Majorizations and Inequalities in Matrix Theory

T. Ando*
Research Institute for Electronic Science
Hokkaido University
Sapporo 060, Japan

Submitted by Richard A. Brualdi

Dedicated to Professor I. Olkin on his 70th birthday.

ABSTRACT

In matrix theory, majorization plays a significant role. For instance, majorization relations among eigenvalues and singular values of matrices produce a lot of norm inequalities and even matrix inequalities. This survey article is intended as a review of recent results in matrix theory related to majorization.

INTRODUCTION

My aim is to give a brief survey of results related to majorization in matrix theory since the appearance in 1979 of the monumental book *Inequalities: Theory of Majorization and its Applications* by W. Marshall and I. Olkin, which will be cited as [MO].

In 1981 I delivered a lecture of similar nature with the title "Majorization, doubly stochastic matrices and comparison of eigenvalues of matrices," which was published later as Ando (1989), and will be cited as [A].

As the area to be covered is vast, I have to confine myself to the field of my own interest. Therefore main emphasis is placed on majorization related to eigenvalues and singular values, matrix inequalities, and norm inequalities.

*Research supported in part by Grant-in-Aid for Scientific Research.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

The survey consists of eight sections:

1. Majorization for sequences
2. Majorization for matrices
3. Matrix means
4. Matrix inequalities
5. Log majorization
6. Spectral perturbation
7. Hadamard products
8. Majorizations in von Neumann algebras

1. MAJORIZATION FOR SEQUENCES

Recall that for a pair of real vectors $a = [a_i]$, $b = [b_i] \in \mathbb{R}^n$ the *majorization* relation $a \succ b$ means that

$$\sum_{i=1}^{k} a_{[i]} \geqslant \sum_{i=1}^{k} b_{[i]} \qquad (k = 1, 2, \dots, n) \tag{1.1}$$

and

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i, \tag{1.2}$$

where $a_{[1]} \geqslant a_{[2]} \geqslant \cdots \geqslant a_{[n]}$ is the decreasing rearrangement of the components of the vector a. When the last equality condition (1.2) is not required, a is said to *weakly majorize* or *submajorize* b, and this weak relation is denoted by $a \succ_w b$. Note that in [A] weak majorization $\succ_w$ is denoted by $\succ\!\!\!\!\cdot\,$.

Both $\succ$ and $\succ_w$ introduce pseudo-orders in $\mathbb{R}^n$. Remark that $a \geqslant b$ means $a_i \geqslant b_i$ $(i = 1, 2, \dots, n)$. Then obviously $a \geqslant b$ implies $a \succ_w b$.

The basic quantity $\sum_{i=1}^{k} a_{[i]}$ for $a \in \mathbb{R}^n$ can be written in the form

$$\sum_{i=1}^{k} a_{[i]} = \max_{|J|=k} \sum_{i \in J} a_i, \tag{1.3}$$

where J is a subset of $\{1, 2, \dots, n\}$ and $|J|$ denotes its cardinality.

For a complex vector $a = [a_i] \in \mathbb{C}^n$, denote by $|a|$ the vector $[|a_i|]$. Besides the formulas

$$\sum_{i=1}^{k} |a|_{[i]} = \max_{|J|=k} \sum_{i \in J} |a_i|, \tag{1.4}$$

one has

$$\sum_{i=1}^{k} |a|_{[i]} = \min\left\{ \sum_{j=1}^{n} |b_j| + k \max_{1 \leqslant i \leqslant n} |a_i - b_i| : b \in \mathbb{C}^n \right\}. \tag{1.5}$$

Remark that when $|a_1| \geqslant |a_2| \geqslant \cdots \geqslant |a_n|$, the minimum in (1.5) is attained at b defined by

$$b_i = \begin{cases} \operatorname{sgn}(a_i) \cdot (|a_i| - |a_k|) & (i = 1, 2, \ldots, k), \\ 0 & (i = k+1, \ldots, n). \end{cases} \tag{1.6}$$

Alberti and Uhlmann (1982) pointed out that if there is a such that $a \succ b$ for all b in a bounded subset $\mathscr{S}$ of $\mathbb{R}^n$, then among all those a's there is a minimum $\hat{a}$ in the sense of $\succ$. With the additional requirement that $\hat{a}_1 \geqslant \hat{a}_2 \geqslant \cdots \geqslant \hat{a}_n$ this minimum is uniquely determined.

In this connection Ando and Nakamura (1991), analyzing the approach of Li and Tsing (1989) in the proof of (2.21), showed that given a, $b \in \mathbb{R}^n$ the set $\{a - c : b \succ c\}$ has a minimum element in the sense of $\succ$: there is $\hat{b}$ such that $b \succ \hat{b}$ and $a - c \succ a - \hat{b}$ for all c for which $b \succ c$.

A corresponding result for weak majorization was pointed out by Bapat (1991): given a bounded subset $\mathscr{S}$ of $\mathbb{R}^n_+$, there is a unique $\hat{a} \in \mathbb{R}^n_+$ with $\hat{a}_1 \geqslant \hat{a}_2 \geqslant \cdots \geqslant \hat{a}_n$ such that $\hat{a} \succ_w b$ for all $b \in \mathscr{S}$, and $a \succ_w \hat{a}$ whenever $a \in \mathbb{R}^n_+$, and $a \succ_w b$ for all $b \in \mathscr{S}$. In fact, this $\hat{a}$ is determined successively by the following formulas:

$$\hat{a}_1 = \min\left\{a_{[1]} : a \succ_w b \ (b \in \mathscr{S})\right\}$$

and

$$\hat{a}_k = \min\left\{ \sum_{i=1}^{k} a_{[i]} : a \succ_w b \ (b \in \mathscr{S}) \right\} - \sum_{i=1}^{k-1} \hat{a}_i \qquad (k = 2, 3, \ldots, n). \tag{1.7}$$

This $\hat{a}$ will be denoted by $\sqcup_w \mathscr{S}$.

A decisive role in majorization theory is played by the theorem of Hardy, Littlewood, and Pólya that $a = [a_i] \succ b = [b_i]$ if and only if $b = Da$ for some doubly stochastic matrix D. Recall here that a matrix is called *doubly stochastic* if all its entries are nonnegative and all its row sums and column sums are equal to 1. If the requirements in the last part are only for row sums (column sums), then the matrix is called *row-stochastic* (*column-stochastic*). Correspondingly $a \succ_w b$ is characterized by the existence of a doubly stochastic matrix D for which $Da \geqslant b$ (see [MO, p. 27] and [A, p. 198]).

When a continuous function $f(t)$ is defined on a region containing all components of a vector $a = [a_i] \in \mathbb{C}^n$, let us write $f(a) \equiv [f(a_i)]$. Then for a vector $a \in \mathbb{R}^n$ and a row-stochastic matrix D one has

$$Df(a) \geqslant f(Da) \qquad \text{for continuous convex } f. \tag{1.8}$$

Therefore it follows from the Hardy-Littlewood-Pólya theorem that

$$a \succ b \quad \text{implies} \quad f(a) \succ_w f(b) \quad \text{for continuous convex } f, \tag{1.9}$$

and

$$\begin{gathered} a \succ_w b \quad \text{implies} \\ f(a) \succ_w f(b) \quad \text{for continuous convex, nondecreasing } f. \end{gathered} \tag{1.10}$$

When S is an $n \times n$ matrix, denote by $\eta(S)$ the vector

$$\eta(S) \equiv \sqcup_w \{|Sa|, |S^*b| : e \geqslant |a|, |b|\}, \tag{1.11}$$

where e is the vector with all components equal to 1. It is clear that if all entries of S are nonnegative then

$$\sqcup_w \{|Sa| : e \geqslant |a|\} = Se;$$

hence

$$\eta(D) = e \qquad \text{for doubly stochastic } D. \tag{1.12}$$

Bapat (1991) showed that, with $\eta = \eta(S)$,

$$\left[\eta_{[i]} \cdot |a|_{[i]}\right] \succ_w |Sa| \qquad (a \in \mathbb{C}^n). \tag{1.13}$$

In particular,

$$|a| \succ_w |Da| \qquad \text{for doubly stochastic } D. \tag{1.14}$$

A norm $\|\cdot\|$ on $\mathbb{C}^n$ is called *permutation-invariant* if for every permutation π of $\{1, 2, \ldots, n\}$

$$\|[a_i]\| = \|[a_{\pi(i)}]\| \qquad (a = [a_i] \in \mathbb{C}^n). \tag{1.15}$$

It is called *absolute* if

$$\|a\| = \||a|\| \qquad (a \in \mathbb{C}^n). \tag{1.16}$$

Such a norm is always *monotone* on $\mathbb{R}^n_+$ in the sense

$$\|a\| \geqslant \|b\| \qquad \text{whenever} \quad a \geqslant b \geqslant 0. \tag{1.17}$$

An absolute, permutation-invariant norm is often called a *symmetric gauge function.*

Among familiar examples of absolute, permutation-invariant norms are

$$\|a\|_p \equiv \left\{\sum_{i=1}^{n} |a_i|^p\right\}^{1/p} \quad (1 \leqslant p < \infty), \qquad \|a\|_\infty \equiv \max_{1 \leqslant i \leqslant n} |a_i|. \tag{1.18}$$

Less well known are

$$\|a\|_{(k)} \equiv \sum_{i=1}^{k} |a|_{[i]} \qquad (k = 1, 2, \ldots, n). \tag{1.19}$$

With this notation (1.5) can be written in the following form

$$\|a\|_{(k)} = \min\{\|b\|_1 + k\|a - b\|_\infty : b \in \mathbb{C}^n\}. \tag{1.20}$$

It follows from Birkhoff's theorem that every doubly stochastic matrix is a convex combination of permutation matrices, so that for every permutation invariant norm $\|\cdot\|$

$$\|a\| \geqslant \|Da\| \qquad \text{for doubly stochastic } D \text{ and } a \in \mathbb{C}^n. \tag{1.21}$$

Further, by the characterization of weak majorization mentioned above, one has for every absolute, permutation-invariant norm $\|\cdot\|$

$$\|a\| \geqslant \|b\| \qquad \text{whenever} \quad \|a\|_{(k)} \geqslant \|b\|_{(k)} \quad (k = 1, 2, \ldots, n). \tag{1.22}$$

Considering a matrix S as a linear map on $\mathbb{C}^n$, one can define its mapping norm (= Lipschitz bound) with respect to a norm $\|\cdot\|$ on $\mathbb{C}^n$. When its mapping norm is not greater than 1, S is called a $\|\cdot\|$-*contraction*. Since

$$\|S\|_{\infty\to\infty}\cdot e \geqslant \sqcup_w \{|Sa| : e \geqslant |a|\}$$

and $\|S\|_{1\to 1} = \|S^*\|_{\infty\to\infty}$ (where, for instance, $\|S\|_{\infty\to\infty}$ is the mapping norm of S with respect to norm $\|\cdot\|_\infty$), it follows from (1.13) that

$$|a| \succ_w |Sa| \qquad \text{whenever } S \text{ is } \|\cdot\|_1\text{-contractive and } \|\cdot\|_\infty\text{-contractive.} \tag{1.23}$$

There are two directions of generalization of the notion of majorization to a pair of finite sequences of real vectors $\{a^{(i)}\}_1^m$, $\{b^{(i)}\}_1^m$. The first direction is along the line of the Hardy-Littlewood-Pólya theorem and requires for simultaneous majorization of $\{a^{(i)}\}_1^m$ over $\{b^{(i)}\}_1^m$ the existence of a doubly stochastic matrix D such that $b^{(i)} = Da^{(i)}$ $(i = 1, 2, \ldots, m)$. There is an extensive study of simultaneous majorization of this type, motivated principally from physics, in the monograph of Alberti and Uhlmann (1982). Simultaneous majorization of $\{b^{(i)}\}$ by $\{a^{(i)}\}$ is characterized in terms of a family of inequalities of the form

$$\Phi(a^{(1)}, \ldots, a^{(m)}) \geqslant \Phi(b^{(1)}, \ldots, b^{(m)}),$$

where $\Phi(\cdots)$ are convex functions of m vector variables.

In the other direction requirement for majorization is that

$$b^{(i)} = \sum_{j=1}^{m} d_{ij} a^{(j)} \qquad \begin{array}{l}(i = 1, 2, \ldots, m)\\ \text{for some } m \times m \text{ doubly stochastic matrix } D = \left[d_{ij}\right].\end{array}$$

This has a close connection with the Choquet theory of simplexes as seen in the monograph of Alfsen (1971). In Fischer and Holbrook (1977, 1980) majorization in this sense is characterized by the condition that

$$\sum_{i=1}^{m} \varphi(a^{(i)}) \geqslant \sum_{i=1}^{m} \varphi(b^{(i)})$$

for all nonnegative convex continuous functions $\varphi(\cdot)$ on $\mathbb{R}^n$. Along similar lines is Komiya (1983).

The case of $m = 2$ is in a special situation. Recall first (see [MO, p. 109] and [A, p. 168]) that $a \succ b$ is characterized by

$$\|a + te\|_1 \geqslant \|b + te\|_1 \qquad (t \in \mathbb{R}). \tag{1.24}$$

This was generalized by Ruch, Schranner, and Seligman (1980): for *stochastic* vectors, that is, $\|\cdot\|_1$ unit vectors with nonnegative components $a^{(i)}$, $b^{(i)}$ $(i = 1, 2)$, there is a column-stochastic matrix S such that $b^{(i)} = Sa^{(i)}$ $(i = 1, 2)$ if and only if

$$\|a^{(1)} + ta^{(2)}\|_1 \geqslant \|b^{(1)} + tb^{(2)}\|_1 \qquad (t \in \mathbb{R}). \tag{1.25}$$

Hässelbarth and Ruch (1993) observed that for general vectors $a^{(i)}$, $b^{(i)}$ $(i = 1, 2)$ there is a $\|\cdot\|_1$ contraction C such that $b^{(i)} = Ca^{(i)}$ $(i = 1, 2)$ if and only if

$$\|sa^{(1)} + ta^{(2)}\|_1 \geqslant \|sb^{(1)} + tb^{(2)}\|_1 \qquad (s, t \in \mathbb{R}). \tag{1.26}$$

The $\|\cdot\|_\infty$ case is valid in a more general setting.

2. MAJORIZATION FOR MATRICES

We take the view that a noncommutative analogue of a complex number is a matrix, say $n \times n$, while an analogue of a real number is a Hermitian matrix and that of a nonnegative number is a positive semidefinite matrix. In this context the conjugate transpose corresponds to the complex conjugate of a number. For a matrix A its *real* (or Hermitian) part $\operatorname{Re} A$ is defined as $\frac{1}{2}(A + A^*)$, while its *imaginary* part $\operatorname{Im} A$ is $(1/2i)(A - A^*)$. The *modulus* $|A|$ of a matrix A is defined as the positive semidefinite square root of A^*A. Let us denote by $\mathbb{M}_n$ the algebra of all $n \times n$ complex matrices.

The *order relation* $A \geqslant B$ for two Hermitian matrices A, B always means that $A - B$ is positive semidefinite. In particular, $A \geqslant 0$ means that A is positive semidefinite. Let us write $A > 0$ to mean that A is positive definite.

All difficulties with respect to this order relation come from the fact that the space of Hermitian matrices does not become a lattice: given two Hermitian matrices A, B, the set $\{X : X \geqslant A \text{ and } X \geqslant B\}$ has no minimum point except when $A \geqslant B$ or $A \leqslant B$. Ando (1993) gave a complete parametrization of all minimal points of this set.

Alberti and Uhlmann (1982) is a useful monograph on the subjects of this section.

To each matrix A two kind of numerical vectors, one in $\mathbb{C}^n$ and the other in $\mathbb{R}^n$, are associated: the one consists of the *eigenvalues* $[\lambda_i(A)]$ of A, and the other of its *singular values* $[\sigma_i(A)]$, where $\sigma_i(A) = \lambda_i(|A|)$ by definition.

When A is Hermitian, all its eigenvalues are real, so let us always arrange them in decreasing order:

$$\lambda_1(A) \geqslant \lambda_2(A) \geqslant \cdots \geqslant \lambda_n(A). \tag{2.1}$$

For two Hermitian matrices A, B, let us write $A \succ B$ or $A \succ_w B$ according as $[\lambda_i(A)] \succ [\lambda_i(B)]$ or $[\lambda_i(A)] \succ_w [\lambda_i(B)]$.

An eigenvalue analogue of the extremal characterization (1.3) for a Hermitian matrix A is the following formula of Ky Fan:

$$\sum_{i=1}^{k} \lambda_i(A) = \max\{\mathrm{Tr}(PA) : P \text{ a Hermitian projection of rank } k\}$$

$$= \max\left\{ \sum_{i=1}^{k} (U^*AU)_{ii} : U \text{ unitary} \right\} \qquad (k = 1, 2, \ldots, n), \tag{2.2}$$

where $(U^*AU)_{ii}$ is the $(i,\ i)$ entry of U^*AU. Immediate consequences of (2.2) are the Schur theorem

$$A \succ \mathrm{diag}(A) \qquad \text{for Hermitian } A, \tag{2.3}$$

and the Ky Fan theorem

$$[\lambda_i(\mathrm{Re}\ A)] \succ [\mathrm{Re}\ \lambda_i(A)] \qquad (A \in \mathbb{M}_n). \tag{2.4}$$

The inverse problems for (2.3) and (2.4) have affirmative answers (see [MO, p. 220]). If $[\lambda_i] \succ [\alpha_i]$, there is a Hermitian matrix $A = [a_{ij}]$ such that

$$\lambda_i(A) = \lambda_i \text{ and } a_{ii} = \alpha_i \qquad (i = 1, 2, \ldots, n).$$

Similarly if $[\alpha_i] \succ [\beta_i]$, there is a matrix A for which

$$\lambda_i(\mathrm{Re}\ A) = \alpha_i \text{ and } \mathrm{Re}\ \lambda_i(A) = \beta_i \qquad (i = 1, 2, \ldots, n)$$

(see [MO, p. 238]).

A practical condition for a non-Hermitian matrix to have real (or even positive) eigenvalues only is total positivity. Recall that a real matrix $A = [a_{ij}]$ is called *totally positive* if all its square submatrices have positive determinants. Ando (1987a) is a comprehensive survey on totally positive matrices.

It had been conjectured that (2.3) is valid also for a totally positive matrix. Garloff (1982, 1985) settled the conjecture affirmatively, and further showed that the inverse problem for this majorization relation is not always true even when $a_{ii} = \text{const}$ $(i = 1, 2, \dots, n)$.

Let A, B be Hermitian matrices. Among vectors of $[\lambda_i(A)]$, $[\lambda_{\pi(i)}(B)]$, and $[\lambda_{\delta(i)}(A+B)]$, π, δ being permutations of $\{1, 2, \dots, n\}$, there are various type of majorization relations. Among the easiest examples is a consequence of Ky Fan's formula (2.2):

$$\big[\lambda_i(A) + \lambda_i(B)\big] \succ \big[\lambda_i(A+B)\big]. \tag{2.5}$$

A deeper result is the celebrated theorem of V. B. Lidskii (the elder) and Wielandt (see [MO, p. 242] and [A, p. 223]):

$$\big[\lambda_i(A+B)\big] \succ \big[\lambda_i(A) + \lambda_{n-i+1}(B)\big]. \tag{2.6}$$

Hersch and Zwahlen (1962) gave an extremal characterization of $\sum_{i=1}^{k} \lambda_{j_i}(A)$ for $1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$, for which Riddle (1984) presented a topological minimax characterization. In Smiley (1966) one can find simple proofs for some of inequalities of the Lidskii-Wielandt type.

Amir-Moéz (1968) and Markus (1964) are excellent surveys of this area, containing original contributions.

Bhatia and Holbrook (1989) showed that if A, B, and $A+B$ are normal, there is a permutation π of $\{1, 2, \dots, n\}$ and a doubly stochastic matrix D such that

$$\big[\lambda_i(A) + \lambda_{\pi_i}(B)\big] = D\big[\lambda_i(A+B)\big].$$

The inverse eigenvalues problem for a sum of Hermitian matrices is to find conditions on three vectors $[a_i]$, $[b_i]$, and $[c_i] \in \mathbb{R}^n$ which guarantee the existence of two Hermitian matrices A, B such that

$$a_i = \lambda_i(A), \quad b_i = \lambda_i(B), \text{ and } c_i = \lambda_i(A+B) \qquad (i = 1, 2, \dots, n).$$

Various inequalities of the Lidskii-Wielandt type give necessary conditions for the solvability of the inverse problem. A serious combinatorial search for sufficient conditions began with A. Horn (1962). Extending this idea, B. V. Lidskii the younger (1982) proposed a complete set of conditions for the inverse problem. But no full proof has been published.

A related problem is the exponential function problem. Remark that a matrix W is unitary if and only if it is of the form $W = e^{iH}$ for a Hermitian matrix H. Therefore, since a product of unitary matrices is again unitary, for any Hermitian matrices A, B there must be a Hermitian matrix C such that

$$e^{iC} = e^{iA}e^{iB}.$$

The problem is whether C is found in the form $C = U^*AU + V^*BV$ for suitable unitary matrices U, V. A close connection with the inverse eigenvalue problem is seen from the relations

$$\lambda_i(U^*AU) = \lambda_i(A) \text{ and } \lambda_i(V^*BV) = \lambda_i(B) \qquad (i = 1, 2, \dots, n).$$

This problem is not fully settled either. R. C. Thompson (1986) proposed a program of reducing the inverse eigenvalue problem to this problem. He showed that the Lidskii conditions can be checked on the basis of Nudelman and Shvartzman (1959), in which the eigenvalues of the product of unitary matrices are investigated.

In accordance with (2.1) let us always arrange the singular values of a matrix A in decreasing order:

$$\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_n(A). \tag{2.7}$$

There are singular value versions of (2.5) and (2.6) for general matrices A, B (see [MO, p. 243] and [A, p. 229]):

$$[\sigma_i(A) + \sigma_i(B)] \succ_w [\sigma_i(A + B)] \tag{2.8}$$

and

$$[\sigma_i(A - B)] \succ_w [|\sigma_i(A) - \sigma_i(B)|]. \tag{2.9}$$

In this connection Gil (1993) proved the following inequalities:

$$2\sum_{i=1}^{k} \left\{\sigma_i(\operatorname{Re} A)^2 - |\operatorname{Re} \lambda_i(A)|^2\right\} \geqslant \sum_{i=1}^{k} \left\{\sigma_{n-i+1}(A)^2 - |\lambda_i(A)|^2\right\}$$

$$(k = 1, 2, \dots, n). \tag{2.10}$$

Recall that a norm $\|\cdot\|$ on $\mathbb{M}_n$ is called *unitarily invariant* if

$$\|A\| = \|UAV\| \qquad \text{for all unitary } U, V. \tag{2.11}$$

It is known that a unitarily invariant norm $\|\cdot\|$ on $\mathbb{M}_n$ stands in one to one correspondence with an absolute, permutation-invariant norm $\|\cdot\|$ on $\mathbb{C}^n$ via

$$\|A\| = \|[\sigma_i(A)]\| \qquad (A \in \mathbb{M}_n). \tag{2.12}$$

Now familiar unitarily invariant norms are produced from (1.18) and (1.19):

$$\begin{aligned} \|A\|_p &\equiv \left\{\sum_{i=1}^{n} \sigma_i(A)^p\right\}^{1/p} = \{\mathrm{Tr}(|A|^p)\}^{1/p} \qquad (1 \leqslant p < \infty), \\ \|A\|_\infty &\equiv \sigma_1(A). \end{aligned} \tag{2.13}$$

The norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ are called the *trace norm*, *Frobenius norm*, and *spectral norm*, respectively. $\|\cdot\|_p$ is generally called the *Schatten p-norm*.

The unitarily invariant norm corresponding to $\|\cdot\|_{(k)}$ is called the *Ky Fan norm*:

$$\|A\|_{(k)} = \sum_{i=1}^{k} \sigma_i(A) \qquad (k = 1, 2, \ldots, n). \tag{2.14}$$

Then the singular value versions of (2.2) and (1.20) hold in the following form:

$$\begin{aligned} \|A\|_{(k)} &= \max\{|\mathrm{Tr}(B^*A)| : B^*B \text{ a projection of rank } k\} \\ &= \max\left\{\sum_{i=1}^{k} |(U^*AV)_{ii}| : U, V \text{ unitary}\right\} \end{aligned} \tag{2.15}$$

and

$$\|A\|_{(k)} = \min\{\|B\|_1 + k \cdot \|A - B\|_\infty : B \in \mathbb{M}_n\}. \tag{2.16}$$

In (2.15) for $k = n$ the maximization can be restricted to $U = V$. See Li (1987).

Now it follows from (1.22) and (2.12) that for every unitarily invariant norm $\|\cdot\|$

$$\|A\| \geqslant \|B\| \qquad \text{whenever} \quad \|A\|_{(k)} \geqslant \|B\|_{(k)} \quad (k = 1, 2, \dots, n). \quad (2.17)$$

A weaker restriction on a norm $\|\cdot\|$ is *unitary similarity invariance*:

$$\|UAU^*\| = \|A\| \qquad \text{for all unitary } U.$$

The *numerical radius* (norm) $w(\cdot)$ is a typical example:

$$w(A) \equiv \sup\{|x^*Ax| : \|x\| \leqslant 1\}. \quad (2.18)$$

For a linear map $\Phi(\cdot)$ from $\mathbb{M}_n$ to $\mathbb{M}_m$, its *adjoint* map $\Phi^*(\cdot)$ from $\mathbb{M}_m$ to $\mathbb{M}_n$ is defined by

$$\operatorname{Tr}(B^* \cdot \Phi(A)) = \operatorname{Tr}(\Phi^*(B)^* \cdot A) \qquad (A \in \mathbb{M}_n,\ B \in \mathbb{M}_m). \quad (2.19)$$

A linear map $\Phi(\cdot)$ is called *positive* if it preserves positive semidefiniteness:

$$\Phi(X) \geqslant 0 \qquad \text{whenever} \quad X \geqslant 0.$$

It is called *unital* if $\Phi(I_n) = I_m$, where I_n is the $n \times n$ identity matrix, and is called *trace-preserving* if

$$\operatorname{Tr} \Phi(X) = \operatorname{Tr} X \qquad \text{for all } X.$$

It is obvious that a linear map is positive if and only if its adjoint is positive, and that it is unital (is trace-preserving) if and only if its adjoint is trace-preserving (is unital).

A positive, unital, and trace-preserving linear map $\Phi(\cdot)$ is called *doubly stochastic*. Then a natural extension of the Hardy-Littlewood-Pólya theorem holds: majorization $A \succ B$ for a pair of Hermitian A, B is equivalent to the existence of a doubly stochastic map $\Phi(\cdot)$ such that $B = \Phi(A)$. Also, for a Hermitian matrix B, its image set under all doubly stochastic maps $\{\Phi(B) : \Phi$ doubly stochastic$\}$ coincides with the convex hull of its *unitary orbit* $\{U^*BU : U$ unitary$\}$. See [A, p. 235].

Li and Tsing (1989) showed that given Hermitian A, B there are doubly stochastic maps $\Phi_1(\cdot)$, $\Phi_2(\cdot)$ such that for all unitary similarity invariant norm $\|\cdot\|$

$$\|A - \Phi_1(B)\| \geqslant \|A - \Phi(B)\| \geqslant \|A - \Phi_2(B)\|$$

$$\text{for all doubly stochastic } \Phi. \tag{2.20}$$

Here $\Phi_1(\cdot)$ is easy to construct on the basis of (2.6); the difficulty in constructing $\Phi_2(\cdot)$ is overcome by the minimum theorem of Ando and Nakamura (1991), mentioned in Section 1.

As an extension of (1.11) to a linear map $\Phi(\cdot)$ on $\mathbb{M}_n$, let us define $\eta(\Phi)$ by

$$\eta(\Phi) \equiv \sqcup_w \{[\sigma_i(\Phi(A))], [\sigma_i(\Phi^*(B))] : \|A\|_\infty, \|B\|_\infty \leqslant 1\}. \tag{2.21}$$

If $\Phi(\cdot)$ is positive, then

$$\sqcup_w \{[\sigma_i(\Phi(A))] : \|A\|_\infty \leqslant 1\} = [\lambda_i(\Phi(I))]. \tag{2.22}$$

As a consequence one has

$$\eta(\Phi) = e \qquad \text{for doubly stochastic } \Phi. \tag{2.23}$$

From (2.16) one can derive, as in Bapat (1991) (see also Bapat (1987, 1989)), an extension of (1.13): for a linear map Φ with $\eta = \eta(\Phi(A))$,

$$[\eta_{[i]} \cdot \sigma_i(A)] \succ_w [\sigma_i(\Phi(A))] \qquad (A \in \mathbb{M}_n). \tag{2.24}$$

As a consequence one has

$$|A| \succ_w |\Phi(A)| \qquad \text{for doubly stochastic } \Phi. \tag{2.25}$$

The majorization (2.25) is valid even when $\Phi(\cdot)$ is $\|\cdot\|_\infty$-contractive and $\|\cdot\|_1$-contractive.

Finally let us mention that

$$w(A) \geqslant w(\Phi(A)) \qquad \text{for doubly stochastic } \Phi \text{ and } A \in \mathbb{M}_n. \tag{2.26}$$

The notions and results mentioned above can be extended to the case of a compact linear operator on a Hilbert space, because such an operator has discrete eigenvalues and is approximated by finite rank operators.

Gohberg and Krein (1965) and Simon (1979) are excellent surveys on unitarily invariant norms as well as majorization relations for Hilbert space operators.

3. MATRIX MEANS

Let A, B be $n \times n$ Hermitian matrices. If $A \geqslant B$, then $\lambda_i(A) \geqslant \lambda_i(B)$ $(i = 1, 2, \dots, n)$. If a real valued continuous function $f(t)$ is defined and nondecreasing on an open or closed, finite or infinite interval containing all eigenvalues of A and B, then obviously

$$\lambda_i(f(A)) = f(\lambda_i(A)) \geqslant f(\lambda_i(B)) = \lambda_i(f(B)) \qquad (i = 1, 2, \dots, n).$$

Therefore $U^*f(A)U \geqslant f(B)$ for some unitary U. But the order relation $f(A) \geqslant f(B)$ does not hold in general.

If

$$f(A) \geqslant f(B) \qquad \text{whenever} \quad A, B \in \mathbb{M}_n \text{ and } A \geqslant B, \tag{3.1}$$

then the function $f(t)$ is called *matrix-monotone of order n* on the interval. Correspondingly *matrix convexity of order n* is defined by the requirement

$$\alpha f(A) + (1 - \alpha) f(B) \geqslant f(\alpha A + (1 - \alpha) B) \qquad (0 < \alpha < 1). \tag{3.2}$$

Further, $f(t)$ is called *matrix-concave of order n* if $-f(t)$ is matrix-convex of order n.

A function which is matrix-monotone of all orders is called *operator-monotone*. Correspondingly *operator convexity* and *operator concavity* are defined.

Remark that for an operator-monotone function $f(t)$ the inequality (3.1) can be extended to Hermitian operators A, B on Hilbert space. The same is true for (3.2) with an operator-convex function.

According to the celebrated theorem of Loewner (see [MO, p. 464]), a function $f(t)$ on an open interval Δ is operator-monotone if and only if it admits an analytic continuation to the open upper half plane and it transforms the half plane into itself. Based on the Nevanlinna integral representation of an analytic function which is defined on the open upper half plane and transforms it into itself, an operator-monotone function on an interval Δ is characterized as the one that admits an integral representation

$$f(t) = a + bt + \int_{(-\infty,\infty)\setminus\Delta}\left(\frac{1}{s-t} - \frac{s}{1+s^2}\right) d\mu(s) \qquad (t \in \Delta), \quad (3.3)$$

where $a \in \mathbb{R}$, $b \geqslant 0$, and $\mu(\cdot)$ is a positive measure on $(-\infty, \infty) \setminus \Delta$ such that $\int(1 + s^2)^{-1}\, d\mu(s) < \infty$. Here a, b, and $\mu(\cdot)$ are determined uniquely.

Let us mention some important examples. The fractional power t^α is operator-monotone on $[0, \infty)$ for $0 < \alpha \leqslant 1$ but not for $\alpha > 1$. For $1 \leqslant p \leqslant 2$ the function t^p is operator-convex, but not for other positive exponents. The logarithm $\log t$ is operator-monotone on $(0, \infty)$, but e^t is not operator-monotone on any interval of $\mathbb{R}$.

Nonnegative operator-monotone functions on $[0, \infty)$ have been especially studied in connection with unitarily invariant norms. In this case, after suitable substitution, the integral representation (3.3) is converted to the following form:

$$f(t) = \alpha + \beta t + \int_0^\infty \frac{t}{s+t}\, d\nu(s) \qquad (t \geqslant 0), \qquad (3.4)$$

where α, $\beta \geqslant 0$ and $\nu(\cdot)$ is a positive measure on $(0, \infty)$ such that $\int(1 + s)^{-1}\, d\nu(s) < \infty$. Furthermore, a, b and $\nu(\cdot)$ are determined uniquely from $f(t)$.

This formula shows that in the cone of nonnegative operator-monotone functions on $[0, \infty)$, the constant functions, the scalar multiples of the function t, and the functions of the form $\alpha t/(\alpha + t)$ constitute extremal rays, and every operator monotone function is a continuous weighted average of those extremal functions. Therefore various matrix inequalities related to operator-monotone functions on $[0, \infty)$ are reduced to the case of such extremal functions.

Donoghue (1974) is a monograph devoted to operator-monotone functions from the standpoint of analytic extensions. Davis (1963) and Ando (1979a) give compact surveys of the basic facts on operator-monotone functions and related matrix inequalities from the standpoint of operator theory. Various operations in the class of operator-monotone functions have been studied by Nakamura (1989).

As is shown in Ando (1979a), an operator-monotone function on $[0, \infty)$ is necessarily operator-concave. But this can be proved without appeal to the integral representation. In fact, Davis (1963) proved the *pinching inequality* for operator-monotone $f(t)$,

$$f(PAP) \geqslant Pf(A)P \qquad \text{for a Hermitian projection } P. \tag{3.5}$$

Hansen (1980) showed that the pinching inequality can be generalized to the form

$$f(X^*AX) \geqslant X^*f(A)X \qquad \text{whenever} \quad I \geqslant X^*X \tag{3.6}$$

and that this is indeed equivalent to operator concavity. His method shows in essence that a matrix-monotone function of order n on $(0, \infty)$ is matrix-concave of order $[n/2]$, as observed by Mathias (1990).

In a similar line Hansen and Pedersen (1981/82) showed equivalences of the following conditions for a continuous real valued function $f(t)$ on $[0, 1)$:

(1) $f(t)$ is operator-convex and $0 \geqslant f(0)$;

(2) $X^*f(A)X \geqslant f(X^*AX)$ for all A with $0 \leqslant A \leqslant I$ and X with $X^*X \leqslant I$;

(3) $X^*f(A)X + Y^*f(B)Y \geqslant f(X^*AX + Y^*BY)$ for A, B with $0 \leqslant A$, $B \leqslant I$ and X, Y with $X^*X + Y^*Y \leqslant I$;

(4) $Pf(A)P \geqslant f(PAP)$ for A with $0 \leqslant A \leqslant I$ and Hermitian projection P.

An observation of this type is also in Friedland and Katz (1987).

Using integral representation (3.3), Ando (1979a, b) showed that if $\Phi(\cdot)$ is a unital positive linear map from $\mathbb{M}_n$ to $\mathbb{M}_m$ then

$$f(\Phi(A)) \geqslant \Phi(f(A)) \qquad \text{for operator-monotone } f \text{ and } A \geqslant 0. \tag{3.7}$$

Since $f(t)$ is concave, this can be considered as a matrix version of (1.8). In particular, with $f(t) = \log t$, one has

$$\log \Phi(A) \geqslant \Phi(\log A) \qquad (A > 0). \tag{3.8}$$

Based on the integral representation (3.4), Kubo and Ando (1979/80) developed a theory of matrix (or operator) means; a map $(A, B) \mapsto A \,\sigma\, B$ in the cone of positive semidefinite matrices is called a *matrix mean* or *operator mean* if the following conditions are satisfied:

(i) *Positive homogeneity*: $\alpha \cdot A \,\sigma\, B = (\alpha A) \,\sigma\, (\alpha B)$ for $\alpha \geqslant 0$;

(ii) *Normalization*: $A \,\sigma\, A = A$;

(iii) *Monotonicity*: $A \,\sigma\, B \geqslant A' \,\sigma\, B'$ whenever $A \geqslant A'$ and $B \geqslant B'$;

(iv) *Continuity from above*: $A_k \downarrow A$, $B_k \downarrow B$ implies $A_k \,\sigma\, B_k \downarrow A \,\sigma\, B$;

(v) *Transformer inequality*: $(T^*AT) \,\sigma\, (T^*BT) \geqslant T^*(A \,\sigma\, B)T$ for every matrix T.

A key for the theory is that there is a one-to-one correspondence between a matrix mean σ and a nonnegative operator-monotone function $f(t) \equiv f_\sigma(t)$ on $[0, \infty)$ with $f(1) = 1$ through the formula

$$A \,\sigma\, B = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2} \qquad \text{for} \quad A > 0, \quad B \geqslant 0. \quad (3.9)$$

If the mean σ corresponds to $f(t)$ with integral representation (3.4), then

$$A \,\sigma\, B = \alpha A + \beta B + \int_0^\infty A(sA + B)^{-1} B \, d\nu(s) \qquad (3.10)$$

The *arithmetic mean* corresponds to the function $f(t) = (t + 1)/2$, while the mean corresponding to $f(t) = 2t/(t + 1)$ will be called the *harmonic mean*. Half of the harmonic mean of A, $B \geqslant 0$ was introduced earlier, under the name of *parallel sum*, and denoted by $A : B$ by Anderson and Duffin (1969). With this notation (3.10) has the form

$$A \,\sigma\, B = \alpha A + \beta B + \int_0^\infty (sA) : B \frac{d\nu(s)}{s}. \qquad (3.11)$$

Then the map $(A, B) \mapsto A \,\sigma\, B$ turns out to be jointly concave in (A, B). Ando (1979b) and Kubo and Ando (1979/80) derived from the integral representation (3.11) that if $\Phi(\cdot)$ is a positive linear map from $\mathbb{M}_n$ to $\mathbb{M}_m$, then for any matrix mean σ

$$\Phi(A) \,\sigma\, \Phi(B) \geqslant \Phi(A \,\sigma\, B) \qquad (A, B \geqslant 0). \qquad (3.12)$$

In particular, with $C \geqslant 0$ and $\Phi(X) \equiv \operatorname{Tr}(XC)$, one has

$$\operatorname{Tr}(CA) \,\sigma\, \operatorname{Tr}(CB) \geqslant \operatorname{Tr}[C \cdot (A \,\sigma\, B)].$$

from which it follows that for every unitarily invariant norm $\|\cdot\|$

$$\|A\| \,\sigma\, \|B\| \geqslant \|A \,\sigma\, B\|. \qquad (3.13)$$

The definitions of arithmetic and harmonic means for matrices suggest that the *geometric mean* should be understood as the mean corresponding to the operator-monotone function $t^{1/2}$. This geometric mean is denoted by $A \# B$:

$$A \# B \equiv A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \qquad (A > 0, \quad B \geqslant 0). \quad (3.14)$$

The geometric mean $A \# B$ is characterized as the maximum of all positive semidefinite matrices X for which

$$\begin{pmatrix} A & X \\ X & B \end{pmatrix} \geqslant 0.$$

Further, for A, $B \geqslant 0$ there is a unitary matrix U such that $A \# B = A^{1/2}UB^{1/2}$. A similar idea appeared in Pusz and Woronowicz (1975) in order to define a functional calculus for sesquilinear forms on a C^*-algebra.

The parallel sum $A : B$ is characterized as the maximum of all positive semidefinite matrices X for which

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \geqslant \begin{pmatrix} X & X \\ X & X \end{pmatrix}.$$

In this connection Anderson, Morley, and Trapp (1990) investigated the following problem: what condition for an $m \times m$ Hermitian matrix K guarantees the existence of maximum in the set of positive semidefinite matrices X such that $\mathbf{A} + K \otimes X \geqslant 0$ for each positive semidefinite block matrix $\mathbf{A}$? They found that a necessary and sufficient condition is that K has only one negative eigenvalue.

Inequalities between matrix means are reduced to those between the corresponding functions. For instance, the arithmetic-geometric-harmonic mean inequalities hold:

$$\frac{A + B}{2} \geqslant A \# B \geqslant 2(A : B). \qquad (3.15)$$

The geometric mean $A \# B$ can be obtained as the limit of successive iteration of an operation defined by arithmetic means and harmonic means:

$$\lim_{k \to \infty} A_k = \lim_{k \to \infty} B_k,$$

where $A_0 \equiv A$, $B_0 \equiv B$, $A_{k+1} \equiv (A_k + B_k)/2$, and $B_{k+1} \equiv 2(A_k : B_k)$.

For $0 < \alpha < 1$, besides the weighted arithmetic mean $(1-\alpha)A + \alpha B$, one can define the matrix mean $A \#_\alpha B$ corresponding to the operator monotone function $f_\alpha(t) \equiv t^\alpha$:

$$A \#_\alpha B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^\alpha A^{1/2}. \tag{3.16}$$

When $\alpha = \frac{1}{2}$ this is just the geometric mean. It is seen that

$$A \#_\alpha B = B \#_{1-\alpha} A \qquad (0 < \alpha < 1). \tag{3.17}$$

When A commutes with B, then $A \#_\alpha B = A^{1-\alpha}B^\alpha$.

The traditional averaging operation

$$(A, B) \mapsto M_\alpha(A, B) \equiv \left\{\tfrac{1}{2}(A^\alpha + B^\alpha)\right\}^{1/\alpha} \qquad \text{with} \quad \alpha > 0$$

is not a matrix mean in the above sense except for $\alpha = 1$. In fact, this operation is not monotone in A or B except for $\alpha = 1$. However, from the operator monotonicity of the function t^α for $0 < \alpha \leqslant 1$ it follows that

$$M_{\alpha_1}(A, B) \geqslant M_{\alpha_2}(A, B) \qquad \text{whenever} \quad \alpha_1 \geqslant \alpha_2 \geqslant 1,$$

as observed in Bhagwat and Subramanian (1978). The limit of $M_\alpha(A, B)$ as $\alpha \to 0$ exists and equals $\exp\{(\log A + \log B)/2\}$. Contrary to the scalar case or the case of commuting A, B, this limit does not coincide with the geometric mean $A \# B$ in general. Moreover, the map $(A, B) \mapsto \exp\{(\log A + \log B)/2\}$ is not monotone in A, B.

The theory of matrix means is essentially for pairs of matrices and cannot be extended to larger sets of matrices. How to define matrix means—in particular, the geometric mean—for a larger number of positive definite matrices is still a challenging problem. One way is to use the integral representation in the definition. In this direction, Kosaki (1983) proposed a definition of the geometric mean $\mathbf{G}(A_0, A_1, \ldots, A_N)$ for an $(N+1)$-tuple $\{A_0, A_1, \ldots, A_N\}$ of positive definite matrices by the integral representation

$$\begin{aligned} \mathbf{G}(A_0, A_1, \ldots, A_N) \equiv \Gamma\left((N+1)^{-1}\right)^{-(N+1)} & \\ \times \int_{\mathbb{R}_+^N} \{A_0 : t_1 A_1 : \cdots : t_N A_N\} & \\ \times \prod_{j=1}^{N} t_j^{-(N+2)/(N+1)}\, dt_1\, dt_2 \cdots dt_N, & \end{aligned} \tag{3.18}$$

where $\Gamma(\cdot)$ is the gamma function. For $N = 1$ (3.18) reduces to the definition (3.14). When all A_j $(j = 0, 1, \ldots, N)$ commute with each other, $\mathbf{G}(A_0, A_1, \ldots, A_N) = (A_0 A_1 \cdots A_N)^{1/(N+1)}$.

An axiomatic approach to a definition of the geometric mean seems more difficult. With such a definition in mind, Anderson, Morley, and Trapp (1984) introduced the notion of *symmetric function means* for an N-tuple $\{A_1, \ldots, A_N\}$ of positive definite matrices, based on the ingenious representation of the symmetric function means of Marcus and Lopes for scalars (1957). They defined two series of matrices

$$\mathbf{P}_{k,N} \equiv \mathbf{P}_{k,N}(A_1, \ldots, A_N) \text{ and } \mathbf{T}_{k,N} \equiv \mathbf{T}_{k,N}(A_1, \ldots, A_N) \qquad (k = 1, 2, \ldots, N),$$

starting from $\mathbf{P}_{1,N}$ = the arithmetic mean of $A_1, \ldots, A_N$ and $\mathbf{T}_{N,N}$ = the harmonic mean of $A_1, \ldots, A_N$, by the successive formulas

$$\mathbf{P}_{k,N} \equiv \sum_{j=1}^{N} \left\{ \frac{A_j}{N-k+1} : \frac{\mathbf{P}_{k-1,N-1}(A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_N)}{k-1} \right\} \qquad (k = 2, \ldots, N)$$

and

$$\mathbf{T}_{k,N} = \prod_{j=1}^{N} : \left\{ kA_j + (N-k)\mathbf{T}_{k,N-1}(A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_N) \right\} \qquad (k = 1, \ldots, N-1),$$

where $\prod_{j=1}^{N}:$ denotes the parallel sum of N objects. $\mathbf{P}_{k,N}$ coincides with $\mathbf{T}_{k,N}$ if all A_i $(i = 1, 2, \ldots, N)$ commute with each other.

Trapp (1984) observed that starting with $X_i^{(1)} = A_i$ $(i = 1, \ldots, N)$ and defining successively $X_i^{(k+1)} = \mathbf{P}_{i,N}(X_1^{(k)}, \ldots, X_N^{(k)})$ $(i = 1, \ldots, N;\ k = 1, 2, \ldots,)$ the limits $\lim_{k \to \infty} X_i^{(k)}$ exist and coincide with each other for all $i = 1, \ldots, N$. This limit is a candidate for the geometric mean of $A_1, \ldots, A_n$. Replacing $\mathbf{P}$ by $\mathbf{T}$ in the above definition, one can get another candidate for the geometric mean of $A_1, \ldots, A_N$. The relation among those candidates and Kosaki's geometric mean (3.18) is not clear.

To establish an order relation between those two candidates of the geometric mean, Trapp (1984) conjectured the following inequalities:

$$\mathbf{P}_{k,N} \geqslant \mathbf{T}_{k,N} \qquad (k = 1, \dots, N) \tag{3.19}$$

and

$$\mathbf{P}_{k,N} \geqslant \mathbf{P}_{k+1,N} \text{ and } \mathbf{T}_{k,N} \geqslant \mathbf{T}_{k+1,N} \qquad (k = 1, \dots, N-1). \tag{3.20}$$

He proved $\mathbf{P}_{2,N} \geqslant \mathbf{T}_{N-1,N}$.

Ando (1983) discovered a new, powerful method for dealing with the conjectures. Using this method, Ando and Kubo (1989, 1990) could affirm the inequalities (3.19) and (3.20) for $2 \leqslant N \leqslant 4$. But the conjectures remain open for most k and N.

The notions and results in this section can be extended to bounded linear operators on Hilbert space.

4. MATRIX INEQUALITIES

As already mentioned, for $0 < \alpha \leqslant 1$ the function t^{α} is operator-monotone on $[0, \infty)$:

$$A \geqslant B \geqslant 0 \quad \text{implies} \quad A^{\alpha} \geqslant B^{\alpha} \quad (0 < \alpha \leqslant 1). \tag{4.1}$$

Chan and Kwong (1985) surmised that even for $p > 1$ the functions t^p has an operator-monotone-like property and conjectured, in particular, that $A \geqslant B \geqslant 0$ implies $A^2 \geqslant (AB^2A)^{1/2}$.

In this connection Furuta (1987) established the following matrix inequalities:

$$A \geqslant B \geqslant 0 \quad \text{implies} \quad A^{(p+2r)/q} \geqslant (A^rB^pA^r)^{1/q}$$

$$\text{whenever } r, p \geqslant 0,\ q \geqslant 1 \text{ and } (1+2r)q \geqslant p + 2r. \tag{4.2}$$

With $r = 1$ and $p = q = 2$ this reduces to the conjectured inequality. Subsequently Furuta together with his collaborators has been refining the inequality (4.2) and applying it to various situations to produce new inequalities.

As an application of (4.2), Ando (1987b) showed that even though the function e^t is not operator-monotone on the real line, for Hermitian matrices $H \geqslant K$ the map $t \mapsto e^{-tK} \# e^{tH}$ is increasing. This was generalized by Fujii, Furuta, and Kamei (1993) to the effect that the map $(p, r) \mapsto e^{-rK} \#_{\alpha(p,t,r)} e^{pH}$ is increasing for $p \geqslant t$ and $r \geqslant 0$, where $\alpha(p, r, t) \equiv (t + r)/(p + r)$.

To see another operator-monotone-like property of t^p for $p > 1$, define for $\lambda > 0$

$$f_{p,\lambda}(t) \equiv \frac{\lambda p t^p}{(p-1)t^p + \lambda^p} \qquad (t \geqslant 0). \tag{4.3}$$

Then obviously $t \geqslant f_{p,\lambda}(t) \geqslant 0$, and further $\sup_{\lambda > 0} f_{p,\lambda}(t) = t$ $(t \geqslant 0)$. Ando and Hiai (1994) show that for any $p > 1$ and $\lambda > 0$ and for $A, B \geqslant 0$

$$f_{p,\lambda}(A) \geqslant B \quad \text{implies} \quad A^p \geqslant B^p. \tag{4.4}$$

If $f(t)$ is a nonnegative operator-monotone function on $[0, \infty)$, it admits an analytic continuation to the upper half plane and maps the half plane into itself. Then for $p > 1$ the function $f(z^{1/p})^p$ is well defined, is analytic on the upper half plane, and transforms it into itself. Then $f(t^{1/p})^p$ is again operator-monotone. This shows that for $p > 1$ and $A, B \geqslant 0$

$$A^p \geqslant B^p \quad \text{implies} \quad f(A)^p \geqslant f(B)^p. \tag{4.5}$$

Lieb (1973) established that if $\alpha, \beta \geqslant 0$ and $\alpha + \beta \leqslant 1$, then for any matrix K the functional $A \mapsto \operatorname{Tr}(A^{\beta/2}K^*A^{\alpha}KA^{\beta/2})$ is concave on the cone of positive semidefinite matrices. This can be extended to the assertion that the function $(A, B) \mapsto \operatorname{Tr}(A^{\beta/2}K^*B^{\alpha}KA^{\beta/2})$ is jointly concave in (A, B).

Using the integral representation of an operator-monotone function (3.4) Ando (1979b) formulated Lieb's results as concavity of some maps in the cone of positive definite matrices and generalized it in the following way: if $f(t)$ is a nonnegative operator-monotone function on $[0, \infty)$, and $\Phi_i(\cdot)$ $(i = 1, 2)$ are concave maps in the cone of positive definite matrices, then the map

$$(A, B) \mapsto f\big(\Phi_1(A)^{-1} \otimes \Phi_2(B)\big) \cdot [\Phi_1(A) \otimes I] \tag{4.6}$$

is jointly concave in (A, B). And if $\Phi_1(\cdot)$ is affine, without the assumption of nonnegativity of $f(t)$, the map

$$(A, B) \mapsto f\big(\Phi_1(A) \otimes \Phi_2(B)^{-1}\big) \cdot [\Phi_1(A) \otimes I] \tag{4.7}$$

is convex. For instance, the first assertion implies joint concavity of the map

$$(A, B) \mapsto A^{\alpha_1} \otimes B^{\alpha_2} \qquad (\alpha_1, \alpha_2 \geqslant 0, \quad \alpha_1 + \alpha_2 \leqslant 1), \tag{4.8}$$

while the second implies joint convexity of the map

$$(A, B) \mapsto (A \cdot \log A) \otimes I - A \otimes \log B. \tag{4.9}$$

Theorems of this type can form a basis for convexity or concavity of certain numerical functionals on the cone of positive definite matrices, say of order n. For instance, the classical Minkowski theorem on concavity of the map $A \mapsto (\det A)^{1/n}$ on $\mathbb{M}_n$ is an immediate consequence of (4.8), because the map $A \mapsto A^{1/n} \otimes \cdots \otimes A^{1/n}$ is concave and $(\det A)^{1/n}$ is just the restriction of $A^{1/n} \otimes \cdots \otimes A^{1/n}$ to the subspace of antisymmetric tensors. Such connections are also pointed out in Bhatia and Davis (1985) and Merris (1982).

Thompson (1978, 1979) established a matrix-valued triangular inequality: for matrices A, B there are unitary matrices U, V, W such that

$$U^*|A|U + V^*|B|V \geqslant |A + B| \quad \text{and} \quad W^*|A|W \geqslant |\mathrm{Re}\, A|. \tag{4.10}$$

He investigated the case of equality.

Thompson (1976) also proved the following matrix inequality:

$$U^*(I + |A|)^{1/2}V^*(I + |B|)V(I + |A|)^{1/2}U \geqslant I + |A + B| \qquad \text{for some unitary } U, V, \tag{4.11}$$

from which one can derive the determinantal inequality of Seiler and Simon (1975),

$$\det(I + |A|)\det(I + |B|) \geqslant \det(I + |A + B|).$$

When A, B are *strictly contractive*, that is, $I - A^*A,\ I - B^*B > 0$, Hua (1955) proved the inequality

$$(I - B^*A)(I - A^*A)^{-1}(I - A^*B) \geqslant I - B^*B, \tag{4.12}$$

which is equivalent to

$$\begin{pmatrix} (I - A^*A)^{-1} & (I - B^*A)^{-1} \\ (I - A^*B)^{-1} & (I - B^*B)^{-1} \end{pmatrix} \geqslant 0.$$

(4.12) implies the determinantal inequality

$$|\det(I - A^*B)|^2 \geqslant \det(I - A^*A)\det(I - B^*B).$$

Ando (1979c) pointed out that (4.12) is a consequence of an obvious inequality

$$\mathrm{Re}(I - A^*B) \geqslant 2^{-1}\{(I - A^*A) + (I - B^*B)\}$$

and generalized (4.12) to the effect that for every operator-monotone function $f(t)$ on $(0, \infty)$

$$\mathrm{Re}\, f(I - A^*B) \geqslant 2^{-1}\{f(I - A^*A) + f(I - B^*B)\}, \qquad (4.13)$$

where $f(I - A^*B)$ is defined by the analytic continuation of $f(t)$.

It is obvious that for $A, B \geqslant 0$

$$(1 - \alpha)A + \alpha B \geqslant A \#_\alpha B \qquad (0 < \alpha \leqslant 1). \qquad (4.14)$$

Bhatia and Kittaneh (1990) established a *matrix arithmetic–geometric mean inequality* in the following form: for matrices A, B

$$U^*\left(\frac{|A|^2}{2} + \frac{|B|^2}{2}\right)U \geqslant |AB^*| \qquad \text{for some unitary } U. \qquad (4.15)$$

Ando (1994a) generalizes (4.15) to a *matrix Young inequality*: for $p, q > 1$ with $p^{-1} + q^{-1} = 1$

$$U^*\left(\frac{|A|^p}{p} + \frac{|B|^q}{q}\right)U \geqslant |AB^*| \qquad \text{for some unitary } U. \qquad (4.16)$$

Olson (1971) pointed out that the space of Hermitian matrices becomes a *conditionally complete lattice* under a more restrictive order $\hat{\geqslant}$; here $X \hat{\geqslant} Y$

is defined as $f(X) \geqslant f(Y)$ for all real valued increasing functions $f(t)$ on the real line. This new order relation is characterized by that the Hermitian projection to the eigenspace of X corresponding to the k largest eigenvalues is greater than that of Y for $k = 1, 2, \dots, n$. For A, $B \geqslant 0$ the relation $A \mathrel{\hat{\geqslant}} B$ is seen to be equivalent to $A^m \geqslant B^m$ $(m = 1, 2, \dots)$.

In this connection, Kato (1979) showed that the supremum of positive semidefinite $A_1, \dots, A_m$ with respect to this new order is obtained as $\lim_{k \to \infty}(A_1^k + \cdots + A_m^k)^{1/k}$.

Most of the results of this section can be extended to bounded linear operators on a Hilbert space. But the assertions on existence of unitaries requires separate considerations. For instance, Akemann, Anderson, and Pedersen (1982) got an inequality corresponding to (4.10) in a C^*-algebra setting with isometries in place of unitaries. And with unitaries only, an approximate version was proved with the left hand side of (4.10) increased by ϵI with arbitrarily small $\epsilon > 0$.

Generalizations of (4.15) or (4.16) to the C^*-algebra case are not known.

5. LOG MAJORIZATION

An important notion in majorization theory is log majorization for a pair of positive vectors. For $0 < a$, $b \in \mathbb{R}^n$, *log majorization*, in symbols

$$a \underset{(\log)}{\succ} b,$$

means that $\log a \succ \log b$. This is equivalent to requiring that

$$\prod_{i=1}^{k} a_{[i]} \geqslant \prod_{i=1}^{k} b_{[i]} \qquad (k = 1, 2, \dots, n) \tag{5.1}$$

and

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} b_i. \tag{5.2}$$

Weak log majorization, or *log submajorization*, for a pair of a, $b > 0$, in symbols

$$a \underset{(\log)}{\succ_w} b,$$

is defined by (5.1). This can be extended to a pair of nonnegative vectors.

Remark that, as in (1.3), for $a > 0$

$$\prod_{i=1}^{k} a_{[i]} = \max_{|J|=k} \prod_{i \in J} a_i \qquad (k = 1, 2, \dots, n). \tag{5.3}$$

It follows from (1.10) that if $f(t)$ is defined on $(0, \infty)$ and $f(e^t)$ is convex and increasing,

$$a \underset{(\log)}{\succ_w} b \quad \text{implies} \quad f(a) \succ_w f(b). \tag{5.4}$$

Since the function $f_\alpha(t) \equiv e^{\alpha t}$ for $\alpha > 0$ is convex and increasing on $(-\infty, \infty)$,

$$a \underset{(\log)}{\succ_w} b \quad \text{implies} \quad a^\alpha \succ_w b^\alpha.$$

In particular,

$$a \underset{(\log)}{\succ_w} b \quad \text{implies} \quad a \succ_w b. \tag{5.5}$$

Further it follows, with $f(t) = -\log t$, from (1.9) that

$$a > 0 \text{ and } a \succ b \quad \text{implies} \quad a^{-1} \underset{(\log)}{\succ_w} b^{-1}, \tag{5.6}$$

which means

$$\prod_{i=1}^{k} b_{[n-i+1]} \geqslant \prod_{i=1}^{k} a_{[n-i+1]} \quad (k = 1, 2, \dots, n) \qquad \text{if} \quad a, b > 0 \text{ and } a \succ b. \tag{5.7}$$

The eigenvalue analogue of (5.3) for a positive definite matrix A is

$$\prod_{i=1}^{k} \lambda_i(A) = \max\{\det[PAP + (I - P)] : P \text{ a Hermitian projection of rank } k\}. \tag{5.8}$$

A fundamental relation between eigenvalues and singular values for a general matrix A was discovered by Weyl:

$$[\sigma_i(A)] \underset{(\log)}{\succ} [|\lambda_i(A)|], \tag{5.9}$$

Here the eigenvalues $\lambda_i(A)$ are arranged as $|\lambda_1(A)| \geqslant |\lambda_2(A)| \geqslant \cdots \geqslant |\lambda_n(A)|$. This is a consequence of the rather trivial inequality

$$\sigma_1(A) \geqslant |\lambda_1(A)|,$$

which means that spectral radius is not greater than the spectral norm. To extend this to (5.9) is based on the use of *compounds* $\mathbf{C}_k(A)$ $(k = 1, 2, \ldots, n)$ with the help of the basic identities

$$\sigma_1(\mathbf{C}_k(A)) = \prod_{i=1}^{k} \sigma_i(A) \quad \text{and} \quad \lambda_1(\mathbf{C}_k(A)) = \prod_{i=1}^{k} \lambda_i(A). \tag{5.10}$$

The inverse problem for the majorization relation (5.9) has an affirmative solution. In fact, if

$$[\sigma_i] \underset{(\log)}{\succ} [|\lambda_i|] \qquad \text{for} \quad [\sigma_i] \in \mathbb{R}_+^n, \quad [\lambda_i] \in \mathbb{C}^n,$$

there is a matrix A for which

$$\sigma_i(A) = \sigma_i \text{ and } \lambda_i(A) = \lambda_i \qquad (i = 1, 2, \ldots, n).$$

See [MO, p. 233].

For A, $B > 0$ let us write

$$A \underset{(\log)}{\succ} B \qquad \left(A \underset{(\log)}{\succ_w} B\right)$$

to mean

$$[\lambda_i(A)] \underset{(\log)}{\succ} [\lambda_i(B)] \qquad \left([\lambda_i(A)] \underset{(\log)}{\succ_w} [\lambda_i(B)]\right).$$

C. J. Thompson (1971) established that

$$|A| \underset{(\log)}{\succ} |B|$$

if and only if for some unitary U two matrices $U|A|$ and $|B|$ have the same set of eigenvalues.

A matrix version of (5.6) is

$$A \underset{(\log)}{\succ_w} \Phi(A^{-1})^{-1} \qquad \text{for doubly stochastic } \Phi \text{ and } A > 0. \tag{5.11}$$

Considering the map $\Phi(X) \equiv \operatorname{diag}(X)$, (5.11) gives the Hadamard determinantal inequality:

$$\prod_{i=1}^{n} a_{ii} \geqslant \det A \qquad (A > 0).$$

The other fundamental fact about this subject is log majorization between singular values for the product of matrices, due to A. Horn (see [MO, p. 246] and [A, p. 229]): for any matrices A, B

$$[\sigma_i(A)\sigma_i(B)] \underset{(\log)}{\succ} [\sigma_i(AB)]. \tag{5.12}$$

This is a generalization of the trivial inequality

$$\sigma_1(A)\sigma_1(B) = \|A\|_\infty \|B\|_\infty \geqslant \|AB\|_\infty = \sigma_1(AB).$$

A Lidskii-Wielandt type relation for singular values was established by Gelfand and Naimark (see [MO, p. 248] and [A, p. 228]):

$$[\sigma_i(AB)] \underset{(\log)}{\succ} [\sigma_i(A)\sigma_{n-i+1}(B)]. \tag{5.13}$$

Use of the compound matrices makes it possible to formulate some assertions on matrix inequalities in terms of log majorization. For instance, as observed in Araki (1990), the Loewner inequality (4.1),

$$A \geqslant B \geqslant 0 \quad \text{implies} \quad A^\alpha \geqslant B^\alpha \quad \text{for } 0 < \alpha < 1,$$

can be restated in the form

$$(XYX)^{\alpha} \underset{(\log)}{\succ} X^{\alpha}Y^{\alpha}X^{\alpha} \quad \text{for } 0 < \alpha < 1 \qquad (X, Y > 0). \qquad (5.14)$$

See also Furuta (1989) and Wang and Gong (1993) for this relation.

We now show that many known trace inequalities can be extended to inequalities with respect to all unitarily invariant norms by establishing log majorization in a relatively easy way.

A typical example is the Golden-Thompson inequality (see [MO, p. 252]):

$$\operatorname{Tr}(e^{H}e^{K}) \geqslant \operatorname{Tr} e^{H+K} \qquad \text{for Hermitian } H, K.$$

This is generalized to the form

$$\|e^{H}e^{K}\| \geqslant \|e^{K/2}e^{H}e^{K/2}\| \geqslant \|e^{H+K}\|$$
$$\text{for every unitarily invariant norm } \|\cdot\|. \qquad (5.15)$$

This fact itself was already pointed out by C. J. Thompson (1971). The reason behind such generalization is better understood as follows. According to the Lie-Trotter formula (see Simon (1979, p. 97).)

$$\lim_{\alpha \downarrow 0} \left(e^{\alpha K/2}e^{\alpha H}e^{\alpha K/2}\right)^{1/\alpha} = e^{H+K},$$

(5.9) and (5.14) imply

$$|e^{H}e^{K}| \underset{(\log)}{\succ} e^{H+K}. \qquad (5.16)$$

In a similar way, Bernstein's trace inequality (1988)

$$\operatorname{Tr} e^{A+A^*} \geqslant \operatorname{Tr}(e^{A}e^{A^*})$$

was extended by Cohen (1988) to the form

$$\|e^{A+A^*}\| \geqslant \|e^{A}e^{A^*}\| \qquad \text{for every unitarily invariant norm } \|\cdot\|. \quad (5.17)$$

The line of such generalizations is fully developed in Hiai and Petz (1993) and Ando and Hiai (1993), in which also an estimate for e^{H+K} from below is established in the log majorization sense:

$$X^{\beta} \#_{\alpha} Y^{\beta} \underset{(\log)}{\succ} (X \#_{\alpha} Y)^{\beta} \quad \text{for } 0 < \beta < 1 \qquad (0 < \alpha < 1), \quad (5.18)$$

and

$$e^{(1-\alpha)H+\alpha K} = \lim_{\beta \downarrow 0} \left(e^{\beta H} \#_{\alpha} e^{\beta K}\right)^{1/\beta} \underset{(\log)}{\succ} e^{H} \#_{\alpha} e^{K}. \qquad (5.19)$$

As the Loewner inequality (4.1) is rewritten in terms of log majorization, the Furuta inequality (4.2) can also be expressed in the log majorization form

$$A^{(1-\alpha)/2} B^{\alpha} A^{(1-\alpha)/2} \underset{(\log)}{\succ} \left\{A^{t\mu} \#_{\alpha} \left(A^{(1-\alpha)\mu/2\alpha} B^{t} A^{(1-\alpha)\mu/2\alpha}\right)\right\}^{1/t}$$

whenever $1 > \alpha > 0$, $t > 0$, and $\min(\alpha, \ \alpha t) \geqslant \mu \geqslant 0$. In the converse direction, (5.18) is equivalent to the matrix inequality:

$$A \geqslant B \geqslant 0 \quad \text{implies}$$

$$A^{p} \geqslant \left\{A^{p/2}\left(A^{-1/2} B^{t} A^{-1/2}\right)^{p} A^{p/2}\right\}^{1/t} \quad (p, t \geqslant 1). \qquad (5.20)$$

See Ando and Hiai (1993) for details. We should mention that Furuta (1994) presents a parametric formula interpolating (4.2) and (5.20).

There are several known log majorization relations among products of fractional powers; see, for instance, Marshall and Olkin (1985). Those are unified in Ando and Hiai (1993) in the following way:

$$|A^{\alpha_1 + \cdots + \alpha_m} B^{\beta_1 + \cdots + \beta_m}| \underset{(\log)}{\succ} |A^{\alpha_1} B^{\beta_1} \cdots A^{\alpha_m} B^{\beta_m}| \qquad (A, B > 0) \quad (5.21)$$

whenever $\alpha_i, \ \beta_i \geqslant 0 \ (i = 1, 2, \ldots, m)$, and

$$\sum_{i=1}^{k} \beta_i \geqslant \sum_{i=1}^{k} \alpha_i \geqslant \sum_{i=1}^{k-1} \beta_i \quad (k = 1, 2, \ldots, m) \quad \text{and} \quad \sum_{i=1}^{m} \beta_i = \sum_{i=1}^{m} \alpha_i.$$

In particular, for any $\alpha_i \geqslant 0$ $(i = 1, 2, \dots, m)$

$$|A^{\alpha_1 + \cdots + \alpha_m} B^{\alpha_1 + \cdots + \alpha_m}| \underset{(\log)}{\succ} |A^{\alpha_1} B^{\alpha_1} \cdots A^{\alpha_m} B^{\alpha_m}|, \tag{5.22}$$

and

$$|A^{\alpha+\beta} B^{\gamma}| \underset{(\log)}{\succ} |A^{\alpha} B^{\gamma} A^{\beta}| \qquad \text{for any} \quad \alpha, \beta, \gamma \geqslant 0. \tag{5.23}$$

Further, for natural powers the following holds: for $A,\ B \geq 0$

$$A^{(k+1)/2} B^k A^{(k+1)/2} \underset{(\log)}{\succ} (AB)^k A \qquad (k = 1, 2, \dots). \tag{5.24}$$

6. SPECTRAL PERTURBATION

Given an $n \times n$ matrix X, denote by $\lambda(X)$ an arbitrarily arranged sequence of its eigenvalues, considered as a vector in $\mathbb{C}^n$. The *spectral distance* $\|(\lambda(A), \lambda(B))\|$ of a pair of *normal* matrices A, B with respect to a norm $\|\cdot\|$ on $\mathbb{M}_n$ is defined as

$$\|(\lambda(A), \lambda(B))\| \equiv \inf\{\|\operatorname{diag}(\lambda(A)) - P^T \operatorname{diag}(\lambda(B))\, P\| : P \text{ a permutation matrix}\}. \tag{6.1}$$

The *matching problem* is to obtain the optimal value $\kappa_{\|\cdot\|}$ of κ for which

$$\kappa \cdot \|A - B\| \geqslant \|(\lambda(A), \lambda(B))\| \qquad \text{for all normal } A, B. \tag{6.2}$$

Hoffman and Wielandt proved that $\kappa_{\|\cdot\|} = 1$ for the Frobenius norm $\|\cdot\|_2$ (see [MO, p. 274]). Using a Clifford algebra method, Bhatia and Bhattacharyya (1993) generalized this result to the case of commuting tuples of normal matrices: if $\{A^{(1)}, \dots, A^{(m)}\}$ is an m-tuple of commuting normal matrices with joint eigenvalues $[\alpha_j^{(k)}]$ $(j = 1, 2, \dots, n)$, and $(B^{(1)}, \dots, B^{(m)})$

is another m-tuple of commuting normal matrices with joint eigenvalues $[\beta_j^{(k)}]$ $(j = 1, 2, \dots, n)$, then there is a permutation π of $\{1, \dots, n\}$ such that

$$\sum_{k=1}^{m} \|A^{(k)} - B^{(k)}\|_2^2 \geqslant \sum_{k=1}^{m} \sum_{j=1}^{n} |\alpha_j^{(k)} - \beta_{\pi_j}^{(k)}|^2. \tag{6.3}$$

See Elsner (1993) for a simple proof.

For Hermitian A, B one has by Lidskii-Wielandt majorization (2.6)

$$[\lambda_i(A - B)] \succ [\lambda_i(A) - \lambda_i(B)],$$

which implies that for every unitary-similarity-invariant norm $\|\cdot\|$

$$\|A - B\| \geqslant \|\operatorname{diag}(\lambda_i(A)) - \operatorname{diag}(\lambda_i(B))\| \qquad \text{for Hermitian } A, B. \tag{6.4}$$

Remark that (6.4) for the case of the spectral norm was already in Weyl (see [MO, p. 552]).

The majorization (2.5) gives

$$[\lambda_i(A) - \lambda_{n-i+1}(B)] \succ [\lambda_i(A - B)],$$

which implies that for any unitary-similarity-invariant norm $\|\cdot\|$

$$\|\operatorname{diag}(\lambda_i(A)) - \operatorname{diag}(\lambda_{n-i+1}(B))\| \geqslant \|A - B\| \qquad \text{for Hermitian } A, B. \tag{6.5}$$

See Sunder (1982a) and Bhatia (1986) in this connection.

The long-standing Weyl conjecture was to ask whether $\kappa_{\|\cdot\|} = 1$ for spectral $\|\cdot\|_\infty$. Under suitable restrictions on the distribution of eigenvalues of normal A, B the inequality

$$\|A - B\|_\infty \geqslant \|(\lambda(A), \lambda(B))\|_\infty$$

has been guaranteed. Here are several examples: (1) by Sunder (1982b) when A is Hermitian and B is skew-Hermitian, (2) by Bhatia and Davis (1984) when both A, B are unitary, (3) by Bhatia and Holbrook (1985) when A is unitary and B is a scalar multiple of unitary, and (4) by Bhatia (1982) when $A - B$ is normal.

Proofs for cases (3) and (4) are based on an effective estimate of the spectral distance by Bhatia (1982): when A, B are connected by a piecewise C^1-arc $X(t)$ $(0 \leqslant t \leqslant 1)$, then

$$\int_0^1 \|X'(t)\|\, dt \geqslant \|(\lambda(A), \lambda(B))\|.$$

By a computer experiment, however, Holbrook (1992) finally gave a negative answer to the Weyl conjecture even for $n = 3$.

Sunder (1982b) and Ando and Bhatia (1989) studied the best value κ_p of the constant κ such that

$$\kappa \cdot \|A - B\|_p \geqslant \|(\lambda(A), \lambda(B))\|_p$$

for Hermitian A and skew-Hermitian B,

to show

$$\kappa_p = 2^{1/p - 1/2} \quad (1 \leqslant p \leqslant 2) \quad \text{and} \quad \kappa_p = 1 \quad (2 < p < \infty). \quad (6.6)$$

With an ingenious use of Fourier analysis, Bhatia, Davis, and McIntosh (1983) and Bhatia, Davis, and Koosis (1989) discovered a method of finding universal bounds for $\kappa_{\|\cdot\|}$ independent of the norm $\|\cdot\|$. One such universal bound is

$$\inf_f \int_{\mathbb{R}^2} |f(x, y)|\, dx\, dy,$$

where $f(x, y)$ runs over the set of integrable functions on $\mathbb{R}^2$ whose Fourier transforms $\hat{f}$ satisfy

$$\hat{f}(\xi, \eta) = (\xi + i\eta)^{-1} \qquad \text{for} \quad |\xi + i\eta| > 1.$$

It is known that this bound is between $\pi/2$ and 3.

Bhatia (1987) is a nice survey on perturbation of matrix eigenvalues.

Given a map $\Phi(\cdot)$, defined on a subset of the space of matrices, the perturbation problem with respect to a norm $\|\cdot\|$ is to find a reasonable bound for $\|\Phi(A) - \Phi(B)\|$ in terms of a quantity related to $A - B$. When the map is defined as $\Phi(A) \equiv f(A)$ by a continuous function $f(t)$ on an interval of the real line, there are a number of deep investigations of the St.

Petersburg school, including Birman, Farfarkova, Naboko, Solomyak, and others. For instance, Birman, Koplyanko, and Solomyak (1975) showed that for A, $B \geqslant 0$ and for every unitarily invariant norm $\|\cdot\|$

$$\| |A - B|^{\alpha} \| \geqslant \|A^{\alpha} - B^{\alpha}\| \qquad (0 < \alpha < 1).$$

Using the integral representation (3.4), Ando (1988) showed that for every nonnegative operator-monotone function $f(t)$ on $[0, \infty)$

$$\|f(|A - B|)\| \geqslant \|f(A) - f(B)\| \qquad \text{for} \quad A, B \geqslant 0,$$

which is equivalent to

$$f(|A - B|) \succ_{w} |f(A) - f(B)|. \tag{6.7}$$

Mathias (1990) pointed out that the same holds for a matrix-monotone function of order n and a pair of $n \times n$ positive semidefinite matrices.

For general matrices A, B comparison between $\| |A| - |B| \|_{\infty}$ and $\|A - B\|_{\infty}$ has a long history of research. The best possible Lipschitz constant for the map $\Phi(A) \equiv |A|$ with respect to spectral norm is known to depend on the dimension and is asymptotically of order $\log n$. In fact, there is a universal constant κ such that

$$\kappa \log n \cdot \|A - B\|_{\infty} \geqslant \| |A| - |B| \|_{\infty} \qquad \text{for } n \times n \text{ matrices } A, B. \tag{6.8}$$

The following form, not containing dimension n, for the spectral norm is due to Kato (1973):

$$\frac{2}{\pi}\|A - B\|_{\infty}\left\{2 + \log\frac{\|A\|_{\infty} + \|B\|_{\infty}}{\|A - B\|_{\infty}}\right\} \geqslant \| |A| - |B| \|_{\infty}. \tag{6.9}$$

For the Frobenius norm Araki and Yamagami (1981) showed that

$$\sqrt{2}\|A - B\|_2 \geqslant \| |A| - |B| \|_2$$

and the constant $\sqrt{2}$ is not necessary when A, B are Hermitian. This is completed by Kittaneh (1986a, b) in the following form: for all A, B, X

$$\begin{aligned} &\|AX - XB\|_2^2 + \|A^*X - XB^*\|_2^2 \\ &\qquad \geqslant \| |A|X - X|B| \|_2^2 + \| |A^*|X - X|B^*| \|_2^2. \end{aligned} \tag{6.10}$$

The following inequality was first proved by Kosaki (1984) for the trace norm and later generalized by Bhatia (1988) to every unitarily invariant norm:

$$\sqrt{2\|A+B\| \cdot \|A-B\|} \geqslant \| \, |A| - |B| \, \|. \tag{6.11}$$

For Schatten p-norms sharper estimates were obtained by Kittaneh and Kosaki (1987) and Bhatia (1988):

$$\max(2^{1/p-1/2}, 1) \cdot \sqrt{\|A+B\|_p \cdot \|A-B\|_p} \geqslant \| \, |A| - |B| \, \|_p. \tag{6.12}$$

Davies (1988) proved Lipschitz continuity of the *modulus map* $A \mapsto |A|$ for the case of $\| \cdot \|_p$ $(1 < p < \infty)$: there is an absolute constant κ such that

$$4\left\{1 + \kappa \max\left(p, \frac{p}{p-1}\right)\right\}\|A - B\|_p \geqslant \| \, |A| - |B| \, \|_p. \tag{6.13}$$

He used a deep result of Matsaev that the $\| \cdot \|_p$ norm of the triangular truncation map (observed in the next section) is uniformly bounded without respect to the order of matrices. See Gohberg and Krein (1967) in this connection. The inequality (6.13) can also be generalized to the form: for some constant γ_p and every A, B, X

$$\gamma_p\{\|AX - XB\|_p + \|A^*X - XB^*\|_p\} \geqslant \| \, |A|X - X|B| \, \|_p. \tag{6.14}$$

Davies's approach was analyzed in detail by Kosaki (1992b) to give a complete characterization of a unitarily invariant norm for which the modulus map is Lipschitz-continuous: such is the norm obtained as an interpolation norm between p_1- and p_2-norms for $1 < p_1, p_2 < \infty$.

7. HADAMARD PRODUCTS

In $\mathbb{M}_n$, besides the usual matrix product, the entrywise product is quite important and interesting. The entrywise product of two matrices A, B is called their *Hadamard* (or *Schur*) product and denoted by $A \circ B$. With this multiplication $\mathbb{M}_n$ becomes a commutative algebra, for which the matrix with all entries equal to one is a unit. Horn (1990) is an excellent survey on recent development of the study of the Hadamard product.

The most basic is the Schur theorem:

$$A \circ B \geqslant 0 \qquad \text{whenever} \quad A,\, B \geqslant 0. \tag{7.1}$$

When combined with the well-known fact that

$$1 \geqslant \|X\|_\infty \quad \text{if and only if} \quad \begin{pmatrix} I & X \\ X^* & I \end{pmatrix} \geqslant 0,$$

this yields submultiplicativity of the spectral norm with respect to Hadamard multiplication:

$$\|A\|_\infty \cdot \|B\|_\infty \geqslant \|A \circ B\|_\infty, \tag{7.2}$$

Submultiplicativity for the Frobenius norm is nothing but the Cauchy-Schwarz inequality.

Marcus, Kidman, and Sandy (1984) investigated several other cases and conjectured submultiplicativity for all unitarily invariant norms. But it was shown by Horn and Johnson (1987) that for a unitarily invariant norm $\|\cdot\|$ submultiplicativity with respect to the Hamadard is simultaneous with that with respect to the matrix product, and is characterized by the condition that $\|X\| \geqslant \|X\|_\infty$ for all X.

Contrary to the case of usual matrix multiplication, $[\sigma_i(A)\sigma_i(B)]$ does not always weakly log majorize $[\sigma_i(A \circ B)]$. However, weak majorization holds:

$$[\sigma_i(A)\sigma_i(B)] \succ_w [\sigma_i(A \circ B)]. \tag{7.3}$$

This was proved by Bapat and Sunder (1985) and also by Horn and Johnson (1987) as well as Okubo (1987) and Zhang (1987). As a consequence one has

$$\sigma_1(A)|B| \succ_w |A \circ B|.$$

The weak majorization (7.3) is improved by Horn and Johnson (1987) and by Ando, Horn, and Johnson (1987) in the following way: for any factorization $X^* Y = A$

$$[c_i(X)c_i(Y)\sigma_i(B)] \succ_w [\sigma_i(A \circ B)], \tag{7.4}$$

where for a matrix Z, $c_1(Z) \geqslant \cdots \geqslant c_n(Z)$ denote the Euclidean length of its columns, arranged in decreasing order. Bapat (1991) derived (7.4) from the general theorem (2.24).

Since

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

one can obtain from (7.4)

$$\left[\sqrt{p_i(A)p_i(A^*)}\,\sigma_i(B)\right] \succ_w [\sigma_i(A \circ B)], \tag{7.5}$$

where $p_1(X) \geqslant \cdots \geqslant p_n(X)$ are the diagonal entries of $|X|$.

Given a matrix A and a unitarily invariant norm $\|\cdot\|$, denote by $\gamma_{\|\cdot\|}(A)$ the norm of the linear map $\Phi_A(X) \equiv A \circ X$ with respect to this norm. Then (7.4) implies that for any unitarily invariant norm $\|\cdot\|$

$$c_1(X)c_1(Y) \geqslant \gamma_{\|\cdot\|}(A) \qquad \text{whenever} \quad X^*Y = A. \tag{7.6}$$

In particular, (7.5) implies

$$\sqrt{p_1(A)p_1(A^*)} \geqslant \gamma_{\|\cdot\|}(A). \tag{7.7}$$

For the spectral norm, (7.7) was obtained by Walter (1986).

In an unpublished manuscript Haagerup showed that with a special choice of X, Y with $X^*Y = A$, the number $c_1(X)c_1(Y)$ in (7.6) gives the mapping norm $\gamma_{\|\cdot\|_\infty}(A)$. This can be formulated in the following form:

$1 \geqslant \gamma_{\|\cdot\|_\infty}(A) \Leftrightarrow$

$$\begin{pmatrix} X & A \\ A^* & Y \end{pmatrix} \geqslant 0 \quad \text{for some } X, Y \geqslant 0 \qquad \text{with } I \geqslant X \circ I, Y \circ I. \tag{7.8}$$

With the observation that if $A = \operatorname{diag}(a_i)$ and $B = \operatorname{diag}(b_i)$ then $AX + XB = [a_i + b_j] \circ X$, Corach, Porta, and Recht (1990) derived from (7.6) that when S is an invertible Hermitian matrix

$$\|SXS^{-1} + S^{-1}XS\|_\infty \geqslant 2\|X\|_\infty \qquad \text{for all } X. \tag{7.9}$$

Kittaneh (1994) points out that (7.9) is an immediate consequence of a generalization of (4.15), due to Bhatia and Davis (1993), in the following form: for A, $B \geqslant 0$

$$\tfrac{1}{2}\|B^2X + XA^2\| \geqslant \|BXA\| \quad \text{for every unitarily invariant norm } \|\cdot\|, \tag{7.10}$$

which is equivalent to

$$\tfrac{1}{2}|B^2X + XA^2| \succ_w |BXA|. \tag{7.11}$$

The exact characterization (7.8) can be also used to show that (7.11) *cannot* be generalized to the Young form:

$$\left|\frac{1}{p}B^pX + \frac{1}{q}XA^q\right| \succ_w |BXA| \qquad \left(\frac{1}{p} + \frac{1}{q} = 1\right).$$

In this connection only the norm submultiplicative inequality is valid in general (see Kittaneh, 1993): for every unitarily invariant $\|\cdot\|$

$$\|B^pX\|^{1/p}\|XA^q\|^{1/q} \geqslant \|BXA\| \qquad (1/p + 1/q = 1). \tag{7.12}$$

Ando and Okubo (1991) gave a characterization of the norm $\gamma_w(A)$ of $\Phi_A(\cdot)$ with respect to the numerical radius norm $w(\cdot)$:

$$1 \geqslant \gamma_w(A) \Leftrightarrow$$

$$\begin{pmatrix} X & A \\ A^* & X \end{pmatrix} \geqslant 0 \quad \text{for some } X \geqslant 0 \text{ with } I \geqslant X \circ I, \tag{7.13}$$

and derived Haagerup's criterion (7.8) as a corollary. Also, Cowen, Dritschel, and Penney (1993) give another proof of (7.8).

Let $G(V, E)$ be an *undirected graph* with vertex set V, indexed by 1, 2, ..., N, and edge set E. Given matrices A_{ij} for each $\{i, j\} \in E$ and every $i = j$, the *positive completion problem* is to find a condition for the existence of a positive semidefinite block matrix $\mathbf{A} = [\hat{A}_{ij}]_{i,j=1}^N$ such that $\hat{A}_{ij} = A_{ij}$ whenever $\{i, j\} \in E$ or $i = j$. Paulsen, Power, and Smith (1989) showed that a positive completion exists if and only if the Hadamard multiplication map caused by $[\tilde{A}_{ij}]_{i,j=1}^N$ is positive, where $\tilde{A}_{ij} = A_{ij}$ whenever $\{i, j\} \in E$ or $i = j$, and $\tilde{A}_{ij} = 0$ otherwise. They also derived Haagerup's characterization (7.8) as a consequence.

The exact value of $\gamma_{\|\cdot\|_\infty}(A)$ is difficult to calculate except for special cases. It is equal to 1 when A is unitary, and it is equal to $\max\{a_{ii}\}$ when $A \geqslant 0$.

When the (i, j) entry of A is 0 or 1 according as $i < j$ or $i \geqslant j$, then $\Phi_A(B) \equiv A \circ B$ gives the *triangular truncation* of B. Denote by γ_n the $\|\cdot\|_\infty$-norm of the triangular truncation map on the space of $n \times n$ matrices. Angelos, Cowen, and Narayan (1993) give an exact asymptotic formula for γ_n:

$$\lim_{n\to\infty} \frac{\gamma_n}{\log n} = \frac{1}{\pi}.$$

A well-known matrix inequality of Fiedler (1961) says

$$A \circ A^{-1} \geqslant I \qquad (A > 0). \tag{7.14}$$

[See Bapat and Kwong (1987) for an improvement.] In this connection, as a consequence of (3.8), Ando (1979b) derived the following inequality:

$$\log(A \circ B) \geqslant (\log A + \log B) \circ I \qquad (A, B > 0). \tag{7.15}$$

Since by (2.3)

$$\log A + \log B \succ (\log A + \log B) \circ I,$$

combining (7.15) with (2.5) and (5.12), Ando (1994b) proves

$$\prod_{i=1}^{k} \lambda_{n-i+1}(A \circ B) \geqslant \prod_{i=1}^{k} \lambda_{n-i+1}(AB) \qquad (k = 1, 2, \dots, n), \tag{7.16}$$

which is equivalent to the statement that

$$A^{1/2}BA^{1/2} \underset{(\log)}{\succ_w} (A^{-1} \circ B^{-1})^{-1} \qquad (A, B > 0). \tag{7.17}$$

Since $B^T \circ I = B \circ I$ for Hermitian B, one can derive from (7.15) in a similar way

$$\prod_{i=1}^{k} \lambda_{n-i+1}(A \circ B) \geqslant \prod_{i=1}^{k} \lambda_{n-i+1}(AB^T) \qquad (k = 1, 2, \dots, n). \tag{7.18}$$

Remark that (7.16) gives an affirmative answer to the conjecture mentioned in Zhang (1993), and also yields, together with (7.18), the known inequality of Fiedler (1983):

$$A \circ B \geqslant \max\{\lambda_n(AB),\ \lambda_n(AB^T)\}I \qquad (A,\ B > 0). \tag{7.19}$$

Johnson and Elsner (1987) showed that for a positive function $f(t)$ on $(0, \infty)$ the matrix inequality

$$A \circ f(A) \geqslant A \cdot f(A) \qquad \text{for all} \quad A > 0$$

is valid if and only if $f(t)$ is a positive scalar multiple of t^{-1}.

8. MAJORIZATION IN VON NEUMANN ALGEBRAS

The algebra $\mathbb{M}_n$, equipped with spectral norm, and the algebra of bounded linear operators on a Hilbert space, equipped with operator norm, are special examples of von Neumann algebras. We refer for the basic notions and results on von Neumann algebras to Takesaki (1979).

The notion of (decreasingly arranged) generalized s-numbers (= singular values) for a not necessarily compact operator on a Hilbert space, or even for a *measurable operator* affiliated with a semifinite von Neumann algebra $\mathscr{M}$ acting on a Hilbert space, with a faithful normal semifinite trace $\tau(\cdot)$, was considered by Ovchinikov (1970), Sonis (1971), Fack (1982), and Fack and Kosaki (1986). This makes it possible to introduce various spaces of measurable operators as generalizations of the Schatten classes of compact operators. The above authors obtained various convexity inequalities, including several L^p-norm inequalities, by exploiting (weak) majorization relations for generalized s-numbers of the sum or product of two elements of $\mathscr{M}$.

Each self-adjoint element a of $\mathscr{M}$ is uniquely written in the form

$$a = \int_{-\infty}^{\infty} t\, de_a(t) \tag{8.1}$$

where $\{e_a(t): -\infty < t < \infty\}$ is the spectral projection of a. For general $a \in \mathscr{M}$ (also for any measurable operator a affiliated with $\mathscr{M}$), its generalized singular value function, or for short *generalized s-number* $\mu_t(a)$, is defined by

$$\mu_t(a) = \inf\{s \geqslant 0 : \tau(1 - e_{|a|}(s)) \leqslant t\} \qquad (t > 0). \tag{8.2}$$

It is shown that

$$\mu_t(a) = \inf\{\|ae\| : e \in \mathscr{M} \text{ a projection with } \tau(1-e) \leqslant t\}. \tag{8.3}$$

When $\mathscr{M}$ is the space of all bounded linear operators on a Hilbert space and $\tau(\cdot)$ is the usual trace, this definition coincides with the usual one of decreasingly arranged singular values for a compact operator, but it makes sense for every bounded linear operator, too.

When $\tau(1) < \infty$ and a is self-adjoint, Petz (1985) introduced the *spectral scale* $\lambda_a(t)$ of a by

$$\lambda_a(t) = \inf\{s \in \mathbb{R} : \tau(1 - e_a(s)) \leqslant t\} \qquad [0 < t < \tau(1)], \tag{8.4}$$

which corresponds to the decreasingly arranged eigenvalues of a Hermitian matrix.

Majorization and weak majorization between two self-adjoint elements of a *finite factor* were introduced by Kamei (1983) and extended by Hiai (1987) to the case of measurable operators affiliated with a semifinite von Neumann algebra.

When a and b are positive elements in $\mathscr{M}$, *weak majorization* $a \succ_w b$ is defined as

$$\int_0^s \mu_t(a)\,dt \geqslant \int_0^s \mu_t(b)\,dt \qquad \text{for} \quad s > 0, \tag{8.5}$$

and *majorization* with the additional requirement

$$\int_0^\infty \mu_t(a)\,dt = \int_0^\infty \mu_t(b)\,dt, \qquad \text{i.e.,} \quad \tau(a) = \tau(b). \tag{8.6}$$

Characterizations for (weak) majorization of elements of $\mathscr{M}$ are similar to the case of Hermitian matrices. A somewhat different (and more general) approach was treated by Alberti and Uhlmann (1982), where the relation of more mixedness (*unitary mixing*) plays a corresponding role.

The notion of a doubly stochastic map on a matrix space is naturally extended to a linear map $\Phi(\cdot)$ on a von Neumann algebra $\mathscr{M}$ with trace $\tau(\cdot)$. A map $\Phi(\cdot)$ is called *doubly stochastic* if it is positivity-preserving with $\Phi(1) = 1$ and τ-preserving. Kamei (1984, 1985) and Hiai (1987) discussed the Birkhoff-type theorem: when $\mathscr{M}$ is a finite factor, then the extreme point of the convex set of doubly stochastic maps consists exactly of all maps $\Phi(\cdot)$ such that $\Phi(a)$ is *equivalent* to a for all $a \in \mathscr{M}$.

Given $a \in \mathscr{M}$, define its *unitary orbit* by

$$\mathscr{U}(a) \equiv \{u^* a u : u \in \mathscr{M} \text{ unitary}\}, \tag{8.7}$$

and denote its convex hull by co $\mathscr{U}(a)$. Then as an extension of the Hardy-Littlewood-Pólya theorem the following holds: when $\mathscr{M}$ is a factor and a, b are positive elements with finite trace, then $a \succ b$ if and only if b is in the $\|\cdot\|_1$ closure of co $\mathscr{U}(a)$. The same assertion is true for any self-adjoint a, b in $L^1(\mathscr{M})$ when $\mathscr{M}$ is a finite factor.

In establishing majorization relations the following generalization of (2.16), found in Fack and Kosaki (1986), plays a key role:

$$\int_0^s \mu_t(a)\, dt = \inf\{\tau(|b|) + s\|a - b\| : b \in \mathscr{M}\} \qquad (a \in \mathscr{M}). \tag{8.8}$$

A Lidskii-Wielandt type theorem for generalized s-numbers was proved in the setting of von Neumann algebras by Hiai and Nakamura (1987), and is improved a little by Dodds, Dodds, and de Pagter (1989). It says that

$$\mu_t(a - b) \succ_w |\mu_t(a) - \mu_t(b)| \qquad (a, b \in \mathscr{M}), \tag{8.9}$$

or more explicitly,

$$\int_0^{|E|} \mu_t(a - b)\, dt \geqslant \int_E |\mu_t(a) - \mu_t(b)|\, dt$$

$$\text{for every Borel subset } E \subset (0, \infty), \tag{8.10}$$

where $|E|$ denotes the Lebesgue measure of E. In a similar line a Gelfand-Naimark type theorem for the generalized s-numbers was established by Nakamura (1987) in the following form: for any Borel subset of $(0, \tau(1))$

$$\int_0^{|E|} \log \mu_t(a)\, dt + \int_E \log \mu_t(b)\, dt \geqslant \int_E \log \mu_t(ab)\, dt, \tag{8.11}$$

provided $\mu_t(ab) > 0$ on $(0, \tau(1))$.

Kosaki (1992a) established a von Neumann algebra version of Araki's log majorization (5.14) in the following form: if $p \geqslant 1$ and $f(t)$ is a continuous increasing function on $\mathbb{R}_+$ such that $f(0) = 0$ and $t \mapsto f(e^t)$ is convex, then for any positive elements a, b satisfying $\lim_{s\to\infty} \mu_s(a) = \lim_{s\to\infty} \mu_s(b) = 0$

$$\int_0^t f(\mu_s(a^p b^p))\, ds \geqslant \int_0^t f(\mu_s(|ab|^p))\, ds \qquad (t > 0). \tag{8.12}$$

On the other hand, in the appendix to Hiai and Nakamura (1989), Kosaki presented the following extension of Ando's majorization (6.7): if $f(t)$ is a nonnegative operator monotone function on $[0, \infty)$, then for positive elements a, b

$$\int_0^t \mu_s(f(|a-b|))\, ds \geqslant \int_0^t \mu_s(f(a) - f(b))\, ds \qquad (t > 0). \tag{8.13}$$

See Dodds and Dodds (1994) for a generalization.

In Section 6 we discussed a matching problem for normal matrices with respect to the spectral norm. Let us take a quick look into its modification to the von Neumann algebra case. Given two *normal* elements $a, b \in \mathcal{M}$, define

$$\begin{aligned} \operatorname{dist}(a, b) &\equiv \|\cdot\|\text{-distance between } a \text{ and } \mathscr{U}(b) \\ &= \|\cdot\|\text{-distance between } \mathscr{U}(a) \text{ and } \mathscr{U}(b). \end{aligned} \tag{8.14}$$

Further define the *spectral distance* $\delta(a, b)$ as

$$\begin{aligned} \delta(a, b) \equiv \inf\Big\{ r > 0 : \tau\big(e_{V_r}(a)\big) \geqslant \tau(e_V(b)) \quad \text{and} \\ \tau\big(e_{V_r}(b)\big) \geqslant \tau(e_V(a)) \quad \text{for every open subset } V \subset \mathbb{C} \Big\}, \end{aligned} \tag{8.15}$$

where for a Borel set $E \subset \mathbb{C}$, $e_E(a)$ denotes the spectral projection of a corresponding to E, and E_r is the r-neighborhood of E. It is clear that when $\mathcal{M} = \mathbb{M}_n$, then $\delta(A, B)$ coincides with $\|(\lambda(A), \lambda(B))\|_\infty$ defined in (6.1).

As shown in Davidson (1986) and Hiai and Nakamura (1989), the von Neumann version of the Bhatia-Davis-McIntosh result holds: there is a universal constant κ such that

$$\kappa \operatorname{dist}(a, b) \geqslant \delta(a, b) \qquad \text{for normal } a, b \in \mathcal{M}. \tag{8.16}$$

It is also checked that $\operatorname{dist}(a, b) = \delta(a, b)$ holds under suitable restrictions on the distribution of the spectra of a, b, as in Section 6.

When $\mathscr{M}$ is a finite factor and a, b are self-adjoint, Hiai and Nakamura (1989) gave an exact formula to calculate the distance and the antidistance between a and $\mathscr{U}(b)$ with respect to the $\|\cdot\|_p$-metric ($1 \leqslant p \leqslant \infty$):

$$\inf\{\|a - u^*bu\|_p : u \in \mathscr{M} \text{ unitary}\} = \left\{\int_{-\infty}^{\infty} |\lambda_t(a) - \lambda_t(b)|^p \, dt\right\}^{1/p} \tag{8.17}$$

and

$$\sup\{\|a - u^*bu\|_p : u \in \mathscr{M} \text{ unitary}\} = \left\{\int_{-\infty}^{\infty} |\lambda_t(a) - \check{\lambda}_t(b)|^p \, dt\right\}^{1/p}, \tag{8.18}$$

where $\check{\lambda}_t(b) = -\lambda_t(-b)$. This is an extension of the Hoffman-Wielandt theorem on matching of eigenvalues for Hermitian matrices.

For self-adjoint a, b in the σ-finite infinite semifinite factor $\mathscr{M}$, Hiai and Nakamura (1989) gave an exact formula to calculate the distance between a and $\operatorname{co}\mathscr{U}(b)$:

$$\begin{aligned} &\inf\{\|a - c\| : c \in \operatorname{co}\mathscr{U}(b)\} \\ &\quad = \max\left\{0,\ \sup_{s>0} \frac{1}{s}\int_0^s [\lambda_t(a) - \lambda_t(b)] \, dt,\ \sup_{s>0} \frac{1}{s}\int_0^s \left[\check{\lambda}_t(a) - \check{\lambda}_t(b)\right] dt\right\} \end{aligned} \tag{8.19}$$

REFERENCES

Akemann, C. A., Anderson, J., and Pedersen, G. K. 1982. Triangular inequalities in operator algebras, *Linear and Multilinear Algebra* 11:167–178; MR 84a:46126.

Alberti, P. M., and Uhlmann, A. 1982. *Stochasticity and Partial Order*, D. Reidel, Boston; MR 84i:46057a-b.

Alfsen, E. M. 1971. *Compact Convex Sets and Boundary Integrals*, Springer-Verlag, New York; MR 56#3615.

Amir-Moéz, A. R. 1968. *Extreme Properties of Linear Transformations and Geometry in Unitary Spaces*, Texas Tech. Univ., Lubbock, Tex.; MR41#1751.

Anderson, W. N., Duffin, R. J. 1969. Series and parallel addition of matrices, *J. Math. Anal. Appl.* 26:576–594; MR 39#3904.

Anderson, W. N., Morley, T. D., and Trapp, G. E. 1984. Symmetric function means of positive operators, *Linear Algebra Appl.* 60:129–143; MR 85j:47015.

Anderson, W. N., Morley, T. D., and Trapp, G. E. 1990. Set of positive operators with suprema, *SIAM J. Matrix Anal. Appl.* 11:207–211; MR 91b:47037.

Ando, T. (1979a). *Topics on Operator Inequalities*, Hokkaido Univ., Sapporo, Japan; MR 58#2451.

Ando, T. 1979b. Concavity of certain maps on positive definite matrices and applications to Hadamard products, *Linear Algebra Appl.* 26:203–241; MR 80f:15023.

Ando, T. 1979c. Hua-Marcus inequalities, *Linear and Multilinear Algebra* 8:347–352; MR 81b:15013.

Ando, T. 1983. An inequality between symmetric function means of positive operators, *Acta Sci. Math.* (*Szeged*) 45:19–22; MR 85b:47014.

Ando, T. 1987a. Totally positive matrices, *Linear Algebra Appl.* 90:165–219; MR 88b:15023.

Ando, T. 1987b. On some operator inequalities, *Math. Ann.* 279:157–159; MR 89c:47019.

Ando, T. 1988. Comparison of norms $|||f(A)-f(B)|||$ and $|||f(|A-B|)|||$, *Math. Z.* 197:403–409; MR 90a:47021.

Ando, T. 1989. Majorization, doubly stochastic matrices, and comparison of eigenvalues, *Linear Algebra Appl.* 118:163–248; MR 90g:15034.

Ando, T. 1993. Parametrization of minimal points of some convex sets of matrices, *Acta Sci. Math.* (*Szeged*) 57:3–10.

Ando, T. 1994a. Matrix Young inequalities, *Oper. Theory Adv. Appl.*, to appear.

Ando, T. 1994b. A majorization relation for Hadamard product of positive definite matrices, to appear.

Ando, T., and Bhatia, R. 1989. Eigenvalue inequalities associated with the Cartesian decomposition, *Linear and Multilinear Algebra* 22:133–147; MR 89c:15022.

Ando, T., and Hiai, F. 1993. Log-majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.*, to appear.

Ando, T., and Hiai, F. 1994. Inequalities between powers of positive semidefinite matrices, *Linear Algebra Appl.*, to appear.

Ando, T., Horn, R., and Johnson, C. 1987. The singular values of a Hadamard product: Basic inequalities, *Linear and Multilinear Algebra* 21:345–364; MR 89g:15018.

Ando, T., and Kubo, F. 1989. Some matrix inequalities in multiport network connections, *Oper. Theory Adv. Appl.* 40:111–131; MR 91d:15038.

Ando, T., and Kubo, F. 1990. Inequalities among operator symmetric function means, *Progr. Systems Control Theory* 5:535–542; MR 92m:47035.

Ando, T., and Nakamura, Y. 1991. Some extremal problems related to majorization, *Oper. Theory Adv. Appl.* 50:83–92; MR 92f:46026.

Ando, T., and Okubo, K. 1991. Induced norms of the Schur multiplication operators, *Linear Algebra Appl.* 147:181–199; MR 92k:47015.

Angelos, J. R., Cowen, C. C., and Narayan, S. K. 1993. Triangular truncation and finding the norm of a Hadamard multiplier, *Linear Algebra Appl.*, to appear.

Araki, H. 1990. On an inequality of Lieb and Thirring, *Lett. Math. Phys.* 19:1167–1170; MR 91d:47020.

Araki, H., and Yamagami, S. 1981. An inequality for the Hilbert-Schmidt norm, *Comm. Math. Phys.* 81:89–91; MR 82j:46076.

Bapat, R. B. 1987. Majorization and singular values, *Linear and Multilinear Algebra* 21:211–214; MR 89j:15015.

Bapat, R. B. 1989. Majorization and singular values II, *SIAM J. Matrix Anal. Appl.* 10:429–434; MR 90i:15009.

Bapat, R. B. 1991. Majorization and singular values III, *Linear Algebra Appl.* 145:59–70; MR 91k:15018.

Bapat, R. B., and Kwong. 1987. A generalization of $A \circ A^{-1} \geqslant I$, *Linear Algebra Appl.* 93:107–111; MR 88g:15014.

Bapat, R. B., and Sunder, V. S. 1985. On majorization and Schur products, *Linear Algebra Appl.* 76:107–117; MR 87c:15018.

Bernstein, D. S. 1988. Inequalities for the trace of matrix exponentials, *SIAM J. Matrix Anal. Appl.* 9:156–158; MR 89b: 15031.

Bhagwat, K. V., and Subramanian, R. 1978. Inequalities between means of positive operators, *Math. Proc. Cambridge Philos. Soc.* 83:393–401; MR 67#7231.

Bhatia, R. 1982. Analysis of spectral variation and some inequalities, *Trans. Amer. Math. Soc.* 272:323–331; MR 83k:15015.

Bhatia, R. 1986. The distance between the eigenvalues of Hermitian matrices, *Proc. Amer. Math. Soc.* 96:41–42; MR 87b:15021.

Bhatia, R. 1987. *Perturbation Bounds for Matrix Eigenvalues*, Longman Scientific & Technical, Essex, U.K.; MR 88k:15020.

Bhatia, R. 1988. Perturbation inequalities for the absolute value map in norm ideals of operator, *J. Operator Theory* 19:129–136; MR 89g:47010.

Bhatia, R., and Bhattacharyya, T. 1993. A generalization of the Hoffman-Wielandt theorem, *Linear Algebra Appl.* 179:11–17.

Bhatia, R., and Davis, C. 1984. A bound for the spectral variation of a unitary operator, *Linear and Multilinear Algebra* 15:71–76; MR 85b:15020.

Bhatia, R., and Davis, C. 1985. Concavity of certain functions of matrices, *Linear and Multilinear Algebra* 17:155–164; MR 87j:15051.

Bhatia, R., and Davis, C. 1993. More matrix forms of the arithmetic geometric mean inequality, *SIAM J. Matrix Anal. Appl.* 14:132–136.

Bhatia, R., Davis, C., and Koosis, P. 1989. An extremal problem in Fourier analysis with application to operator theory, *J. Funct. Anal.* 82:138–150; MR 91a:42006.

Bhatia, R., Davis, C., and McIntosh, A. 1983. Perturbation of spectral subspaces and solutions of linear operator equations, *Linear Algebra Appl.* 52:45–67; MR 85a:47020.

Bhatia, R., and Holbrook, J. A. 1985. Short normal paths and spectral variation, *Proc. Amer. Math. Soc.* 94:277–282; MR 86e:15017.

Bhatia, R., and Holbrook, J. A. 1989. A softer, stronger Lidskii theorem, *Proc. Indian Acad. Sci.* 99:75–83; MR 90h:15034.

Bhatia, R., and Kittaneh, F. 1990. On the singular values of a product of operators, *SIAM J. Matrix Anal. Appl.* 11:272–277; MR 90m:47033.

Birman, M. Sh., Koplyanko, L. S., and Solomyak, M. Z. 1975. Estimate of the spectrum of the difference between fractional powers of self-adjoint operators, *Izv. Vyssh. Uchebn. Zaved. Mat.* 19:3–10; MR 52#6458.

Chan, N. N., and Kwong, K. 1985. Hermitian matrix inequalities and a conjecture, *Amer. Math. Monthly* 92:533–541; MR 87d:15011.

Cohen, J. E. 1988. Spectral inequalities for matrix exponentials, *Linear Algebra Appl.* 111:25–28; MR 90c:15025.

Corach, G., Porta, H., and Recht, L. 1990. An operator inequality, *Linear Algebra Appl.* 142:153–158; MR 91m:47020.

Cowen, C. C., Dritschel, M. A., and Penney, R. C. 1993. Norms of Hadamard multipliers, *SIAM J. Matrix Anal. Appl.*, to appear.

Davidson, K. R. 1986. The distance between unitary orbits of normal operators, *Acta Sci. Math. (Szeged)* 50:213–223; MR 88c:47040.

Davies, E. B. 1988. Lipschitz continuity of functions of operators in the Schatten classes, *J. London Math. Soc.* 37:148–157; MR 89c:47009.

Davis, C. 1963. Notions generalizing convexity for functions defined on spaces of matrices, in *Convexity* (V. Klee, Ed.), Proc. Symp. Pure Math. 7, Amer. Math. Soc., Providence, pp. 187–201; MR 27#5771.

Dodds, P. G., and Dodds, T. K. 1994. On a submajorization inequality of T. Ando, *Oper. Theory Adv. Appl.*, to appear.

Dodds, P. G., Dodds, T. K., and de Pagter, B. 1989. Non-commutative Banach function spaces, *Math. Z.* 201:583–597; MR 90j:46054.

Donoghue, W. 1974. *Monotone Matrix Functions and Analytic Continuation*, Springer-Verlag, New York; MR 58#6279.

Elsner, L. 1993. A note on the Hoffman-Wielandt theorem, *Linear Algebra Appl.* 182:235–237.

Fack, T. 1982. Sur la notion de valeur caractéristique, *J. Operator Theory* 7:307–333; MR 84m:47012.

Fack, T., and Kosaki, H. 1986. Generalized s-numbers of τ-measurable operators, *Pacific J. Math.* 123:269–300; MR 87h:46122.

Fiedler, M. 1961. Über eine Ungleichung fur positive definite Matrizen, *Math. Nachr.* 23:197–199; MR 25#3049.

Fiedler, M. 1983. A note on the Hadamard product of matrices, *Linear Algebra Appl.* 49:233–235; MR 84c:15025.

Fisher, P., and Holbrook, A. J. R. 1977. Matrices sous-stochastiques et fonctions convexes, *Canad. J. Math.* 29:631–637; MR 58#30042.

Fisher, P., and Holbrook, A. J. R. 1980. Balayage defined by the non-negative convex functions, *Proc. Amer. Math. Soc.* 78:445–448; MR 81f:46012.

Friedland, S., and Katz, M. 1987. On a matrix inequality, Linear Algebra Appl. 85:185–190; MR 87m:15053.

Fujii, M., Furuta, T., and Kamei, E. 1993. Furuta's inequality and its application to Ando's theorem, *Linear Algebra Appl.* 179:161–169.

Furuta, T. 1987. $A \geqslant B \geqslant 0$ assures $(B^rA^pB^r)^{1/q} \geqslant B^{(p+2r)/q}$ for $r \geqslant 0$, $p \geqslant 0$, $q \geqslant 1$ with $(1 + 2r)q \geqslant (p + 2r)$, *Proc. Amer. Math. Soc.* 101:85–88; MR 89b:47023.

Furuta, T. 1989. Norm inequalities equivalent to Loewner-Heinz theorem, *Rev. Math. Phys.* 1:135–137; MR 91b:47028.

Furuta, T. 1994. Extension of the Furuta inequality and log majorization by Ando and Hiai, *Linear Algebra Appl.*, to appear.

Garloff, J. 1982. Majorization between the diagonal elements and the eigenvalues of an oscillating matrix, *Linear Algebra Appl.* 47:181–184; MR 83m:15012.

Garloff, J. 1985. An inverse eigenvalue problem for totally nonnegative matrices, *Linear and Multilinear Algebra* 17:19–23; MR 87i:15009.

Gil, M. I. 1993. On inequalities for eigenvalues of matrices, *Linear Algebra Appl.* 184:201–206.

Gohberg, I., and Krein, M. G. 1965. *Introduction to the Theory of Linear Nonselfadjoint Operators* (transl.), Amer. Math. Soc., Providence; MR 36#3137.

Gohberg, I., and Krein, M. G. 1967. *Theory of Volterra Operators in Hilbert Space and Its Applications* (transl.), Amer. Math. Soc. Providence; MR #362007.

Hansen, F. 1980. An operator inequality, *Math. Ann.* 246:249–250; MR 82a:46065.

Hansen, F., and Pedersen, G. K. 1981/82. Jensen's inequality for operators and Loewner's theorem, *Math. Ann.* 258:229–241; MR 83g:47020.

Hässelbarth, W., and Ruch, E. 1993. On the extension of linear contractions, *Linear Algebra Appl.*, to appear.

Hersch, J., and Zwahlen, B. P. 1962. Évaluations par defaut pour une somme quelconque de valeurs propres γ_k d'un opérateur $C = A + B$ a l'aide de valeurs propres α_i de A et β_j de B, *C. R. Acad. Sci. Paris* 254:1559–1561; MR 24#A2583.

Hiai, F. 1987. Majorization and stochastic maps in von Neumann algebras, *J. Math. Anal. Appl.* 127:18–48; MR 88k:46076.

Hiai, F., and Nakamura, Y. 1987. Majorizations for generalized s-numbers in semifinite von Neumann algebras, *Math. Z.* 195:17–27; MR 88g:46070.

Hiai, F., and Nakamura, Y. 1989. Distance between unitary orbits in von Neumann algebra, *Pacific J. Math.* 138:259–294; MR 90i:46105.

Hiai, F., and Petz, D. 1993. The Golden-Thompson trace inequality is complemented, *Linear Algebra Appl.* 181:153–185.

Holbrook, A. J. 1992. Spectral variation of normal matrices, *Linear Algebra Appl.* 174:131–144.

Horn, A. 1962. Eigenvalues of sum of Hermitian matrices, *Pacific J. Math.* 9:225–241; MR 15#847.

Horn, R. 1990. The Hadamard product, in *Matrix Theory and Applications* (C. Johnson, Ed.), Proc. Symp. Appl. Math. 40, pp. 87–169; MR 92a:15001.

Horn, R., and Johnson, C. 1987. Hadamard and conventional submultiplicativity for unitarily invariant norms on matrices, *Linear and Multilinear Algebra* 20:19–106; MR 89b:15037.

Hua, L. 1955. Inequalities involving determinants, *Acta Math. Sinica* 5:463–470; MR 17#703.

Johnson, C., and Elsner, L. 1987. The relation between the Hadamard and conventional multiplication for positive definite matrices, *Linear Algebra Appl.* 92:231–240; MR 88i:15039.

Kamei, E. 1983. Majorization in finite factors, *Math. Japon.* 28:495–499; MR 84g:46086.

Kamei, E. 1984. Doubly stochasticity in finite factors, *Math. Japon.* 29:903–907; MR 88a:46067a.

Kamei, E. 1985. An order on statistical operators implicitly introduced by von Neumann, *Math. Japon.* 30:433–443; MR 87k:47048.

Kato, T. 1973. Continuity of the map $S \to |S|$ for linear operators, *Proc. Japan Acad.* 49:157–160; MR 53#8943.

Kato, T. 1979. Spectral order and a matrix limit theorem, *Linear and Multilinear Algebra* 8:15–19; MR 80j:47030.

Kittaneh, F. 1986a. Inequalities for the Schatten p-norm III, *Comm. Math. Phys.* 104:307–310; MR 88j:47009.

Kittaneh, F. 1986b. Inequalities for the Schatten p-norm IV, *Comm. Math. Phys.* 104:307–310; MR 88j:47010.

Kittaneh, F. 1992. A note on the arithmetic–geometric-mean inequality for matrices, *Linear Algebra Appl.* 171:1–8.

Kittaneh, F. 1993. Norm inequalities for fractional powers of positive operators, *Lett. Math. Phys.* 27:279–285.

Kittaneh, F. 1994. On some operator inequalities, to appear.

Kittaneh, F., and Kosaki, H. 1987. Inequalities for the Schatten p-norm V, *Publ. Res. Inst. Math. Sci.* 23:433–443; MR 88h:47011.

Komiya, H. 1983. Necessary and sufficient conditions for multivariate majorization, *Linear Algebra Appl.* 55:145–154; MR 84k:15015.

Kosaki, H. 1983. Private communication.

Kosaki, H. 1984. On the continuity of the map $\phi \to |\phi|$ from the predual of a W^*-algebra, *J. Funct. Anal.* 59:123–131; MR 86c:46072.

Kosaki, H. 1992a. An inequality of Araki-Lieb-Thirring, *Proc. Amer. Math. Soc.* 114:477–481; MR 92g:46079.

Kosaki, H. 1992b. Unitarily invariant norm under which the map $A \mapsto |A|$ is Lipschitz continuous, *Publ. Res. Inst. Math. Sci.* 28:299–313; MR 93d:47043.

Kubo, F., and Ando, T. 1979/80. Means of positive linear operators, *Math. Ann.* 249:205–224; MR 84d:47028.

Li, C. K. 1987. A generalization of spectral radius, numerical radius, and spectral norm, *Linear Algebra Appl.* 90:105–118; MR 88g:26015.

Li, C. K., and Tsing, N. K. 1989. Distance to the convex hull of unitary orbits, *Linear and Multilinear Algebra* 25:93–103; MR 90k:15011.

Lidskii, B. V. 1982. A polyhedron of the spectrum of the sum of two Hermitian matrices, *Funktsional. Anal. i Prilozhen.* 16:76–77; MR 83k:15009.

Lieb, E. 1973. Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Adv. In Math.* 11:267–288; MR 48#6218.

Marcus, M., Kidman, K., and Sandy, K. 1984. Unitarily invariant generalized matrix norms and Hadamard products, *Linear and Multilinear Algebra* 16:197–213; MR 87b:15039.

Marcus, M., and Lopes, L. 1957. Inequalities for symmetric functions and Hermitian matrices, *Canad. J. Math.* 9:305–312; MR 18#877.

Markus, A. S. 1964. Eigenvalues and singular values of the sum and product of linear operators, *Russian Math. Surveys* 19:91–120; MR 24#6318.

Marshall, A. W., and Olkin, I. 1979. *Inequalities: Theory of Majorization and Its Applications*, Academic, New York; MR 81b:00002.

Marshall, A. W., and Olkin, I. 1985. Inequalities for the trace function, *Aequationes Math.* 29:36–39; MR 87e:15037.

Mathias, R. 1990. Concavity of monotone matrix function of finite order, *Linear and Multilinear Algebra* 27:129–138; MR 91f:47011.

Merris, R. 1982. Extension of the Minkowski determinant theorem, *Portugal. Math.* 38:149–153; MR 84c:15009.

Nakamura, Y. 1987. An inequality for generalized s-numbers, *Integral Equations Operator Theory* 10:140–145; MR 88f:47016.

Nakamura, Y. 1989. Classes of operator monotone functions and Stieltjes functions, *Oper. Theory Adv. Appl.* 41:395–404; MR 91a:47017.

Nudelman, A. A., and Shvartzman, P. A. 1959. On the spectra of a product of unitary matrices, *Uspekhi Mat. Nauk* 13(84):111–117; MR 21#3436.

Okubo, K. 1987. Hölder-type norm inequalities for Schur products of matrices, *Linear Algebra Appl.* 91:13–28; MR 88f:15028.

Olson, N. P. 1971. The selfadjoint operators of a von Neumann algebra form a conditionally complete lattice, *Proc. Amer. Math. Soc.* 28:537–544; MR 43#2528.

Ovchinikov, V. I. 1970. s-Number of measurable operators, *Funktsional. Anal. i Prilozhen.* 4:236–242; MR 42#6644.

Paulsen, V. I., Power, S. C., and Smith, R. R. 1989. Schur products and matrix completions, *J. Funct. Anal.* 85:151–178; MR 90j:46051.

Petz, D. 1985. Spectral scale of selfadjoint operators and trace inequalities, *J. Math. Anal. Appl.* 109:74–82; MR c:47055.

Pusz, W., and Woronowicz, S. L. 1975. Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.* 8:159–170; MR 54#8316.

Riddle, R. C. 1984. Minimax problems on Grassmann manifolds, sums of eigenvalues, *Adv. Math.* 54:107–199; MR 86#58019.

Ruch, E., Schranner, R., and Seligman, R. 1980. Generalization of a theorem of Hardy-Littlewood-Polya, *J. Math. Anal. Appl.* 76:22–229; MR 81m:26015.

Seiler, E., and Simon, B. 1975. An inequality among determinants, *Proc. Nat. Acad. Sci. U.S.A.* 72:3277–3278; MR 55#6225.

Simon, B. 1979. *Trace Ideals and Their Applications*, Cambridge U.P., New York; MR 80k:47048.

Smiley, M. F. 1966. Inequalities related to Lidskii's, *Proc. Amer. Math. Soc.* 19:1029–1034; MR 38#2154.

Sonis, M. G. 1971. On a class of operators in von Neumann algebras with Segal measure on the projections, *Math. USSR Sb.* 13:344–357; MR 46#681.

Sunder, V. S. 1982a. On permutations, convex hulls and normal operators, *Linear Algebra Appl.* 48:403–411; MR 83c:47040.

Sunder, V. S. 1982b. Distance between normal operators, *Proc. Amer. Math. Soc.* 84:483–484; MR 83c:47040.

Takesaki, M. 1979. *Theory of Operator Algebras I*, Springer-Verlag, New York; MR 81e:46038.

Thompson, C. J. 1971. Inequalities and partial orders on matrix spaces, *Indiana Univ. Math. J.* 21:469–480; Mr 45#3442.

Thompson, R. C. 1976. A matrix inequality, *Comment. Math. Univ. Carolin.* 17:393–376; MR 54#7503.

Thompson, R. C. 1978. Matrix type metric inequalities, *Linear and Multilinear Algebra* 5:303–319; MR 58#16725.

Thompson, R. C. 1979. The case of equality in the matrix valued triangle inequality, *Pacific J. Math.* 82:279–280.

Thompson, R. C. 1986. Proof of a conjectured exponential formula, *Linear and Multilinear Algebra* 19:187–189; MR 88b:15027.

Trapp, G. E. 1984. Hermitian semidefinite matrix means and related inequalities—an introduction, *Linear and Multilinear Algebra* 16:113–123; MR 86b:15009.

Walter, M. E. 1986. On the norm of a Schur product, *Linear Algebra Appl.* 79:209–213; MR 87j:15048.

Wang, B. Y., and Gong, M. P. 1993. Some eigenvalue inequalities for positive semidefinite matrix power products, *Linear Algebra Appl.* 184:249–260.

Zhang, F. 1987. Another proof of a singular value inequality concerning Hadamard products of matrices, *Linear and Multilinear Algebra* 21:307–311; MR 89f:15022.

Zhang, F. 1993. A majorization conjecture for Hadamard products and compound matrices, *Linear and Multilinear Algebra* 33:301–303.

Received 7 October 1993