n} e^{-K/2n} e^{-H/4n})^{2n} = \mathrm{tr}\, e^{-(H+K)}.$$

This together with Corollary 3.10(1) implies that

$$(3.21) \qquad\qquad \lim_{r \downarrow 0} \mathrm{tr}(e^{-rH/2} e^{-rK} e^{-rH/2})^{1/r} = \mathrm{tr}\, e^{-(H+K)}.$$

Next let $H = \int_a^\infty \lambda\, dE_H(\lambda)$ be lower-bounded and $H_n = \int_a^n \lambda\, dE_H(\lambda)$. Since $c(\varphi, H) > -\infty$ follows from $\mathcal{D}(H) \cap \mathcal{D}(K) \neq \{0\}$, we can define the perturbed functional φ^H by (3.18) as well as φ^{H_n}. Then Proposition 3.14 shows that $\|\varphi^{H_n} - \varphi^H\| \to 0$. We have $\varphi^{H_n} = e^{-(H_n+K)}$ because $\Phi^{H_n} = e^{-(H_n+K)/2}$ for bounded H_n. Furthermore, $(I + H_n + K)^{-1} \downarrow (I + H + K)^{-1}$ and hence $e^{-(H_n+K)} \to e^{-(H+K)}$ in SOT (see the proofs of Lemmas 4.15 and 4.8(1) below). Therefore $\varphi^H = e^{-(H+K)}$ and

$$(3.22) \qquad\qquad \lim_{n \to \infty} \|e^{-(H_n+K)} - e^{-(H+K)}\|_1 = 0.$$

By (3.21) for bounded H_n and by (3.22) we get

$$\lim_{r \downarrow 0} \mathrm{tr}(e^{-rH/2} e^{-rK} e^{-rH/2})^{1/r} \leq \lim_{r \downarrow 0} \mathrm{tr}(e^{-rH_n/2} e^{-rK} e^{-rH_n/2})^{1/r}$$

$$= \mathrm{tr}\, e^{-(H_n+K)} \to \mathrm{tr}\, e^{-(H+K)},$$

which implies (3.21) for lower-bounded H. Hence

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/2r}\|_2 = \|e^{-(H+K)/2}\|_2.$$

Since it is easily seen from Theorem 3.6(2) that $(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/2r} \to e^{-(H+K)/2}$ weakly in $\mathcal{C}_2(\mathcal{H})$, we have

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/2r} - e^{-(H+K)/2}\|_2 = 0,$$

which shows (3.12) by using the Hölder inequality. The proof of (3.13) is similar, and (3.14) is easily shown from (3.12) thanks to

$$(3.23) \qquad (e^{-H/n}e^{-K/n})^{n+1} = e^{-H/2n}(e^{-H/2n}e^{-K/n}e^{-H/2n})^n e^{-H/2n}e^{-K/n}.$$

Finally (3.15) is immediate from (3.12). ∎

COROLLARY 3.15. *If $e^{-K} \in \mathcal{C}_p(\mathcal{H})$ where $0 < p < \infty$, then*

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r} - e^{-(H+K)}\|_p = 0,$$

$$\lim_{r\downarrow 0}\|(e^{-rK/2}e^{-rH}e^{-rK/2})^{1/r} - e^{-(H+K)}\|_p = 0,$$

$$(3.24) \qquad \lim_{n\to\infty}\|(e^{-H/n}e^{-K/n})^{n+1} - e^{-(H+K)}\|_p = 0.$$

Furthermore, for every $q > p$,

$$\lim_{n\to\infty}\|(e^{-H/n}e^{-K/n})^n - e^{-(H+K)}\|_q = 0.$$

Proof (Sketch). Choose $k \in \mathbb{N}$ such that $2^k > 1/p$. Applying (3.12) to pH, pK instead of H, K we have

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{p/r} - e^{-p(H+K)}\|_1 = 0,$$

so that by (2.8) with $\theta = 1/2^k p$

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/2^k r} - e^{-(H+K)/2^k}\|_{2^k p} = 0.$$

Now, using the Hölder inequality repeatedly, we can obtain the first assertion. (A similar argument in more detail will be supplied in the proof of Corollary 4.23.) The second is similar and the third is shown by using (3.23). The last assertion is also shown by using (2.8) and the Hölder inequality (see [40] for details). ∎

PROBLEMS 3.16. (1) Is it possible to replace $(e^{-H/n}e^{-K/n})^{n+1}$ by $(e^{-H/n}e^{-K/n})^n$ in (3.14) and (3.24)?

(2) In view of Corollary 3.15, we are tempted to conjecture that if $e^{-K} \in \mathcal{C}_\Phi^{(0)}(\mathcal{H})$ then

$$\lim_{r\downarrow 0}\|(e^{-rH/2}e^{-rK}e^{-rH/2})^{1/r} - e^{-(H+K)}\| = 0,$$

where Φ is a symmetric gauge function with the corresponding norm $\|\cdot\|$.

4. Inequalites of complementary Golden–Thompson type. A log-majorization result in terms of power operator means was obtained in [6] in the matrix case, which gives rise to matrix norm inequalities considered as complementary to the Golden–Thompson

ones. The aim of this section is to study log-majorizations and norm inequalities involving operator means for infinite-dimensional exponential operators. In the course of this study, we establish a Trotter-like exponential product formula for operator means, which may be interesting by itself. Most results in this section are new.

4.1. *Preliminaries on operator means.* First, for convenience, let us give a brief survey on operator means. An axiomatic approach for operator means was investigated by Kubo and Ando [57]. Assume that $\mathcal{H}$ is a separable and infinite-dimensional Hilbert space. A binary operation $\sigma : \mathcal{B}(\mathcal{H})_+ \times \mathcal{B}(\mathcal{H})_+ \to \mathcal{B}(\mathcal{H})_+$ is called an *operator connection* if it satisfies the following conditions (i)–(iii) for $A, B, C, D \in \mathcal{B}(\mathcal{H})_+$:

(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) $A_n, B_n \in \mathcal{B}(\mathcal{H})_+$, $A_n \downarrow A$, and $B_n \downarrow B$ imply $A_n \, \sigma \, B_n \downarrow A \, \sigma \, B$ (upper semicontinuity).

An operator connection σ is called an *operator mean* if
(iv) $I \, \sigma \, I = I$.

The fundamental theorem of Kubo and Ando is summarized as follows: For each operator connection σ there exists a unique operator monotone function $f \geq 0$ on $[0, \infty)$ such that $f(x)I = I \, \sigma \, (xI)$ for $x \geq 0$. Then the map $\sigma \mapsto f$ is an affine order-isomorphism between the operator connections and the non-negative operator monotone functions on $[0, \infty)$. The order preservation means that when $\sigma_i \mapsto f_i$ for $i = 1, 2$, $A \, \sigma_1 \, B \leq A \, \sigma_2 \, B$ for all $A, B \in \mathcal{B}(\mathcal{H})_+$ if and only if $f_1(x) \leq f_2(x)$ for all $x \geq 0$. The operator connection σ is defined via the corresponding function f by

(4.1) $$A \, \sigma \, B = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}$$

if A is invertible, and for general A, B

$$A \, \sigma \, B = \text{s-}\lim_{\varepsilon \downarrow 0} A_\varepsilon^{1/2} f(A_\varepsilon^{-1/2} B_\varepsilon A_\varepsilon^{-1/2}) A_\varepsilon^{1/2} \quad \text{decreasingly,}$$

where $A_\varepsilon = A + \varepsilon I$ and $B_\varepsilon = B + \varepsilon I$. Moreover σ is an operator mean if and only if $f(1) = 1$, which implies that $A \, \sigma \, A = A$ for all A.

The following are typical examples of operator means.

(1) *Arithmetic mean*: $A \nabla B = \frac{1}{2}(A + B)$.
(2) *Harmonic mean*: $A \, ! \, B = 2(A : B)$, where $A : B = \text{s-}\lim_{\varepsilon \downarrow 0}(A_\varepsilon^{-1} + B_\varepsilon^{-1})^{-1}$ is called the *parallel sum* of A, B.
(3) *Geometric mean*: $A \, \# \, B = \text{s-}\lim_{\varepsilon \downarrow 0} A_\varepsilon^{1/2}(A_\varepsilon^{-1/2} B_\varepsilon A_\varepsilon^{-1/2})^{1/2} A_\varepsilon^{1/2}$.

The corresponding functions of $\nabla, !, \#$ are $(1 + x)/2, 2x/(x + 1), x^{1/2}$, respectively. For $0 \leq \alpha \leq 1$ let $\#_\alpha$ denote the α-*power mean*, which is the operator mean corresponding to the operator monotone function x^α. Namely, for each $A, B \in \mathcal{B}(\mathcal{H})_+$ 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 adopt the usual convention $B^0 = I$ for any $B \in \mathcal{B}(\mathcal{H})_+$. Note that $A \, \#_0 \, B = A$, $A \, \#_1 \, B = B$, and $A \, \#_{1/2} \, B = A \, \# \, B$.

Using the integral representation (2.6) of f and transforming the measure ν to $m = t(1+t)^{-1}\nu$, we can represent $A\,\sigma\,B$ as

$$A\,\sigma\,B = aA + bB + \int\limits_{(0,\infty)} \frac{1+t}{t}\{(tA) : B\}dm(t) = \int\limits_{[0,\infty]} \frac{1+t}{t}\{(tA) : B\}dm(t),$$

where $a = m(\{0\})$ and $b = m(\{\infty\})$. In this way, it is seen that there exists an affine one-to-one correspondence between the operator means and the probability measures on $[0,\infty]$. The above representation shows that every operator connection σ satisfies

(v) $(A\,\sigma\,B) + (C\,\sigma\,D) \le (A+C)\,\sigma\,(B+D)$,

(vi) $T^*(A\,\sigma\,B)T \le (T^*AT)\,\sigma\,(T^*BT)$ for every $T \in \mathcal{B}(\mathcal{H})$ and the equality holds if T is invertible.

Now let σ be an operator mean with the corresponding function f. The integral representation of f implies that f is automatically infinitely many times differentiable. So let $\alpha = f'(1)$. The concavity of f implies that $0 \le \alpha \le 1$ and $f(x) \le (1-\alpha) + \alpha x$ for $x \ge 0$. Since $f(x^{-1})^{-1}$ is again operator monotone and hence concave, we have

(4.2) $\dfrac{x}{(1-\alpha)x + \alpha} \le f(x) \le (1-\alpha) + \alpha x, \quad x \ge 0.$

The operator mean σ' determined by $A\,\sigma'\,B = B\,\sigma\,A$ is called the *transpose* of σ. The corresponding function of σ' is $xf(x^{-1})$ because

$$I\,\sigma'\,(xI) = (xI)\,\sigma\,I = x(I\,\sigma\,(x^{-1}I)) = xf(x^{-1})I, \quad x > 0.$$

In particular, if σ is *symmetric*, i.e. $A\,\sigma\,B = B\,\sigma\,A$ for all A, B, or equivalently $f(x) = xf(x^{-1})$, then $\alpha = 1/2$ and so (4.2) says that the arithmetic mean $A\,\nabla\,B$ is the greatest and the harmonic mean $A\,!\,B$ is the least among the symmetric operator means. If $\alpha = 0$, then $f(x) \equiv 1$ and $A\,\sigma\,B = A$. Also if $\alpha = 1$, then $f(x) = x$ and $A\,\sigma\,B = B$. All the results in this section become trivial in these two extremal cases. So we may assume $0 < \alpha < 1$ in the sequel discussions.

Although the next lemma will be used only for α-power mean $\#_\alpha$, we give it for general σ.

LEMMA 4.1. *Let σ be an operator mean and f the corresponding operator monotone function. Let $A, B \in \mathcal{B}(\mathcal{H})_+$.*

(1) *Assume that A is invertible. Then $B \mapsto A\,\sigma\,B$ is continuous in the norm $\|\cdot\|_\infty$ on $\mathcal{B}(\mathcal{H})_+$.*

(2) *Assume that $f(0) = 0$ and B is compact. If $A_n \downarrow A$ then $\|A_n\,\sigma\,B - A\,\sigma\,B\|_\infty \to 0$.*

P r o o f. (1) is immediate from the expression (4.1) because of the $\|\cdot\|_\infty$-continuity of functional calculus.

(2) Choose $a > 0$ such that $A_n \le aI$. Then we have $A_n\,\sigma\,B \le (aI)\,\sigma\,B = af(a^{-1}B)$ and $af(a^{-1}B) \in \mathcal{C}(\mathcal{H})$ from $f(0) = 0$. Hence Proposition 2.12 shows the assertion because $A_n\,\sigma\,B \downarrow A\,\sigma\,B$. ■

4.2. *Log-majorization for power operator means.* The log-majorization in the next theorem was given in [6] in the matrix case, which is considered as complementary to Theorem 3.4.

THEOREM 4.2. *If $A, B \in \mathcal{B}(\mathcal{H})_+$ and A is either invertible or compact, then*

$$(4.3) \qquad \mu(A^r \#_\alpha B^r) \prec_{w(\log)} \mu((A \#_\alpha B)^r), \qquad r \geq 1,$$

or equivalently

$$(4.4) \qquad \mu((A^p \#_\alpha B^p)^{1/p}) \prec_{w(\log)} \mu((A^q \#_\alpha B^q)^{1/q}), \qquad p \geq q > 0.$$

P r o o f. First assume that both A and B are invertible. The proof below is the same as that of [6, Theorem 2.1]. For each $n \in \mathbb{N}$ it is easily checked from Lemma 3.1 that

$$\Lambda^n(A^r \#_\alpha B^r) = (\Lambda^n A)^r \#_\alpha (\Lambda^n B)^r,$$
$$\Lambda^n((A \#_\alpha B)^r) = ((\Lambda^n A) \#_\alpha (\Lambda^n B))^r.$$

So it suffices to show that

$$(4.5) \qquad \|A^r \#_\alpha B^r\|_\infty \leq \|(A \#_\alpha B)^r\|_\infty, \qquad r \geq 1,$$

because (4.3) follows from Lemma 3.2 by taking $\Lambda^n A, \Lambda^n B$ instead of A, B in (4.5). To show (4.5), we may prove that $A \#_\alpha B \leq I$ implies $A^r \#_\alpha B^r \leq I$. When $1 \leq r \leq 2$, let us write $r = 2 - \varepsilon$ with $0 \leq \varepsilon \leq 1$. Let $C = A^{-1/2}BA^{-1/2}$. Suppose $A \#_\alpha B \leq I$. Then $C^\alpha \leq A^{-1}$ and

$$(4.6) \qquad A \leq C^{-\alpha},$$

so that thanks to $0 \leq \varepsilon \leq 1$,

$$(4.7) \qquad A^{1-\varepsilon} \leq C^{-\alpha(1-\varepsilon)}.$$

Now we have

$$A^r \#_\alpha B^r = A^{1-\frac{\varepsilon}{2}}\{A^{-1+\frac{\varepsilon}{2}}B \cdot B^{-\varepsilon} \cdot BA^{-1+\frac{\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}}$$
$$= A^{1-\frac{\varepsilon}{2}}\{A^{-\frac{1-\varepsilon}{2}}CA^{1/2}(A^{-1/2}C^{-1}A^{-1/2})^\varepsilon A^{1/2}CA^{-\frac{1-\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}}$$
$$= A^{1/2}\{A^{1-\varepsilon} \#_\alpha [C(A \#_\varepsilon C^{-1})C]\}A^{1/2}$$
$$\leq A^{1/2}\{C^{-\alpha(1-\varepsilon)} \#_\alpha [C(C^{-\alpha} \#_\varepsilon C^{-1})C]\}A^{1/2}$$

by using (4.6), (4.7), and the joint monotonicity of power means. 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 get

$$(4.8) \qquad A^r \#_\alpha B^r \leq A^{1/2}C^\alpha A^{1/2} = A \#_\alpha B \leq I.$$

Therefore (4.3) is proved when $1 \leq r \leq 2$. When $r > 2$, write $r = 2^k s$ with $k \in \mathbb{N}$ and $1 \leq s \leq 2$. Repeating the above argument we have

$$\mu(A^r \#_\alpha B^r) \prec_{w(\log)} \mu(A^{2^{k-1}s} \#_\alpha B^{2^{k-1}s})^2$$
$$\vdots$$
$$\prec_{w(\log)} \mu(A^s \#_\alpha B^s)^{2^k}$$
$$\prec_{w(\log)} \mu(A \#_\alpha B)^r.$$

Second assume that A is invertible while B is general. Since by Lemma 4.1(1)

$$A^r \#_\alpha B^r = \lim_{\varepsilon\downarrow 0} A^r \#_\alpha B_\varepsilon^r \quad \text{and} \quad (A \#_\alpha B)^r = \lim_{\varepsilon\downarrow 0}(A \#_\alpha B_\varepsilon)^r$$

in the norm $\|\cdot\|_\infty$, (4.3) follows from the first case and Proposition 1.4(10). Third assume that A is compact. The second case shows that

$$\mu(B_\varepsilon^r \#_{1-\alpha} A^r) \prec_{w(\log)} \mu((B_\varepsilon \#_{1-\alpha} A)^r), \quad r \geq 1, \ \varepsilon > 0.$$

Since Lemma 4.1(2) implies that $B_\varepsilon \#_{1-\alpha} A \to B \#_{1-\alpha} A$ and $B_\varepsilon^r \#_{1-\alpha} A^r \to B^r \#_{1-\alpha} A^r$ in $\|\cdot\|_\infty$, we get

$$\mu(B^r \#_{1-\alpha} A^r) \prec_{w(\log)} \mu((B \#_{1-\alpha} A)^r), \quad r \geq 1,$$

which is nothing but (4.3). Finally (4.4) readily follows from (4.3) as in the last part of proof of Theorem 3.4. ∎

By Theorem 4.2 and Proposition 2.10 we have:

COROLLARY 4.3. *Let $A, B \in \mathcal{B}(\mathcal{H})_+$ and assume that A is either invertible or compact. Let $\|\cdot\|$ be any symmetric norm. If f is a continuous increasing function on $[0, \infty)$ such that $f(0) \geq 0$ and $f(e^x)$ is convex, then*

$$\|f(A^r \#_\alpha B^r)\| \leq \|f((A \#_\alpha B)^r)\|, \quad r \geq 1.$$

In particular,

$$\|A^r \#_\alpha B^r\| \leq \|(A \#_\alpha B)^r\|, \quad r \geq 1.$$

PROBLEM 4.4. It is desirable to prove Theorem 4.2 (hence Corollary 4.3) without the assumption of A being invertible or compact. We can do so if the convergence $\|A \#_\alpha B\|_\infty = \lim_{\varepsilon \downarrow 0} \|(A + \varepsilon I) \#_\alpha B\|_\infty$ holds true for general $A, B \in \mathcal{B}(\mathcal{H})_+$. However, this is not the case as the following example shows.

EXAMPLE 4.5. Let us consider the infinite direct sum of 2×2 matrices. Let P_0, ξ_n, and Q_n be as in Example 3.11. Define $P = \bigoplus_{n=1}^\infty P_0$ and $Q = \bigoplus_{n=1}^\infty Q_n$, so that

$$P \# Q = \bigoplus_{n=1}^\infty (P_0 \# Q_n),$$

$$(P + \varepsilon I) \# Q = \bigoplus_{n=1}^\infty ((P_0 + \varepsilon I_2) \# Q_n) \quad \text{where} \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We have

$$\begin{aligned}
&(P_0 + \varepsilon I_2) \# Q_n \\
&= (P_0 + \varepsilon I_2)^{1/2} \{(P_0 + \varepsilon I_2)^{-1/2} \xi_n \otimes (P_0 + \varepsilon I_2)^{-1/2} \xi_n\}^{1/2} (P_0 + \varepsilon I_2)^{1/2} \\
&= (P_0 + \varepsilon I_2)^{1/2} \Big\{ \|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|^2 \\
&\qquad \frac{(P_0 + \varepsilon I_2)^{-1/2} \xi_n}{\|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|} \otimes \frac{(P_0 + \varepsilon I_2)^{-1/2} \xi_n}{\|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|} \Big\}^{1/2} (P_0 + \varepsilon I_2)^{1/2} \\
&= \|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\| \frac{\xi_n}{\|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|} \otimes \frac{\xi_n}{\|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|} \\
&= \|(P_0 + \varepsilon I_2)^{-1/2} \xi_n\|^{-1} \xi_n \otimes \xi_n.
\end{aligned}$$

Since

$$(P_0 + \varepsilon I_2)^{-1/2} \xi_n = \begin{bmatrix} \sqrt{\frac{n-1}{n(1+\varepsilon)}} \\ \sqrt{\frac{1}{n\varepsilon}} \end{bmatrix},$$

we get

$$\|(P_0 + \varepsilon I_2) \# Q_n\|_\infty = \left(\frac{n-1}{n(1+\varepsilon)} + \frac{1}{n\varepsilon} \right)^{-1/2}.$$

Letting $\varepsilon \downarrow 0$ yields $\|P_0 \# Q_n\|_\infty = 0$ (this follows also from $P_0 \wedge Q_n = 0$; see [57, Theorem 3.7]). Hence $P \# Q = 0$. On the other hand, we have

$$\|(P + \varepsilon I) \# Q\|_\infty = \sup_{n \geq 1} \|(P_0 + \varepsilon I_2) \# Q_n\|_\infty = \sup_{n \geq 1} \left(\frac{n-1}{n(1+\varepsilon)} + \frac{1}{n\varepsilon} \right)^{-1/2} = (1+\varepsilon)^{1/2}.$$

Therefore

$$\|P \# Q\|_\infty = 0 < 1 = \lim_{\varepsilon \downarrow 0} \|(P + \varepsilon I) \# Q\|_\infty.$$

Furthermore, for each $0 < \delta < 1$ and $n \in \mathbb{N}$, we can choose $\alpha_n > 0$ such that

$$\|(P_0 + \alpha_n I_2) \# (Q_n + \alpha_n I_2)\|_\infty \leq \delta.$$

Define $A = \bigoplus_{n=1}^\infty (P_0 + \alpha_n I_2)$ and $B = \bigoplus_{n=1}^\infty (Q_n + \alpha_n I_2)$, which are strictly positive. Since $P \leq A$ and $Q \leq B$, we have

$$\|A \# B\|_\infty \leq \delta < 1 \leq \lim_{\varepsilon \downarrow 0} \|(A + \varepsilon I) \# B\|_\infty.$$

The next proposition extends [2, Theorem 1] where H, K are bounded and $\alpha = 1/2$.

PROPOSITION 4.6. *If $H \in \mathcal{B}(\mathcal{H})$ is self-adjoint and K is a lower-bounded self-adjoint operator on $\mathcal{H}$, then the following conditions are equivalent*:

(i) $(1 - \alpha)H + \alpha K \geq 0$;
(ii) $e^{-tH} \#_\alpha e^{-tK} \leq I$ *for all* $t \geq 0$;
(iii) $t \mapsto e^{-tH} \#_\alpha e^{-tK}$ *is a decreasing function from* $[0, \infty)$ *to* $\mathcal{B}(\mathcal{H})_+$.

P r o o f. (i)$\Rightarrow$(ii). Suppose that $(1-\alpha)H + \alpha K \geq 0$, i.e. $K \geq -\alpha^{-1}(1-\alpha)H$. To prove (ii), it suffices to show that

$$(4.9) \qquad \lim_{r \downarrow 0} \|(e^{-rH} \#_\alpha e^{-rK})^{1/r}\|_\infty \leq 1,$$

because $\|e^{-rH} \#_\alpha e^{-rK}\|_\infty^{1/r} = \|(e^{-rH} \#_\alpha e^{-rK})^{1/r}\|_\infty$ increases as $r \downarrow 0$ by (4.4). When $0 < r < \alpha\{(1-\alpha)\|H\|_\infty\}^{-1}$, since $I + rK \geq I - \alpha^{-1}(1-\alpha)rH \geq 0$ and $I - \alpha^{-1}(1-\alpha)rH$ is invertible, we get

$$e^{-rK} \leq (I + rK)^{-1} \leq \left(I - \frac{1-\alpha}{\alpha} rH \right)^{-1}$$

as well as $e^{-rH} \leq (I + rH)^{-1}$ for small $r > 0$. Therefore

$$e^{-rH} \#_\alpha e^{-rK} \leq (I + rH)^{-1} \#_\alpha \left(I - \frac{1-\alpha}{\alpha} rH \right)^{-1} = (I + rH)^{-(1-\alpha)} \left(I - \frac{1-\alpha}{\alpha} rH \right)^{-\alpha},$$

so that for small $r > 0$

$$\|(e^{-rH} \#_\alpha e^{-rK})^{1/r}\|_\infty \le \left\|(I + rH)^{-(1-\alpha)/r}\left(I - \frac{1-\alpha}{\alpha}rH\right)^{-\alpha/r}\right\|_\infty.$$

Since

$$\lim_{r\downarrow 0}(1+r\lambda)^{-(1-\alpha)/r}\left(1 - \frac{1-\alpha}{\alpha}r\lambda\right)^{-\alpha/r} = (e^\lambda)^{-(1-\alpha)}(e^{-\frac{1-\alpha}{\alpha}\lambda})^{-\alpha} = 1$$

uniformly in $\lambda \in [-\|H\|_\infty, \|H\|_\infty]$, it follows that

$$\lim_{r\downarrow 0}\left\|(I + rH)^{-(1-\alpha)/r}\left(I - \frac{1-\alpha}{\alpha}rH\right)^{-\alpha/r} - I\right\|_\infty = 0,$$

implying (4.9).

(ii)$\Rightarrow$(iii). What we have to prove is that if $A, B \in \mathcal{B}(\mathcal{H})_+$ and A is invertible, then $A \#_\alpha B \le I$ implies $A^r \#_\alpha B^r \le A \#_\alpha B$ for every $r \ge 1$. When both A and B are invertible, the result was shown in the proof of Theorem 4.2 (see (4.8)). When A is invertible but B is not, let $B_\varepsilon = B + \varepsilon I$ for $\varepsilon > 0$. Since

$$A \#_\alpha (\|A \#_\alpha B_\varepsilon\|_\infty^{-1/\alpha} B_\varepsilon) = \|A \#_\alpha B_\varepsilon\|_\infty^{-1}(A \#_\alpha B_\varepsilon) \le I,$$

the first case applied to A and $\|A \#_\alpha B_\varepsilon\|_\infty^{-1/\alpha} B_\varepsilon$ shows that for $r \ge 1$

$$A^r \#_\alpha B_\varepsilon^r \le \|A \#_\alpha B_\varepsilon\|_\infty^{r-1}(A \#_\alpha B_\varepsilon).$$

Letting $\varepsilon \downarrow 0$ we have by Lemma 4.1(1)

$$A^r \#_\alpha B^r \le \|A \#_\alpha B\|_\infty^{r-1}(A \#_\alpha B) \le A \#_\alpha B,$$

as desired.

(iii)$\Rightarrow$(i). For every $\varepsilon > 0$ and $t > 0$, it follows from (4.2) that

$$\begin{aligned}
e^{-tH} \#_\alpha (e^{-tK} + \varepsilon I) &= e^{-tH/2}\{e^{tH/2}(e^{-tK} + \varepsilon I)e^{tH/2}\}^\alpha e^{-tH/2} \\
&\ge e^{-tH/2}\{(1-\alpha)I + \alpha e^{-tH/2}(e^{-tK} + \varepsilon I)^{-1}e^{-tH/2}\}^{-1}e^{-tH/2} \\
&\ge e^{-tH/2}\{(1-\alpha)I + \alpha e^{-tH/2}e^{tK}e^{-tH/2}\}^{-1}e^{-tH/2} \\
&= \{(1-\alpha)e^{tH} + \alpha e^{tK}\}^{-1},
\end{aligned}$$

which shows that $e^{-tH} \#_\alpha e^{-tK} \ge \{(1-\alpha)e^{tH} + \alpha e^{tK}\}^{-1}$. Hence (iii) implies that $(1-\alpha)e^{tH} + \alpha e^{tK} \ge I$ for every $t > 0$. So for every $n \in \mathbb{N}$ we have

$$(4.10) \qquad \alpha n(e^{K/n} - I) \ge -(1-\alpha)n(e^{H/n} - I).$$

Let $K = \int_b^\infty \lambda dE_K(\lambda)$ be the spectral decomposition. For any $\xi \in E_K(c)\mathcal{H}$ with $b < c < \infty$, (4.10) says that

$$\alpha \int_b^c n(e^{\lambda/n} - 1)d\|E_K(\lambda)\xi\|^2 \ge -(1-\alpha)\langle n(e^{H/n} - I)\xi, \xi\rangle.$$

The above left-hand side tends as $n \to \infty$ to $\alpha \int_b^c \lambda d\|E_K(\lambda)\xi\|^2 = \alpha\langle K\xi, \xi\rangle$, while the right-hand side tends to $-(1-\alpha)\langle H\xi, \xi\rangle$. Therefore $\langle((1-\alpha)H + \alpha K)\xi, \xi\rangle \ge 0$. This implies (i) because $\bigcup_{c>b} E_K(c)\mathcal{H}$ is a core of $(1-\alpha)H + \alpha K$. ∎

Remark 4.7. By using the Furuta inequality [28], a more general result for positive invertible operators was shown in [29] (also [27]) as follows: If $A, B \in \mathcal{B}(\mathcal{H})_+$ are invertible, the following conditions, among others, are equivalent:

(I) $\log A \geq \log B$;

(II) $A^r \geq (A^{r/2} B^p A^{r/2})^{\frac{r}{p+r}}$ for all $p, r \geq 0$;

(III) for any $t \geq 0$, $A^{-r}(A^r B^p A^r)^{\frac{t+2r}{p+2r}} A^{-r}$ is a decreasing function of both $p \geq t$ and $r \geq 0$ to $\mathcal{B}(\mathcal{H})_+$. If we set $\alpha = r/(p+r)$, then the above (II) is written as $A^{-r} \#_\alpha B^{\frac{1-\alpha}{\alpha} r} \leq I$ for all $r \geq 0$ and $0 \leq \alpha \leq 1$. Thus it is seen that (I) $\Leftrightarrow$ (II) is nothing but (i) $\Leftrightarrow$ (ii) of Proposition 4.6.

To obtain the complementary counterpart of Theorem 3.7, let us introduce the additive version of antisymmetric tensor powers. For a lower-bounded self-adjoint operator H on $\mathcal{H}$ and $n \in \mathbb{N}$, we define the lower-bounded self-adjoint operator $\Sigma^n H$ on $\Lambda^n \mathcal{H}$ by

$$\Sigma^n H = \sum_{m=1}^{n} \left\{ (\otimes^{m-1} I) \otimes H \otimes (\otimes^{n-m} I) \right\} \Big|_{\Lambda^n \mathcal{H}}.$$

More precisely, for any $A \in \mathcal{B}(\mathcal{H})$, since it is easily checked that

(4.11)
$$\sum_{m=1}^{n} (\otimes^{m-1} I) \otimes A \otimes (\otimes^{n-m} I)$$

commutes with the projection of $\otimes^n \mathcal{H}$ onto $\Lambda^n \mathcal{H}$, we can define $\Sigma^n A$ as the restriction of (4.11) to $\Lambda^n \mathcal{H}$. Let $H_k = \int_a^k \lambda dE_H(\lambda)$ with the spectral decomposition $H = \int_a^\infty \lambda dE_H(\lambda)$. Then $\Sigma^n H$ is the limit of an increasing sequence $\{\Sigma^n H_k\}_{k=1}^\infty$ of (mutually commuting) self-adjoint operators in $\mathcal{B}(\Lambda^n \mathcal{H})$.

We prove (2) below as stated, while it will be used in the obvious case when H is bounded.

LEMMA 4.8. *Let H and K be lower-bounded self-adjoint operators on $\mathcal{H}$. For $n \in \mathbb{N}$ define $\Sigma^n H$ and $\Sigma^n K$ as above.*

(1) $\Lambda^n(e^{-H}) = e^{-\Sigma^n H}$.

(2) $\Sigma^n(H \hat{+} K) = (\Sigma^n H) \hat{+} (\Sigma^n K)$.

Proof. We may assume $H, K \geq 0$ by taking $H + aI, K + aI$.

(1) If H is bounded, then we have

$$e^{-\Sigma^n H} = \prod_{m=1}^{n} \left\{ (\otimes^{m-1} I) \otimes e^{-H} \otimes (\otimes^{n-m} I) \right\} \Big|_{\Lambda^n \mathcal{H}} = \otimes^n e^{-H} |_{\Lambda^n \mathcal{H}} = \Lambda^n(e^{-H}).$$

For general $H \geq 0$ let $H_k = \int_0^k \lambda dE_H(\lambda)$. Then $I \geq (I + H_k)^{-1} \downarrow (I + H)^{-1}$ as $k \to \infty$, which implies that $e^{-H_k} \to e^{-H}$ in SOT, because e^{-H} is the functional calculus of $(I + H)^{-1}$ for $f(x) = \exp(1 - x^{-1})$ on $[0, 1]$. Hence $\Lambda^n(e^{-H_k}) \to \Lambda^n(e^{-H})$ in SOT by Lemma 3.1(4). Similarly $e^{-\Sigma^n H_k} \to e^{-\Sigma^n H}$ in SOT, showing the assertion.

(2) Let P_0 be the projection of $\mathcal{H}$ onto the closure $\mathcal{H}_0$ of $\mathcal{D}(H^{1/2}) \cap \mathcal{D}(K^{1/2})$. Then $\Sigma^n(H \hat{+} K)$ is a positive self-adjoint operator on $\Lambda^n \mathcal{H}_0 (\subset \Lambda^n \mathcal{H})$. Since $\Lambda^n P_0$ is the projection of $\Lambda^n \mathcal{H}$ onto $\Lambda^n \mathcal{H}_0$ and (1) says that $e^{-\Sigma^n(H \hat{+} K)} = \Lambda^n(e^{-(H \hat{+} K)})$ on $\Lambda^n \mathcal{H}_0$, we have

$$e^{-\Sigma^n(H \hat{+} K)} \cdot \Lambda^n P_0 = \Lambda^n(e^{-(H \hat{+} K)}),$$

where $e^{-(H\hat{+}K)}=e^{-(H\hat{+}K)}P_0$ by convention in Theorem 3.6. Hence using Theorem 3.6(1), Lemma 3.1, and (1), we have

$$e^{-\Sigma^n(H\hat{+}K)} \cdot \Lambda^n P_0 = \operatorname*{s-lim}_{k\to\infty}\Lambda^n((e^{-H/k}e^{-K/k})^k) = \operatorname*{s-lim}_{k\to\infty}(\Lambda^n(e^{-H})^{1/k}\Lambda^n(e^{-K})^{1/k})^k$$

$$= \operatorname*{s-lim}_{k\to\infty}(e^{-(\Sigma^n H)/k}e^{-(\Sigma^n K)/k})^k = e^{-((\Sigma^n H)\hat{+}(\Sigma^n K))}.$$

Indeed, this shows that $(\Sigma^n H)\hat{+}(\Sigma^n K)$ is a self-adjoint operator on $\Lambda^n\mathcal{H}_0$ and $\Sigma^n(H\hat{+}K) = (\Sigma^n H)\hat{+}(\Sigma^n K)$. ∎

THEOREM 4.9. *If $H \in \mathcal{B}(\mathcal{H})$ is self-adjoint and K is a lower-bounded self-adjoint operator on $\mathcal{H}$, then*

$$(e^{-rH} \#_\alpha e^{-rK})^{1/r} \prec_{w(\log)} e^{-((1-\alpha)H+\alpha K)}, \quad r > 0.$$

Proof. Lemmas 3.1 and 4.8 imply that

$$\Lambda^n((e^{-rH} \#_\alpha e^{-rK})^{1/r}) = (\Lambda^n(e^{-rH}) \#_\alpha \Lambda^n(e^{-rK}))^{1/r} = (e^{-r\Sigma^n H} \#_\alpha e^{-r\Sigma^n K})^{1/r},$$

$$\Lambda^n(e^{-((1-\alpha)H+\alpha K)}) = e^{-((1-\alpha)\Sigma^n H+\alpha\Sigma^n K)}.$$

So, as in the proof of Theorem 4.2, we may show that

$$\|(e^{-rH} \#_\alpha e^{-rK})^{1/r}\|_\infty \leq \|e^{-((1-\alpha)H+\alpha K)}\|_\infty.$$

Thus it suffices to show that if $e^{-((1-\alpha)H+\alpha K)} \leq I$, i.e. $(1-\alpha)H + \alpha K \geq 0$, then $e^{-rH} \#_\alpha e^{-rK} \leq I$ for all $r > 0$. But this is (i)⇒(ii) of Proposition 4.6. ∎

PROBLEM 4.10. It is desirable to prove Theorem 4.9 when both H and K are lower-bounded. In view of Lemma 4.8, the problem consists in proving that $(1-\alpha)H\hat{+}\alpha K \geq 0$ implies $e^{-tH} \#_\alpha e^{-tK} \leq I$ for $t \geq 0$.

4.3. *Exponential product formula for operator means.* The next theorem is a Trotter-like product formula for operator means.

THEOREM 4.11. *Let σ be an operator mean with $\alpha = f'(1)$ for the corresponding operator monotone function f. If $H \in \mathcal{B}(\mathcal{H})$ is self-adjoint and K is a lower-bounded self-adjoint operator on $\mathcal{H}$, then*

$$\operatorname*{s-lim}_{r\downarrow 0}(e^{-rtH} \sigma e^{-rtK})^{1/r} = e^{-t((1-\alpha)H+\alpha K)}, \quad t > 0,$$

the convergence being uniform in $t \in [a,b]$ for any $0 < a < b$.

We divide the proof of the theorem into several lemmas.

LEMMA 4.12. *Let $G(t)$, $G_1(t)$, and $G_2(t)$ be SOT-continuous functions from $[0,\infty)$ to $\mathcal{B}(\mathcal{H})$ such that $G(t)$, $G_1(t)$, and $G_2(t)$ mutually commute for any $t \geq 0$ and such that for some $a \geq 0$,*

$$0 \leq G_1(t) \leq G(t) \leq G_2(t) \leq e^{at}I, \quad t \geq 0.$$

Let S be a self-adjoint operator on $\mathcal{H}$ and $\mathcal{D}$ a core of S. If

$$\lim_{t\downarrow 0}\left\|\frac{G_i(t)-I}{t}\xi + S\xi\right\| = 0, \quad \xi \in \mathcal{D}, \ i = 1,2,$$

then

$$\lim_{t \downarrow 0} \left\| \frac{G(t) - I}{t} \xi + S\xi \right\| = 0, \quad \xi \in \mathcal{D}.$$

P r o o f. Replacing $G(t), G_1(t), G_2(t)$ by those with multiple e^{-at} and S by $S + aI$, it suffices to show the case $a = 0$. Let $\xi \in \mathcal{D}$. Since by assumption

$$0 \le \frac{I - G_2(t)}{t} \le \frac{I - G(t)}{t} \le \frac{I - G_1(t)}{t},$$

we get

$$\left\langle \frac{I - G_2(t)}{t} \xi, \xi \right\rangle \le \left\langle \frac{I - G(t)}{t} \xi, \xi \right\rangle \le \left\langle \frac{I - G_1(t)}{t} \xi, \xi \right\rangle.$$

This implies that $\langle t^{-1}(I - G(t))\xi, \xi \rangle \to \langle S\xi, \xi \rangle$ as $t \downarrow 0$. By polarization we have

(4.12) $$\left\langle \frac{I - G(t)}{t} \xi, \eta \right\rangle \to \langle S\xi, \eta \rangle, \quad \eta \in \mathcal{D}.$$

Thanks to the commuting assumption of $G(t), G_1(t), G_2(t)$, we also get

$$\left(\frac{I - G_2(t)}{t} \right)^2 \le \left(\frac{I - G(t)}{t} \right)^2 \le \left(\frac{I - G_1(t)}{t} \right)^2,$$

so that

$$\left\| \frac{I - G_2(t)}{t} \xi \right\| \le \left\| \frac{I - G(t)}{t} \xi \right\| \le \left\| \frac{I - G_1(t)}{t} \xi \right\|.$$

Hence $\|t^{-1}(I - G(t))\xi\| \to \|S\xi\|$ and $t^{-1}(I - G(t))\xi$, $t > 0$, are bounded. These and (4.12) imply that $t^{-1}(I - G(t))\xi \to S\xi$ weakly. Therefore

$$\left\| \frac{I - G(t)}{t} \xi - S\xi \right\| = \left\| \frac{I - G(t)}{t} \xi \right\|^2 - 2\text{Re}\left\langle \frac{I - G(t)}{t} \xi, S\xi \right\rangle + \|S\xi\|^2$$

$$\to \|S\xi\|^2 - 2\langle S\xi, S\xi \rangle + \|S\xi\|^2 = 0,$$

showing the conclusion. ∎

Let H, K be as in Theorem 4.11 and define

$$F(t) = e^{-tH} \sigma e^{-tK}, \quad t \ge 0.$$

Let $L = (1 - \alpha)H + \alpha K$ and $\mathcal{D} = \mathcal{D}(K)$, which is the domain of L. Then:

LEMMA 4.13.
$$\lim_{t \downarrow 0} \left\| \frac{F(t) - I}{t} \xi + L\xi \right\| = 0, \quad \xi \in \mathcal{D}.$$

P r o o f. Define $A(t) = e^{tH/2}e^{-tK}e^{tH/2}$, so that $F(t) = e^{-tH/2}f(A(t))e^{-tH/2}$ for $t \ge 0$. Choose $a > 0$ such that $H \le aI$ and $K + aI \ge 0$. We have

(4.13) $$\frac{F(t) - I}{t}\xi = e^{-tH/2}f(A(t))\frac{e^{-tH/2} - I}{t}\xi + e^{-tH/2}\frac{f(A(t)) - I}{t}\xi + \frac{e^{-tH/2} - I}{t}\xi.$$

Since $A(t) \to I$ in SOT and hence $f(A(t)) \to I$ in SOT, the first and third terms in the right-hand side of (4.13) strongly converge as $t \downarrow 0$ to $(-H/2)\xi$, so that what remains to show is

(4.14) $$\lim_{t \downarrow 0} \left\| \frac{f(A(t)) - I}{t}\xi + \alpha(K - H)\xi \right\| = 0, \quad \xi \in \mathcal{D}.$$

Define $G(t) = f(A(t))$, $G_1(t) = A(t)((1 - \alpha)A(t) + \alpha I)^{-1}$, and $G_2(t) = (1 - \alpha)I + \alpha A(t)$, which mutually commute for each $t \geq 0$. Since $G_2(t) \leq e^{2at}I$ and (4.2) yields $0 \leq G_1(t) \leq G(t) \leq G_2(t)$, we can apply Lemma 4.12. Let $\xi \in \mathcal{D}$. It is easy to check that

$$\left\| \frac{G_2(t) - I}{t}\xi + \alpha(K - H)\xi \right\| = \alpha\left\| \frac{A(t) - I}{t}\xi + (K - H)\xi \right\| \to 0.$$

Furthermore, we get

$$\left\| \frac{G_1(t) - I}{t}\xi + \alpha(K - H)\xi \right\|$$

$$\leq \left\| ((1 - \alpha)A(t) + \alpha I)^{-1}\left\{ \frac{A(t) - ((1 - \alpha)A(t) + \alpha I)}{t}\xi + \alpha(K - H)\xi \right\} \right\|$$

$$+ \|\{((1 - \alpha)A(t) + \alpha I)^{-1} - I\}\alpha(K - H)\xi\|$$

$$\leq \left\| \frac{A(t) - I}{t}\xi + (K - H)\xi \right\| + (1 - \alpha)\|(A(t) - I)(K - H)\xi\| \to 0$$

thanks to $\|((1 - \alpha)A(t) + \alpha I)^{-1}\|_\infty \leq \alpha^{-1}$. Hence Lemma 4.12 implies (4.14). ∎

To prove the theorem, we may assume $H, K \geq 0$ by taking $H + aI, K + aI$, so that $0 \leq F(t) \leq I$ for all $t \geq 0$. Now let us proceed as in [19]. Fix any sequence $\{s_n\}$ such that $0 < s_n \to \infty$. Let $L_n = s_n(I - F(s_n^{-1}))$ so that $L_n \geq 0$. Then Lemma 4.13 says that $\|(L - L_n)\xi\| \to 0$ for all $\xi \in \mathcal{D}$.

LEMMA 4.14.
$$\mathop{\text{s-lim}}_{n \to \infty}(F(s_n^{-1})^{s_n} - e^{-L_n}) = 0.$$

Proof. For any $\xi \in \mathcal{H}$ we have

$$\|F(s_n^{-1})^{s_n}\xi - e^{-L_n}\xi\| = \left\| F(s_n^{-1})^{s_n}\xi - e^{-s_n}\sum_{k=0}^{\infty}\frac{s_n^k}{k!}F(s_n^{-1})^k\xi \right\|$$

$$\leq e^{-s_n}\sum_{k=0}^{\infty}\frac{s_n^k}{k!}\|(F(s_n^{-1})^{s_n} - F(s_n^{-1})^k)\xi\|$$

$$\leq e^{-s_n}\sum_{k=0}^{\infty}\frac{s_n^k}{k!}\|(I - F(s_n^{-1})^{|k-s_n|})\xi\|.$$

Since

$$0 \leq I - F(s_n^{-1})^r \leq \begin{cases} r(I - F(s_n^{-1})), & r \geq 1, \\ I - F(s_n^{-1}), & 0 \leq r \leq 1, \end{cases}$$

it follows that

$$\|F(s_n^{-1})^{s_n}\xi - e^{-L_n}\xi\| \leq e^{-s_n}\|(I - F(s_n^{-1}))\xi\|\sum_{k=0}^{\infty}\frac{s_n^k}{k!}|k - s_n| + \|(I - F(s_n^{-1}))\xi\|.$$

The Schwarz inequality gives

$$\sum_{k=0}^{\infty}\frac{s_n^k}{k!}|k - s_n| \leq \left(\sum_{k=0}^{\infty}\frac{s_n^k}{k!} \right)^{1/2}\left(\sum_{k=0}^{\infty}\frac{s_n^k}{k!}(k - s_n)^2 \right)^{1/2} = e^{s_n/2}(s_ne^{s_n})^{1/2} = s_n^{1/2}e^{s_n}.$$

Hence for every $\xi \in \mathcal{D}$,

$$\|F(s_n^{-1})^{s_n}\xi - e^{-L_n}\xi\| \leq (s_n^{1/2} + 1)\|(I - F(s_n^{-1}))\xi\| = \frac{s_n^{1/2} + 1}{s_n}\|L_n\xi\| \to 0.$$

This shows the assertion because $F(s_n^{-1})^{s_n}$ and e^{-L_n} are contractions. ∎

The next step is much simpler than that in [19], because the functional calculus can be applied in our setting.

LEMMA 4.15.
$$\operatorname*{s-lim}_{n\to\infty} e^{-L_n} = e^{-L}.$$

Proof. First note that $0 \leq (I + L_n)^{-1} \leq I$ and $0 \leq (I + L)^{-1} \leq I$. If $\eta = (I + L)\xi$ with $\xi \in \mathcal{D}$, then

$$\begin{aligned}\|(I + L_n)^{-1}\eta - (I + L)^{-1}\eta\| &= \|(I + L_n)^{-1}\{(I + L_n)\xi + (L - L_n)\xi\} - \xi\| \\ &= \|(I + L_n)^{-1}(L - L_n)\xi\| \leq \|(L - L_n)\xi\| \to 0.\end{aligned}$$

This implies that $(I + L_n)^{-1} \to (I + L)^{-1}$ in SOT. Now the result follows due to the functional calculus as in the proof of Lemma 4.8(1). ∎

Proof of Theorem 4.11 (Completion). Lemmas 4.14 and 4.15 show that $F(s_n^{-1})^{s_n} \to e^{-L}$ in SOT if $0 < s_n \to \infty$; namely

(4.15) $$\operatorname*{s-lim}_{n\to\infty} F(r_n)^{1/r_n} = e^{-L} \text{ if } 0 < r_n \to 0.$$

Replacing H, K by tH, tK with $t > 0$ yields

$$\operatorname*{s-lim}_{r\downarrow 0}(e^{-rtH}\,\sigma\,e^{-rtK})^{1/r} = e^{-tL}, \quad t > 0.$$

Finally it is an easy task to show the uniform convergence. Indeed, let $0 < a < b$ and suppose that the above convergence is not uniform in $t \in [a, b]$. Then there exist $\xi \in \mathcal{H}$, $\varepsilon > 0$, $r_n \downarrow 0$, and $t_n \in [a, b]$ such that

(4.16) $$\|F(r_n t_n)^{1/r_n}\xi - e^{-t_n L}\xi\| \geq \varepsilon, \quad n \in \mathbb{N}.$$

Choosing a subsequence we may assume that $t_n \to t$. Then since $\lambda^{t_n} \to \lambda^t$ uniformly in $\lambda \in [0, 1]$, it follows from (4.15) that $F(r_n t_n)^{1/r_n} = (F(r_n t_n)^{1/r_n t_n})^{t_n} \to e^{-tL}$ in SOT as well as $\|e^{-t_n L} - e^{-tL}\|_\infty \to 0$, contradicting (4.16). ∎

It would be expected that Theorem 4.9 is a consequence of Theorems 4.2 and 4.11. However this is not truly the case, because $\prod_{i=1}^n \mu_i(A)$ is not continuous in SOT.

It is worth noting the following specialization to the matrix case.

COROLLARY 4.16. *For arbitrary Hermitian matrices H and K,*

$$\frac{d}{dt}e^{tH}\,\sigma\,e^{tK}\Big|_{t=0} = (1 - \alpha)H + \alpha K,$$

$$\lim_{t\to 0}(e^{tH}\,\sigma\,e^{tK})^{1/t} = e^{(1-\alpha)H + \alpha K},$$

$$\frac{d}{dt}\log(e^{tH}\,\sigma\,e^{tK})\Big|_{t=0} = (1 - \alpha)H + \alpha K.$$

Proof. Theorem 4.11 gives the second, while the first follows from Lemma 4.13. For the third, take the logarithm of the second. ∎

R e m a r k 4.17. Theorem 4.11 fails to hold when both H and K are lower-bounded. Indeed, Kosaki [56] gave an example of positive self-adjoint operators H, K such that $\mathcal{D}(H^{1/2}) \cap \mathcal{D}(K^{1/2})$ is dense in $\mathcal{H}$ and hence $e^{-(H \hat{+} K)/2}$ is non-zero, but $e^{-rH} : e^{-rK} = 0$ for each $r > 0$.

4.4. *Norm inequalities of complementary Golden–Thompson type.* Theorems 4.9 and 4.11 show:

COROLLARY 4.18. *Let σ be an operator mean such that $\sigma \leq \#_\alpha$, i.e. $f(x) \leq x^\alpha$ for $x \geq 0$ where f is the corresponding operator monotone function. If H and K are as in Theorem 4.11 and $\| \cdot \|$ is a symmetric norm, then*

$$(4.17) \qquad \|(e^{-rH} \, \sigma \, e^{-rK})^{1/r}\| \leq \|e^{-((1-\alpha)H + \alpha K)}\|, \quad r > 0,$$

and moreover

$$(4.18) \qquad \lim_{r \downarrow 0} \|(e^{-rH} \, \sigma \, e^{-rK})^{1/r}\| = \|e^{-((1-\alpha)H + \alpha K)}\|.$$

If $\sigma = \#_\alpha$ then the left-hand side of (4.17) increases to $\|e^{-((1-\alpha)+\alpha K)}\|$ as $r \downarrow 0$.

P r o o f. Since $e^{-rH} \, \sigma \, e^{-rK} \leq e^{-rH} \#_\alpha e^{-rK}$ by assumption, Theorem 4.9 gives for every $r > 0$

$$\|(e^{-rH} \, \sigma \, e^{-rK})^{1/r}\| \leq \|(e^{-rH} \#_\alpha e^{-rK})^{1/r}\| \leq \|e^{-((1-\alpha)H + \alpha K)}\|.$$

On the other hand, in view of Proposition 2.11, Theorem 4.11 implies that

$$\|e^{-((1-\alpha)H + \alpha K)}\| \leq \liminf_{r \downarrow 0} \|(e^{-rH} \, \sigma \, e^{-rK})^{1/r}\|.$$

Therefore the first part is shown. The second part is immediate from (4.4). ∎

In particular, if σ is a symmetric operator mean with $\sigma \leq \#$ and H, K are as above, then

$$(4.19) \qquad \|(e^{-2rH} \, \sigma \, e^{-2rK})^{1/r}\| \leq \|e^{-(H+K)}\|, \quad r > 0,$$

and the left-hand side converges to $\|e^{-(H+K)}\|$ as $r \downarrow 0$. For instance, when $\sigma = !$ or $\sigma = \#$, the left-hand side of (4.19) increases to the right as $r \downarrow 0$. Norm inequalities such as (4.17) and (4.19) are considered as complementary to Golden–Thompson ones (see [6, 44, 71]).

R e m a r k s 4.19. (1) Let $H = 0$, $K = xI$ for $x \in \mathbb{R}$, and $\| \cdot \| = \| \cdot \|_\infty$. Then (4.17) menas that $f(e^{-rx}) \leq e^{-\alpha rx}$ for $r > 0$ and so $\sigma \leq \#_\alpha$. Thus the assumption $\sigma \leq \#_\alpha$ is essential in Corollary 4.18.

(2) Inequality (4.19) does not hold for a general symmetric mean. Indeed, if $e^{-K} \in \mathcal{C}_p(\mathcal{H})$ where $0 < p < \infty$ then, for any bounded H, $e^{-(H+K)} \in \mathcal{C}_p(\mathcal{H})$ by (3.11), but $\|(e^{-2rH} \nabla e^{-2rK})^{1/r}\|_p = \infty$ for all $r > 0$.

When we are concerned with (quasi-)norms $\| \cdot \|_p$, the following holds:

PROPOSITION 4.20. *Let σ be as in Corollary 4.18. Let H and K be lower-bounded self-adjoint operators such that $H + K$ is essentially bounded. If $e^{-K} \in \mathcal{C}_p(\mathcal{H})$ where $0 < p < \infty$, then*

$$\|(e^{-rH/(1-\alpha)} \, \sigma \, e^{-rK/\alpha})^{1/r}\|_p \leq \|e^{-(H+K)}\|_p, \quad r > 0.$$

P r o o f. The case $p = 1$ is enough because we may replace H, K by pH, pK (see the proof of Corollary 3.10). So assume that $e^{-K} \in \mathcal{C}_1(\mathcal{H})$. Hence $e^{-(H+K)} \in \mathcal{C}_1(\mathcal{H})$ by (3.11). Let $H_n = \int_a^n \lambda dE_H(\lambda)$. Then we proved the convergence (3.22). For each $r > 0$, since $e^{-rH/(1-\alpha)} \leq e^{-rH_n/(1-\alpha)}$, we get

$$\|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r}\|_1 \leq \|(e^{-rH_n/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r}\|_1 \leq \|e^{-(H_n+K)}\|_1$$

by (4.17). This and (3.22) yield the conclusion. ∎

4.5. *Norm convergence of exponential product formula for operator means.* In this subsection let us consider norm convergence of $(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r}$ as $r \downarrow 0$ when $\| \cdot \|$ is uniformly convex or $\| \cdot \| = \| \cdot \|_p$ for $0 < p < \infty$.

In Banach space theory, there are several important geometric notions of convexity or smoothness for norms (see e.g. [13]). For instance, a Banach space $\mathcal{X}$ is said to be *uniformly convex* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that, for any $x, y \in \mathcal{X}$, $\|(x + y)/2\| \leq 1 - \delta$ holds whenever $\|x\| = \|y\| = 1$ and $\|x - y\| \geq \varepsilon$. A Hilbert space is a typical example of a uniformly convex Banach space. As is well known, a uniformly convex Banach space is reflexive. A useful property of a uniformly convex space $\mathcal{X}$ is that if $\{x_j\} \subset \mathcal{X}$ weakly converges to $x \in \mathcal{X}$ and $\|x_j\| \to \|x\|$, then $\|x_j - x\| \to 0$. This property is typical in a Hilbert space, as was already used in the proof of Theorem 3.12.

When $1 < p < \infty$, the uniform convexity of $\mathcal{C}_p(\mathcal{H})$ is an immediate consequence of the Clarkson–McCarthy inequalities:

$$\|A + B\|_p^p + \|A - B\|_p^p \leq 2^{p-1}(\|A\|_p^p + \|B\|_p^p), \quad 2 \leq p < \infty,$$

$$\|A + B\|_p^q + \|A - B\|_p^q \leq 2(\|A\|_p^p + \|B\|_p^p)^{q/p}, \quad 1 < p \leq 2,$$

where $1/p + 1/q = 1$. (The proof in the most general setup is found in [26].)

Now let Φ be a symmetric gauge function, Φ' the conjugate one, and $\|\cdot\|$ the symmetric norm corresponding to Φ. Assume that $\mathcal{C}_\Phi(\mathcal{H})$ (or $\| \cdot \|$) is uniformly convex. Then the reflexivity of $\mathcal{C}_\Phi(\mathcal{H})$ implies as remarked after Theorem 2.9 that both Φ and Φ' are regular and so $\mathcal{C}_\Phi(\mathcal{H})^* \cong \mathcal{C}_{\Phi'}(\mathcal{H})$.

LEMMA 4.21. *Assume that $\mathcal{C}_\Phi(\mathcal{H})$ is uniformly convex. If $\{A_j\} \subset \mathcal{C}_\Phi(\mathcal{H})$, $\sup_j \|A_j\| < \infty$, and $A_j \to A$ in WOT, then $A_j \to A$ in $w(\mathcal{C}_\Phi(\mathcal{H}), \mathcal{C}_{\Phi'}(\mathcal{H}))$, the weak topology.*

P r o o f. First the assumptions give $A \in \mathcal{C}_\Phi(\mathcal{H})$ by Proposition 2.11. Since Φ' is regular as remarked above, $\mathcal{C}_{\mathrm{fin}}(\mathcal{H})$ is dense in $\mathcal{C}_{\Phi'}(\mathcal{H})$. By the WOT-convergence of $\{A_j\}$ we get $\mathrm{tr}((A_j - A)B) \to 0$ for any $B \in \mathcal{C}_{\mathrm{fin}}(\mathcal{H})$. In view of Theorem 2.9, this and the $\| \cdot \|$-boundedness imply the conclusion. ∎

COROLLARY 4.22. *Let σ be as in Corollary 4.18. Assume that $\|\cdot\|$ is uniformly convex. If H and K are as in Theorem 4.11 and $e^{-K} \in \mathcal{C}_\Phi(\mathcal{H})$, then*

$$\lim_{r \downarrow 0} \|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r} - e^{-(H+K)}\| = 0.$$

P r o o f. By (3.11), (4.17), Theorem 4.11, and Lemma 4.21 altogether, we have as $r \downarrow 0$,

$$(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r} \to e^{-(H+K)} \quad \text{in } w(\mathcal{C}_\Phi(\mathcal{H}), \mathcal{C}_{\Phi'}(\mathcal{H})).$$

Hence the result follows from the uniform convexity and (4.18). ∎

Concerning $\|\cdot\|_p$, $0 < p < \infty$, we have:

COROLLARY 4.23. *Let σ be as in Corollary* 4.18. *If H and K are as in Theorem* 4.11 *and $e^{-K} \in \mathcal{C}_p(\mathcal{H})$ where $0 < p < \infty$, then*

$$\lim_{r \downarrow 0} \|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r} - e^{-(H+K)}\|_p = 0.$$

P r o o f. The case $1 < p < \infty$ is included in Corollary 4.22. Let $0 < p \leq 1$ and choose $k \in \mathbb{N}$ such that $2^k > 1/p$. Applying Corollary 4.22 to $\|\cdot\|_{2^k p}$ and $2^{-k}H, 2^{-k}K$, we have

(4.20) $$\lim_{r \downarrow 0} \|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r} - e^{-(H+K)/2^k}\|_{2^k p} = 0.$$

Noting that $\|X+Y\|_q \leq 2^{1/q}(\|X\|_q + \|Y\|_q)$ for any $q > 0$, we get by the Hölder inequality

$$\|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^{k-1}r} - e^{-(H+K)/2^{k-1}}\|_{2^{k-1}p}$$

$$\leq 2^{1/2^{k-1}p}\{\|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r}$$

$$\times [(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r} - e^{-(H+K)/2^k}]\|_{2^{k-1}p}$$

$$+ \|[(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r} - e^{-(H+K)/2^k}]e^{-(H+K)/2^k}\|_{2^{k-1}p}\}$$

$$\leq 2^{1/2^{k-1}p}\|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r} - e^{-(H+K)/2^k}\|_{2^k p}$$

$$\times \{\|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^k r}\|_{2^k p} + \|e^{-(H+K)/2^k}\|_{2^k p}\}.$$

This implies by (4.20) and Proposition 4.20 that

$$\lim_{r \downarrow 0} \|(e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/2^{k-1}r} - e^{-(H+K)/2^{k-1}}\|_{2^{k-1}p} = 0.$$

Repeating this argument yields the result. ∎

5. Miscellaneous results. In this section we first discuss the interplay between log-majorizations and the Furuta inequalities [28, 30]. Several log-majorization results are obtained in Subsection 5.2. Furthermore, we give determinant inequalities as simple applications of the log-majorization results.

5.1. *Interplay between log-majorization and Furuta inequlity.* The Furuta inequality [28] says that if $A \geq B \geq 0$ in $\mathcal{B}(\mathcal{H})_+$, then

$$(B^{r/2}A^p B^{r/2})^\alpha \geq B^{\alpha(p+r)},$$

(5.1) $$A^{\alpha(p+r)} \geq (A^{r/2}B^p A^{r/2})^\alpha,$$

whenever $0 \leq \alpha \leq 1$, $p, r \geq 0$, and $1 + r \geq \alpha(p + r)$.

As was clarified in [6], we can transform Furuta type operator inequalities into log-majorizations and vice versa. For instance, the Furuta inequality (5.1) is reformulated in terms of log-majorization as follows.

PROPOSITION 5.1. *For every $A, B \in \mathcal{B}(\mathcal{H})_+$, if $0 < \alpha < 1$, $p \geq 0$, and $0 \leq r \leq \min\{\alpha, \alpha p\}$, then*

$$\mu(A^{\frac{p-r}{1-\alpha}} \#_\alpha (A^{r/2\alpha}B^{p/\alpha}A^{r/2\alpha})) \prec_{w(\log)} \mu((A^{1/2}BA^{1/2})^p).$$

Furthermore, the assumption $r \geq 0$ is removed when A is invertible.

Proof. We may assume that A is invertible. Let $0 < \alpha < 1$, $q, s \geq 0$, and $1 + s \geq \alpha(q + s)$. Then it is seen from (5.1) that $A^{1/2}BA^{1/2} \leq I$ or $A^{-1} \geq B$ implies

$$A^{\alpha(q+s)} \#_\alpha (A^{(\alpha(q+s)-s)/2}B^q A^{(\alpha(q+s)-s)/2}) = A^{\alpha(q+s)/2}(A^{-s/2}B^q A^{-s/2})^\alpha A^{\alpha(q+s)/2} \leq I.$$

Arranging the order of homogeneity, we have

$$\|A^{\alpha(q+s)} \#_\alpha (A^{(\alpha(q+s)-s)/2}B^q A^{(\alpha(q+s)-s)/2})\|_\infty \leq \|(A^{1/2}BA^{1/2})^{\alpha q}\|_\infty.$$

As in the proofs of Theorems 3.4 and 4.2, the antisymmetric tensor technique yields

$$\mu(A^{\alpha(q+s)} \#_\alpha (A^{(\alpha(q+s)-s)/2}B^q A^{(\alpha(q+s)-s)/2})) \prec_{w(\log)} \mu((A^{1/2}BA^{1/2})^{\alpha q}).$$

Now putting $p = \alpha q$ and $r = \alpha(\alpha(q+s)-s)$, we get the conclusion, where the assumptions $q, s \geq 0$ and $1 + s \geq \alpha(q + s)$ are transformed into $p \geq 0$ and $r \leq \min\{\alpha, \alpha p\}$. ∎

COROLLARY 5.2. *If $A, B \in \mathcal{B}(\mathcal{H})_+$ and A is invertible, then for every $0 \leq \alpha \leq 1$,*

$$\mu(A^\alpha(A^{-1/2}BA^{-1/2})^\alpha A^\alpha) \prec_{w(\log)} \mu(A^{\alpha/2}B^\alpha A^{\alpha/2}) \prec_{w(\log)} \mu((A^{1/2}BA^{1/2})^\alpha).$$

Proof. The case $0 < \alpha < 1$ is enough. The second log-majorization is nothing but (3.6). For the first, take $p = 1$ and $r = 2\alpha - 1$ in Proposition 5.1. Then we have

$$\mu(A(A^{-1/2\alpha}B^{1/\alpha}A^{-1/2\alpha})^\alpha A) \prec_{w(\log)} \mu(A^{1/2}BA^{1/2}),$$

which gives the result by replacing A, B by A^α, B^α, respectively. ∎

Also, we can consider the converse direction from log-majorizations to operator inequalities. For instance, the following is a reformulation of the log-majorization (4.3).

PROPOSITION 5.3. *If $A \geq B \geq 0$ and A is invertible, then*

$$(5.2) \qquad\qquad A^r \geq \{A^{r/2}(A^{-1/2}B^p A^{-1/2})^r A^{r/2}\}^{1/p}, \qquad p, r \geq 1.$$

Proof. Let $A, B \in \mathcal{B}(\mathcal{H})_+$ with A invertible. Let $0 < \alpha \leq 1$. Then (4.3) says that $A^{-1} \geq (A^{-1/2}BA^{-1/2})^\alpha$ implies $A^{-r} \geq (A^{-r/2}B^r A^{-r/2})^\alpha$ for every $r \geq 1$. Put $p = 1/\alpha$ and replace $A^{-1}, (A^{-1/2}BA^{-1/2})^\alpha$ by A, B, respectively. Then the result follows. ∎

Recently Furuta [30] obtained a quite general operator inequality extending the original Furuta inequality as well as (5.2), which is stated in the following:

THEOREM 5.4. *Assume that $A \geq B \geq 0$ and A is invertible. Then*

$$(5.3) \qquad\qquad A^{1-t+r} \geq \{A^{r/2}(A^{-t/2}B^p A^{-t/2})^s A^{r/2}\}^{\frac{1-t+r}{(p-t)s+r}},$$

whenever $0 \leq t \leq 1$, $p \geq 1$, $r \geq t$, and $s \geq 1$. Furthermore, for every $0 \leq t \leq 1$ and $p \geq 1$,

$$(r, s) \mapsto A^{-r/2}\{A^{r/2}(A^{-t/2}B^p A^{-t/2})^s A^{r/2}\}^{\frac{1-t+r}{(p-t)s+r}} A^{-r/2}$$

is a decreasing function of both $r \geq t$ and $s \geq 1$.

Note that (5.1) is a consequence of (5.3) in case of $t = 0$ and $s = 1$. Also (5.2) is a special case of (5.3) when $t = 1$ and $r = s \geq 1$. In this way, the operator inequality (5.3) interpolates (5.1) and (5.2) by the parameter $0 \leq t \leq 1$.

In [30] several log-majorizations extending (4.3) were produced from (5.3) in the same way as in the proof of Proposition 5.1. Among others, the following is obtained from (5.3) by taking $t = 1$ and $p = 1/\alpha$. But for A compact, see the proof of Theorem 4.2.

THEOREM 5.5. *If $A, B \in \mathcal{B}(\mathcal{H})_+$ and A is either invertible or compact, then for every $0 \le \alpha \le 1$ and $r, s \ge 1$*

$$\mu(A^r \#_{\alpha q/s} B^s) \prec_{w(\log)} \mu((A \#_\alpha B)^q)$$

where $q = ((1 - \alpha)r^{-1} + \alpha s^{-1})^{-1}$.

5.2. *Other log-majorizations.* In this subsection let us show some other log-majorization results. The first two propositions were given in [6] in the matrix case, while the remaining are new.

For any $T \in \mathcal{B}(\mathcal{H})$, the real part of T is $\operatorname{Re} T = (T + T^*)/2$, and the exponential operator e^T is defined as the power series expansion $e^T = \sum_{n=0}^{\infty} T^n/n!$. The following log-majorization (1) for exponential operator is essentially due to [14].

PROPOSITION 5.6. *Let $T \in \mathcal{B}(\mathcal{H})$.*

(1) $\mu(e^T) \prec_{w(\log)} \mu(e^{\operatorname{Re} T})$.

(2) *Assume that* $\operatorname{Re} T \ge 0$. *Then* $|e^T| \ge I$ *(or* $\log|e^T| \ge 0$*) and* $\mu(\log|e^T|) \prec_w \mu(\operatorname{Re} T)$. *Hence* $\|\log|e^T|\| \le \|\operatorname{Re} T\|$ *for any symmetric norm* $\|\cdot\|$.

P r o o f. (1) For every $X \in \mathcal{B}(\mathcal{H})$ we get by (1.14)

$$\mu(X^k) \prec_{w(\log)} \mu(X)^k, \quad k \in \mathbb{N}.$$

Putting $X = e^{T/k}$ in the above yields

$$\mu(e^T) \prec_{w(\log)} \mu(|e^{T/k}|^k), \quad k \in \mathbb{N}.$$

But $|e^{T/2k}|^{2k} = (e^{T^*/2k}e^{T/2k})^k$ converges in the norm $\|\cdot\|_\infty$ to $e^{(T^*+T)/2} = e^{\operatorname{Re} T}$ as $k \to \infty$. Hence the assertion is shown by Proposition 1.4(10).

(2) If $\operatorname{Re} T \ge 0$, then since

$$|e^T|^{-1} = (e^{T^*}e^T)^{-1/2} = (e^{-T}e^{-T^*})^{1/2} = |e^{-T^*}|,$$

we get by (1),

$$\||e^T|^{-1}\|_\infty \le \|e^{\operatorname{Re}(-T^*)}\|_\infty = \|e^{-\operatorname{Re} T}\|_\infty \le 1,$$

so that $|e^T| \ge I$. Then the required weak majorization is immediate from (1), because $\mu_n(\log|e^T|) = \log\mu_n(e^T)$ and $\mu_n(\operatorname{Re} T) = \log\mu_n(e^{\operatorname{Re} T})$. ∎

PROPOSITION 5.7. *Let $A, B \in \mathcal{B}(\mathcal{H})_+$.*

(1) $\mu(A^{p_1}B^{q_1}A^{p_2}B^{q_2}\ldots A^{p_k}B^{q_k}) \prec_{w(\log)} \mu(A^{p_1+\cdots+p_k}B^{q_1+\cdots+q_k})$ *if* $0 \le p_1 \le q_1 \le p_1 + p_2 \le q_1 + q_2 \le \ldots \le q_1 + \ldots + q_{k-1} \le p_1 + \ldots + p_k = q_1 + \ldots + q_k$ *(in particular, if $p_j = q_j \ge 0$ for all j).*

(2) $\mu(A^p B^q A^r) \prec_{w(\log)} \mu(A^{p+r}B^q)$ *for every $p, q, r \ge 0$.*

(3) $\mu((AB)^k A) \prec_{w(\log)} \mu(A^{(k+1)/2}B^k A^{(k+1)/2})$ *for every $k \in \mathbb{N}$.*

P r o o f. (1) We may assume that $p_1 + \ldots + p_k = q_1 + \ldots + q_k = 1$. It suffices as in the proof of Theorem 3.4 to show that when A, B are invertible, $\|AB\|_\infty \le 1$ implies $\|A^{p_1}B^{q_1}\ldots A^{p_k}B^{q_k}\|_\infty \le 1$. Suppose $\|AB\|_\infty \le 1$ or $BA^2B \le I$. Then $A^2 \le B^{-2}$ and $B^2 \le A^{-2}$. Since $0 \le q_1 - p_1 \le 1$, we get

$$B^{q_1}A^{2p_1}B^{q_1} \le B^{q_1}B^{-2p_1}B^{q_1} = B^{2(q_1-p_1)} \le A^{2(p_1-q_1)}$$

and hence
$$A^{p_2}B^{q_1}A^{2p_1}B^{q_1}A^{p_2} \le A^{2(p_1+p_2-q_1)} \le B^{2(q_1-p_1-p_2)}.$$

Repeating this argument yields
$$B^{q_k}A^{p_k}\dots B^{q_1}A^{2p_1}B^{q_1}\dots A^{p_k}B^{q_k} \le B^{2(q_1+\dots+q_k-p_1-\dots-p_k)} = I,$$

so that $\|A^{p_1}B^{q_1}\dots A^{p_k}B^{q_k}\|_\infty \le 1$.

(2) Put $p_1 = p$, $p_2 = r$, $q_1 = p + r$, and $q_2 = 0$ in (1). Then
$$\mu(A^pB^qA^r) = \mu(A^p(B^{\frac{q}{p+r}})^{p+r}A^r) \prec_{w(\log)} \mu(A^{p+r}(B^{\frac{q}{p+r}})^{p+r}) = \mu(A^{p+r}B^q).$$

(3) When $k = 2m - 1$ we get
$$\mu((AB)^{2m-1}A) = \mu((AB)^{m-1}AB^{1/2})^2$$
$$= \mu(\{(A^{\frac{2m}{2m-1}})^{\frac{2m-1}{2m}}B\}^{m-1}(A^{\frac{2m}{2m-1}})^{\frac{2m-1}{2m}}B^{1/2})^2$$
$$\prec_{w(\log)} \mu(A^mB^{(2m-1)/2})^2 = \mu(A^mB^{2m-1}A^m).$$

The log-majorization in the above follows from (1) with
$$p_1 = \dots = p_m = \frac{2m-1}{2m}, \quad q_1 = \dots = q_{m-1} = 1, \quad q_m = \frac{1}{2}.$$

The derivation is similar when $k = 2m$. ∎

Let H and K be lower-bounded self-adjoint operators on $\mathcal{H}$. For $t > 0$ put $S(t) = e^{-tH/2}e^{-tK}e^{-tH/2}$, which is the second order approximant of $e^{-t(H\hat{+}K)}$. Then $\{S(t/n)S((1-2t)/n)S(t/n)\}^n$ and $\{S(t/n)^2S((1-4t)/n)S(t/n)^2\}^n$ are sometimes used as higher order approximants of $e^{-(H\hat{+}K)}$. When $p_1,\dots,p_k \ge 0$ and $\sum_{j=1}^k p_j = 1$, Proposition 5.7(1) gives
$$\mu(S(p_1)S(p_2)\dots S(p_k)) \prec_{w(\log)} \mu(e^{-H}e^{-K}).$$

A special case of the norm inequality proved by Kato [50] is the following:

(5.4) $$\|A^rXB^r\|_\infty \le \|X\|_\infty^{1-r}\|AXB\|_\infty^r, \quad 0 \le r \le 1,$$

for every $A, B \in \mathcal{B}(\mathcal{H})_+$ and $X \in \mathcal{B}(\mathcal{H})$. For the proof of (5.4), it is enough to show that $\log\|A^rXB^r\|_\infty$ is convex in $r \in [0,1]$. This follows because for $0 \le r \le s \le 1$,
$$\|A^{\frac{r+s}{2}}XB^{\frac{r+s}{2}}\|_\infty^2 = r(B^{\frac{r+s}{2}}X^*A^{r+s}XB^{\frac{r+s}{2}}) = r(B^rX^*A^rA^sXB^s)$$
$$\le \|A^rXB^r\|_\infty\|A^sXB^s\|_\infty.$$

Applying the antisymmetric tensor technique to (5.4), we have the following generalization of Theorem 3.4.

PROPOSITION 5.8. *For every $A, B \in \mathcal{B}(\mathcal{H})_+$ and $X \in \mathcal{B}(\mathcal{H})$,*

(5.5) $$\mu(A^rXB^r) \prec_{w(\log)} \mu(X)^{1-r}\mu(AXB)^r, \quad 0 \le r \le 1,$$

or equivalently
$$\mu(AXB)^r \prec_{w(\log)} \mu(X)^{r-1}\mu(A^rXB^r), \quad r \ge 1.$$

This yields the norm inequality proved by Kittaneh [52] as follows.

COROLLARY 5.9. *Let $A, B \in \mathcal{B}(\mathcal{H})_+$, $X \in \mathcal{B}(\mathcal{H})$, and $\|\cdot\|$ be a symmetric norm. Then*

(5.6) $$\|A^rXB^r\| \le \|X\|^{1-r}\|AXB\|^r, \quad 0 \le r \le 1.$$

P r o o f. Let Φ be the corresponding gauge function. For each $0 < r < 1$, we have by Proposition 1.3 and by Lemmas 2.2 and 2.13(1)

$$\|A^r X B^r\| \le \Phi(\mu(X)^{1-r}\mu(AXB)^r) \le \Phi(\mu(X))^{1-r}\Phi(\mu(AXB))^r = \|X\|^{1-r}\|AXB\|^r. \quad \blacksquare$$

We have also

(5.7) $$\mu(A^r X B^{1-r}) \prec_{w(\log)} \mu(AX)^r \mu(XB)^{1-r}, \quad 0 \le r \le 1,$$

modifying (5.5) in the same way as in [52] where it is observed that (5.6) is equivalently reformulated as

$$\|A^r X B^{1-r}\| \le \|AX\|^r \|XB\|^{1-r}, \quad 0 \le r \le 1.$$

In fact, to check (5.7), we may assume that B is invertible. Then for $0 \le r \le 1$ we get by (5.5)

$$\mu(A^r X B^{1-r}) = \mu(A^r (XB) B^{-r}) \prec_{w(\log)} \mu(XB)^{1-r}\mu(A(XB)B^{-1})^r$$
$$= \mu(AX)^r \mu(XB)^{1-r}.$$

Furthermore, for any $A, B, X \in \mathcal{B}(\mathcal{H})$, it is immediate from polar decompositions that $\mu(AXB) = \mu(|A| \cdot X \cdot |B^*|)$. Hence (5.7) with $r = 1/2$ yields

(5.8) $$\mu(AXB) \prec_{w(\log)} \mu(A^*AX)^{1/2}\mu(XBB^*)^{1/2},$$

which is a generalization of Horn's majorization (1.14). Let $r > 0$, $1 < p, q < \infty$ with $1/p + 1/q = 1$, and $\|\cdot\|$ be a symmetric norm. Then by (5.8) and Lemma 2.13(1) we have

$$\| \, |AXB|^r \| \le \| \, |A^*AX|^{rp/2}\|^{1/p}\| \, |XBB^*|^{rq/2}\|^{1/q}.$$

In particular,

$$\| \, |AB|^r \| \le \| \, |A|^{rp}\|^{1/p}\| \, |B|^{rq}\|^{1/q}, \quad \| \, |AXB|^r \|^2 \le \| \, |A^*AX|^r \| \cdot \| \, |XBB^*|^r \|.$$

The last inequality was given in [16].

As is well known, the Golden–Thompson trace inequality fails to hold for three matrices, while an attempt in the case of three matrices was made in [59]. In the rest of this subsection, we discuss log-majorizations of Golden–Thompson type for more than two operators which are commuting except one.

PROPOSITION 5.10. (1) *If $A_1, A_2 \in \mathcal{B}(\mathcal{H})_+$ and $A_1 A_2 = A_2 A_1$, then for every $B \in \mathcal{B}(\mathcal{H})$,*

$$\mu(BA_1 A_2 B^*) = \mu((A_1 A_2)^{1/2} B^* B (A_1 A_2)^{1/2}) \prec_{w(\log)} \mu(A_1 B^* B A_2).$$

(2) *If $A_1, A_2, A_3 \in \mathcal{B}(\mathcal{H})_+$ and $A_i A_j = A_j A_i$, then for every $B \in \mathcal{B}(\mathcal{H})$,*

$$\mu(BA_1 A_2 A_3 B^*) \prec_{w(\log)} \mu(A_1^2 B^* B A_2^2 B^* B A_3^2)^{1/2}.$$

P r o o f. (1) The equality is obvious because $BA_1 A_2 B^* = |(A_1 A_2)^{1/2} B^*|^2$ and $(A_1 A_2)^{1/2} B^* B (A_1 A_2)^{1/2} = |B(A_1 A_2)^{1/2}|^2$.

For the log-majorization, by the antisymmetric tensor technique, it suffices to show that $|A_1 B^* B A_2| \le I$ implies $(A_1 A_2)^{1/2} B^* B (A_1 A_2)^{1/2} \le I$. We may assume that A_2 is invertible. Then $A_2 B^* B A_1^2 B^* B A_2 \le I$ implies $B^* B A_1^2 B^* B \le A_2^{-2}$ and so $A_1 B^* B A_1^2 B^* B A_1 \le (A_1 A_2^{-1})^2$. This gives $A_1 B^* B A_1 \le A_1 A_2^{-1}$ and hence $(A_1 A_2)^{1/2} B^* B (A_1 A_2)^{1/2} \le I$.

(2) By using (1) twice, we have

$$\mu(BA_1A_2A_3B^*) \prec_{w(\log)} \mu(A_2B^*BA_1A_3) = \mu(A_1A_3(A_2B^*B)^*(A_2B^*B)A_1A_3)^{1/2}$$

$$\prec_{w(\log)} \mu(A_1^2B^*BA_2^2B^*BA_3^2)^{1/2}. \quad \blacksquare$$

PROPOSITION 5.11. *Let H_1, H_2, H_3, K be lower-bounded self-adjoint operators on $\mathcal{H}$. Assume that H_1, H_2, H_3 are mutually commuting (i.e. their spectral projections are all commuting). Then*

$$(5.9) \qquad \mu(e^{-((H_1+H_2)\hat{+}K)}) \prec_{w(\log)} \mu(e^{-H_1}e^{-K}e^{-H_2}),$$

$$(5.10) \qquad \mu(e^{-((H_1+H_2+H_3)\hat{+}2K)}) \prec_{w(\log)} \mu(e^{-H_1}e^{-K}e^{-H_2}e^{-K}e^{-H_3}).$$

Hence $\|e^{-(H_1+H_2+H_3)\hat{+}2K)}\| \le \|e^{-H_1}e^{-K}e^{-H_2}e^{-K}e^{-H_3}\|$ for any symmetric norm $\|\cdot\|$.

P r o o f. (5.9) is a special case of (5.10). The commuting assumption guarantees that $H_1 + H_2 + H_3$ is defined as a lower-bounded self-adjoint operator and

$$e^{-r(H_1+H_2+H_3)} = e^{-rH_1}e^{-rH_2}e^{-rH_3}, \qquad r \ge 0.$$

By (3.6) and Proposition 5.10(2) we have for every $0 < r \le 1/2$,

$$\mu((e^{-rK}e^{-r(H_1+H_2+H_3)}e^{-rK})^{1/r}) \prec_{w(\log)} \mu(e^{-K/2}e^{-(H_1+H_2+H_3)/2}e^{-K/2})^2$$

$$\prec_{w(\log)} \mu(e^{-H_1}e^{-K}e^{-H_2}e^{-K}e^{-H_3}).$$

Hence (5.10) follows by letting $r \downarrow 0$ as in the proof of Theorem 3.7. $\blacksquare$

Let $A_1, \ldots, A_4 \in \mathcal{B}(\mathcal{H})_+$ with $A_iA_j = A_jA_i$, and $B \in \mathcal{B}(\mathcal{H})$. Then, using Proposition 5.10(2) twice, we have

$$\mu(BA_1A_2A_3A_4B^*) \prec_{w(\log)} \mu(A_4B^*BA_1^2A_2^2A_3^2B^*BA_4)^{1/2}$$

$$\prec_{w(\log)} \mu(A_1^4B^*BA_4^2B^*BA_2^4B^*BA_4^2B^*BA_3^4)^{1/4}.$$

This implies as Proposition 5.11 that if $H_1, \ldots, H_4, K$ are lower-bounded self-adjoint operators and $H_1, \ldots, H_4$ are mutually commuting, then

$$\mu(e^{-(H_1+\ldots+H_4)\hat{+}4K)}) \prec_{w(\log)} \mu(e^{-H_1}e^{-K}e^{-H_4/2}e^{-K}e^{-H_2}e^{-K}e^{-H_4/2}e^{-K}e^{-H_3}).$$

PROBLEM 5.12. Do (5.9) and (5.10) extend to a greater number of operators? That is, if $H_1, \ldots, H_n, K$ are lower-bounded self-adjoint operators with $H_1, \ldots, H_n$ mutually commuting, then

$$\mu(e^{-((H_1+\ldots+H_n)\hat{+}(n-1)K)}) \prec_{w(\log)} \mu(e^{-H_1}e^{-K}e^{-H_2}\ldots e^{-K}e^{-H_n})?$$

5.3. *Determinant inequalities.* In the final subsection, some determinant inequalities of Golden–Thompson type are given. Let $A \in \mathcal{C}_1(\mathcal{H})$. Following [32], we define the *determinant* $\det(I + A)$ by

$$\det(I + A) = \prod_{k=1}^{\infty}(1 + \lambda_k(A)),$$

where $\lambda_k(A)$ are the eigenvalues of A counted with algebraic multiplicities. The product of the above right-hand side converges because Theorem 3.3 and Proposition 1.3 imply that

$$\sum_{k=1}^{\infty}|\lambda_k(A)| \le \sum_{k=1}^{\infty}\mu_k(A) = \|A\|_1 < \infty.$$

Another equivalent definition is

$$(5.11) \qquad \det(I + A) = \sum_{n=0}^{\infty} \operatorname{tr}(\Lambda^n A).$$

Indeed, note that if $A \in \mathcal{C}_1(\mathcal{H})$ then $\|\Lambda^n A\|_1 \leq \|A\|_1^n / n!$ for $n \in \mathbb{N}$ and

$$\det(I + zA) = \sum_{n=0}^{\infty} z^n \operatorname{tr}(\Lambda^n A), \qquad z \in \mathbb{C},$$

is an entire analytic function. In particular, comparing the coefficients of z, we have Lidskii's theorem:

$$\operatorname{tr} A = \sum_{k=1}^{\infty} \lambda_k(A).$$

It is known that for every $A, B \in \mathcal{C}_1(\mathcal{H})$,

$$(5.12) \qquad |\det(I + A) - \det(I + B)| \leq \|A - B\|_1 \exp(\|A\|_1 + \|B\|_1 + 1),$$

which shows that $A \mapsto \det(I+A)$ is $\|\cdot\|_1$-continuous on $\mathcal{C}_1(\mathcal{H})$. See [79, §3], [74, §XIII.17], or [31, VII] for details of the facts on the determinant mentioned above.

LEMMA 5.13. *Let* $a = (a_1, a_2, \ldots)$ *and* $b = (b_1, b_2, \ldots)$ *be such that* $a_1 \geq a_2 \geq \ldots \geq 0$ *and* $b_1 \geq b_2 \geq \ldots \geq 0$. *If* $a \prec_{w(\log)} b$, *then* $\prod_{k=1}^{\infty}(1 + a_k) \leq \prod_{k=1}^{\infty}(1 + b_k)$.

P r o o f. Since $\log(1+x)$ is increasing on $[0, \infty)$ and $\log(1 + e^x)$ is convex, Proposition 1.3 implies that $(\log(1 + a_i)) \prec_w (\log(1 + b_i))$ and hence $\prod_{i=1}^{k}(1 + a_i) \leq \prod_{i=1}^{k}(1 + b_i)$ for every $k \in \mathbb{N}$. This gives the result. ∎

PROPOSITION 5.14. *Let* H *and* K *be lower-bounded self-adjoint operators on* $\mathcal{H}$. *Assume that* $e^{-K} \in \mathcal{C}_1(\mathcal{H})$. *Then*

$$(5.13) \qquad \det(I + e^{-(H \hat{+} K)}) \leq \det(I + (e^{-rH/2} e^{-rK} e^{-rH/2})^{1/r}), \qquad r > 0,$$

and the above right-hand side decreases as $r \downarrow 0$. *In particular*,

$$(5.14) \qquad \det(I + e^{-(H \hat{+} K)}) \leq \det(I + e^{-H/2} e^{-K} e^{-H/2}) = \det(I + e^{-H} e^{-K}).$$

Moreover, if $H + K$ *is essentially self-adjoint, then the right-hand side of* (5.13) *decreases to* $\det(I + e^{-(H+K)})$ *as* $r \downarrow 0$.

P r o o f. Note as remarked after Theorem 3.12 that $(e^{-rH/2} e^{-rK} e^{-rH/2})^{1/r} \in \mathcal{C}_1(\mathcal{H})$ for all $r > 0$. The first assertion is immediate from Theorems 3.4 and 3.7 due to Lemma 5.13. The equality in (5.14) is seen because $e^{-H/2} e^{-K} e^{-H/2}$ and $e^{-H} e^{-K}$ have the same eigenvalues including algebraic multiplicities. (This is seen also via (5.11) and Lemma 3.1.) The last assertion is a consequence of (3.12) and (5.12). ∎

PROPOSITION 5.15. *Let* σ *be an operator mean such that* $\sigma \leq \#_\alpha$ *where* $0 \leq \alpha \leq 1$. *Let* H *and* K *be lower-bounded self-adjoint operators. Assume that* $H + K$ *is essentially self-adjoint and* $e^{-K} \in \mathcal{C}_1(\mathcal{H})$. *Then*

$$\det(I + (e^{-rH/(1-\alpha)} \sigma e^{-rK/\alpha})^{1/r}) \leq \det(I + e^{-(H+K)}), \qquad r > 0.$$

Moreover, if H *is bounded, then the above left-hand side converges to* $\det(I + e^{-(H+K)})$ *as* $r \downarrow 0$.

P r o o f. Let $H = \int_a^\infty \lambda dE_H(\lambda)$ and $H_n = \int_a^n \lambda dE_H(\lambda)$. For each $r > 0$ we get

$$\det(I + (e^{-rH/(1-\alpha)}\,\sigma\,e^{-rK/\alpha})^{1/r}) \leq \det(I + (e^{-rH_n/(1-\alpha)}\,\#_\alpha\,e^{-rK/\alpha})^{1/r})$$
$$\leq \det(I + e^{-(H_n+K)})$$

by Theorem 4.9 and Lemma 5.13. So the desired inequality follows because $\det(I + e^{-(H_n+K)}) \to \det(I + e^{-(H+K)})$ due to (3.22) and (5.12). If H is bounded, then Theorem 4.11 and Lemma 3.8 show that

$$\prod_{k=1}^{n} \mu_k(I + e^{-(H+K)}) \leq \liminf_{r\downarrow 0} \prod_{k=1}^{n} \mu_k(I + (e^{-rH/(1-\alpha)}\,\sigma\,e^{-rK/\alpha})^{1/r})$$
$$\leq \liminf_{r\downarrow 0} \det(I + (e^{-rH/(1-\alpha)}\,\sigma\,e^{-rK/\alpha})^{1/r}).$$

Letting $n \to \infty$ we have the second assertion. ∎

The final result is restricted to the matrix case. Further determinant inequalities for matrices are found in [6].

PROPOSITION 5.16. *Let T be any $n \times n$ matrix.*

(1) $\det(I + |e^T|) \leq \det(I + e^{\mathrm{Re}\,T})$.
(2) $\det(\log|e^T|) \geq \det(\mathrm{Re}\,T)$ *if* $\mathrm{Re}\,T \geq 0$.

P r o o f. (1) follows from Proposition 5.6(1) and Lemma 5.13.
(2) For any matrix T, since

$$\det|e^T| = |e^{\mathrm{tr}\,T}| = e^{\mathrm{tr}(\mathrm{Re}\,T)} = \det e^{\mathrm{Re}\,T},$$

Proposition 5.6(1) is improved to $\mu(e^T) \prec_{(\log)} \mu(e^{\mathrm{Re}\,T})$. Since $|e^T| \geq I$ by Proposition 5.6(2), this means that $\mu(\log|e^T|) \prec \mu(\mathrm{Re}\,T)$. For any $\varepsilon > 0$, applying Proposition 1.1(iii) to $f(x) = -\log(x + \varepsilon)$ we have

$$-\sum_{i=1}^{n} \log\{\mu_i(\log|e^T|) + \varepsilon\} \leq -\sum_{i=1}^{n} \log\{\mu_i(\mathrm{Re}\,T) + \varepsilon\},$$

so that

$$\prod_{i=1}^{n} \{\mu_i(\log|e^T|) + \varepsilon\} \geq \prod_{i=1}^{n} \{\mu_i(\mathrm{Re}\,T) + \varepsilon\}.$$

Letting $\varepsilon \downarrow 0$ yields the required inequality. ∎

PROBLEM 5.17. Is it possible to properly extend the determinant inequalities of Proposition 5.16 to operators? Concerning (1), both sides are equal to ∞ for infinite-dimensional $T \in \mathcal{B}(\mathcal{H})$, so that we have to consider unbounded T.

References

[1] T. A n d o, *Topics on Operator Inequalities*, Lecture notes (mimeographed), Hokkaido Univ., Sapporo, 1978.
[2] —, *On some operator inequalities*, Math. Ann. 279 (1987), 157–159.
[3] —, *Comparison of norms $\|\|f(A) - f(B)\|\|$ and $\|\|f(|A - B|)\|\|$*, Math. Z. 197 (1988), 403–409.

[4] T. Ando, *Majorization, doubly stochastic matrices and comparison of eigenvalues*, Linear Algebra Appl. 118 (1989), 163–248.

[5] —, *Majorizations and inequalities in matrix theory*, ibid. 199 (1994), 17–67.

[6] T. Ando and F. Hiai, *Log majorization and complementary Golden–Thompson type inequalities*, ibid. 197/198 (1994), 113–131.

[7] H. Araki, *Relative Hamiltonian for faithful normal states of a von Neumann algebra*, Publ. Res. Inst. Math. Sci. 9 (1973), 165–209.

[8] —, *Golden–Thompson and Peierls-Bogolubov inequalities for a general von Neumann algebras*, Comm. Math. Phys. 34 (1973), 167–178.

[9] —, *Some properties of modular conjugation operator of von Neumann algebras and a non-commutative Radon-Nikodym theorem with a chain rule*, Pacific J. Math. 50 (1974), 309–354.

[10] —, *Relative entropy of states of von Neumann algebras*, Publ. Res. Inst. Math. Sci. 11 (1976), 809–833.

[11] —, *Relative entropy for states of von Neumann algebras II*, ibid. 13 (1977), 173–192.

[12] —, *On an inequality of Lieb and Thirring*, Lett. Math. Phys. 19 (1990), 167–170.

[13] B. Beauzamy, *Introduction to Banach Spaces and Their Geometry*, North-Holland, Amsterdam, 1982.

[14] D. S. Bernstein, *Inequalities for matrix exponentials*, SIAM J. Matrix Anal. Appl. 9 (1988), 156–158.

[15] R. Bhatia, *Perturbation Bounds for Matrix Eigenvalues*, Pitman Res. Notes in Math. Ser. 162, Longman, Harlow, 1987.

[16] R. Bhatia and C. Davis, *A Cauchy–Schwarz inequality for operators with applications*, Linear Algebra Appl. 223/224 (1995), 119–129.

[17] G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Univ. Nac. Tucuman Rev. Ser. A 5 (1946), 147–151.

[18] M. Sh. Birman, L. S. Koplienko, and M. Z. Solomyak, *Estimates for the spectrum of the difference between fractional powers of two self-adjoint operators*, Soviet Math. (Iz. VUZ) 19, no. 3 (1975), 1–6.

[19] P. R. Chernoff, *Note on product formulas for operator semigroups*, J. Funct. Anal. 2 (1968), 238–242.

[20] J. Dixmier, *Formes linéaires sur un anneau d'opérateurs*, Bull. Soc. Math. France 81 (1953), 9–39.

[21] P. G. Dodds, T. K. Dodds, and B. de Pagter, *Non-commutative Banach function spaces*, Math. Z. 201 (1989), 583–597.

[22] —, —, —, *Non-commutative Köthe duality*, Trans. Amer. Math. Soc. 339 (1993), 717–750.

[23] M. J. Donald, *Relative hamiltonians which are not bounded from above*, J. Funct. Anal. 91 (1990), 143–173.

[24] W. Donoghue, *Monotone Matrix Functions and Analytic Continuation*, Springer, New York, 1974.

[25] T. Fack, *Sur la notion de valeur caractéristique*, J. Operator Theory 7 (1982), 307–333.

[26] T. Fack and H. Kosaki, *Generalized s-numbers of τ-measurable operators*, Pacific J. Math. 123 (1986), 269–300.

[27] M. Fujii, T. Furuta, and E. Kamei, *Furuta's inequality and its application to Ando's theorem*, Linear Algebra Appl. 179 (1993), 161–169.

[28] T. Furuta, $A \geq B \geq 0$ *assures* $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$ *for* $r \geq 0$, $p \geq 0$, $q \geq 1$ *with* $(1 + 2r)q \geq p + 2r$, Proc. Amer. Math. Soc. 101 (1987), 85–88.

[29] T. Furuta, *Applications of order preserving operator inequalities*, in: Operator Theory: Adv. Appl. 59, Birkhäuser, Basel, 1992, 180–190.

[30] —, *Extension of the Furuta inequality and Ando-Hiai log-majorization*, Linear Algebra Appl. 219 (1995), 139–155.

[31] I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Classes of Linear Operators I*, Birkhäuser, Basel, 1990.

[32] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Trnansl. Math. Monographs 18, Amer. Math. Soc., 1969.

[33] S. Golden, *Lower bounds for Helmholtz function*, Phys. Rev. 137 (1965), B1127–B1128.

[34] U. Haagerup, *The standard form of von Neumann algebras*, Math. Scand. 37 (1975), 271–283.

[35] —, *L^p-spaces associated with an arbitrary von Neumann algebra*, in: Colloq. Internat. CNRS 274, 1979, 175–184.

[36] F. Hansen and G. K. Pedersen, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. 258 (1982), 229–241.

[37] B. Helffer, *Around the transfer operator and the Trotter–Kato formula*, in: Operator Theory: Adv. Appl. 78, Birkhäuser, Basel, 1995, 161–174.

[38] F. Hiai, *Majorization and stochastic maps in von Neumann algebras*, J. Math. Anal. Appl. 127 (1987), 18–48.

[39] —, *Spectral majorization between normal operators in von Neumann algebras*, in: Operator Algebras and Operator Theory, W. B. Arveson *et al.* (eds.), Pitman Res. Notes in Math. Ser. 271, Longman, Harlow, 1992, 78–115.

[40] —, *Trace norm convergence of exponential product formula*, Lett. Math. Phys. 33 (1995), 147–158.

[41] F. Hiai and Y. Nakamura, *Majorization for generalized s-numbers in semifinite von Neumann algebras*, Math. Z. 195 (1987), 17–27.

[42] —, —, *Distance between unitary orbits in von Neumann algebras*, Pacific J. Math. 138 (1989), 259–294.

[43] —, —, *Closed convex hulls of unitary orbits in von Neumann algebras*, Trans. Amer. Math. Soc. 323 (1991), 1–38.

[44] F. Hiai and D. Petz, *The Golden–Thompson trace inequality is complemented*, Linear Algebra Appl. 181 (1993), 153–185.

[45] M. Hilsum, *Les espaces L^p d'une algèbre de von Neumann définies par la derivée spatiale*, J. Funct. Anal. 40 (1981), 151–169.

[46] T. Ichinose and H. Tamura, *Error estimate in operator norm for Trotter–Kato product formula*, Integral Equations Operator Theory 27 (1997), 195–207.

[47] —, —, *Error bound in trace norm for Trotter–Kato product formula of Gibbs semigroups*, preprint, 1996.

[48] E. Kamei, *Double stochasticity in finite factors*, Math. Japon. 29 (1984), 903–907.

[49] —, *An order on statistical operators implicitly introduced by von Neumann*, Math. Japon. 30 (1985), 891–895.

[50] T. Kato, *A generalization of the Heinz inequality*, Proc. Japan Acad. 37 (1961), 305–308.

[51] —, *Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups*, in: Topics in Functional Analysis, I. Gohberg and M. Kac (eds.), Adv. Math. Suppl. Stud. 3, Academic Press, New York, 1978, 185–195.

[52] F. Kittaneh, *Norm inequalities for fractional powers of positive operators*, Lett. Math. Phys. 27 (1993), 279–285.

[53]   H. K o s a k i, *Non-commutative Lorentz spaces associated with a semi-finite von Neumann algebra and applications*, Proc. Japan Acad. Ser. A Math. Sci. 57 (1981), 303–306.

[54]   —, *Relative entropy of states: a variational expression*, J. Operator Theory 16 (1986), 335–348.

[55]   —, *An inequality of Araki–Lieb–Thirring (von Neumann algebra case)*, Proc. Amer. Math. Soc. 114 (1992), 477–481.

[56]   —, private communication, 1995.

[57]   F. K u b o and T. A n d o, *Means of positive linear operators*, Math. Ann. 246 (1980), 205–224.

[58]   A. L e n a r d, *Generalization of the Golden–Thompson inequality* $\mathrm{Tr}(e^A e^B) \geq \mathrm{Tr}\, e^{A+B}$, Indiana Univ. Math. J. 21 (1971), 457–467.

[59]   E. H. L i e b, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Adv. in Math. 11 (1973), 267–288.

[60]   E. H. L i e b and W. T h i r r i n g, in: Studies in Mathematical Physics, E. H. Lieb, B. Simon, and A. S. Wightman (eds.), Princeton Univ. Press, Princeton, 1976, 301–302.

[61]   A. S. M a r k u s, *The eigen- and singular values of the sum and product of linear operators*, Russian Math. Surveys 19 (4) (1964), 91–120.

[62]   A. W. M a r s h a l l and I. O l k i n, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[63]   L. M i r s k y, *Symmetric gauge functions and unitarily invariant norms*, Quart. J. Math. Oxford Ser. (2) 11 (1960), 50–59.

[64]   Y. N a k a m u r a, *An inequality for generalized s-numbers*, Integral Equations Operator Theory 10 (1987), 140–145.

[65]   H. N e i d h a r d t and V. A. Z a g r e b n o v, *The Trotter–Kato product formula for Gibbs semigroups*, Comm. Math. Phys. 131 (1990), 333–346.

[66]   E. N e l s o n, *Notes on non-commutative integration*, J. Funct. Anal. 15 (1974), 103–116.

[67]   J. v o n N e u m a n n, *Some matrix-inequalities and metrization of matrix-space*, Tomsk Univ. Rev. 1 (1937), 286–299 (Collected Works, Vol. IV, 205–218).

[68]   M. O h y a and D. P e t z, *Quantum Entropy and Its Use*, Springer, Berlin, 1993.

[69]   D. P e t z, *Spectral scale of self-adjoint operators and trace inequalities*, J. Math. Anal. Appl. 109 (1985), 74–82.

[70]   —, *A variational expression for the relative entropy*, Comm. Math. Phys. 114 (1988), 345–349.

[71]   —, *A survey of certain trace inequalities*, in: Functional Analysis and Operator Theory, Banach Center Publ. 30, Inst. Math., Polish Acad. Sci., Warszawa, 1994, 287–298.

[72]   R. T. P o w e r s and E. S t ø r m e r, *Free states of the canonical anticommutation relations*, Comm. Math. Phys. 16 (1970), 1–33.

[73]   M. R e e d and B. S i m o n, *Methods of Modern Mathematical Physics II*, Academic Press, New York, 1975.

[74]   —, —, *Methods of Modern Mathematical Physics IV*, Academic Press, New York, 1978.

[75]   D. L. R o g a v a, *Error bounds for Trotter-type formulas for self-adjoint operators*, Functional Anal. Appl. 27 (1993), 217–219.

[76]   M. B. R u s k a i, *Inequalities for traces on von Neumann algebras*, Comm. Math. Phys. 26 (1972), 280–289.

[77]   R. S c h a t t e n, *Norm Ideals of Completely Continuous Operators*, Springer, Berlin, 1970.

[78]   I. S e g a l, *A non-commutative extension of abstract integration*, Ann. of Math. (2) 57 (1953), 401–457.

[79] B. Simon, *Trace Ideals and Their Applications*, London Math. Soc. Lecture Note Ser. 35, Cambridge Univ. Press, Cambridge, 1979.

[80] S. Strătilă, *Modular Theory in Operator Algebras*, Editura Academiei and Abacus Press, Tunbridge Wells, 1981.

[81] M. Suzuki, *Quantum statistical Monte Carlo methods and applications to spin systems*, J. Statist. Phys. 43 (1986), 883–909.

[82] —, *Convergence of general decompositions of exponential operators*, Comm. Math. Phys. 163 (1994), 491–508.

[83] K. Symanzik, *Proof and refinements of an inequality of Feynman*, J. Math. Phys. 6 (1965), 1155–1156.

[84] C. J. Thompson, *Inequality with applications in statistical mechanics*, ibid. 6 (1965), 1812–1813.

[85] —, *Inequalities and partial orders on matrix spaces*, Indiana Univ. Math. J. 21 (1971), 469–480.

[86] A. Uhlmann, *Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory*, Comm. Math. Phys. 54 (1977), 21–32.

[87] H. Umegaki, *Conditional expectation in an operator algebra, IV (entropy and information)*, Kōdai Math. Sem. Rep. 14 (1962), 59–85.

[88] B.-Y. Wang and M.-P. Gong, *Some eigenvalue inequalities for positive semidefinite matrix power products*, Linear Algebra Appl. 184 (1993), 249–260.

[89] K. Watanabe, *Some results on non-commutative Banach function spaces*, Math. Z. 210 (1992), 555–572.

[90] —, *Some results on non-commutative Banach function spaces II (infinite cases)*, Hokkaido Math. J. 22 (1993), 349–364.

Editorial note: See also M. Fiedler and V. Pták, *A new positive definite geometric mean of two positive definite matrices*, Linear Algebra Appl. 251 (1997), 1–20; W. Pusz and S. L. Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys. 8 (1975), 159–170; and R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.