Lecture Notes in Mathematics

1790

Xingzhi Zhan

Matrix Inequalities

$$\|f(A+B)\| \leq \|f(A)+f(B)\|$$

Springer

Xingzhi Zhan

Matrix Inequalities

Springer

Author

Xingzhi ZHAN
Institute of Mathematics
Peking University
Beijing 100871, China

E-mail: zhan@math.pku.edu.cn

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Zhan, Xingzhi:
Matrix inequalities / Xingzhi Zhan. - Berlin ; Heidelberg ; New York ;
Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2002
(Lecture notes in mathematics ; Vol. 1790)
ISBN 3-540-43798-3

Mathematics Subject Classification (2000):
15-02, 15A18, 15A60, 15A45, 15A15, 47A63

ISSN 0075-8434
ISBN 3-540-43798-3 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York a member of BertelsmannSpringer
Science + Business Media GmbH

http://www.springer.de

Typesetting: Camera-ready TEX output by the author

SPIN: 10882616 41/3142/du-543210 - Printed on acid-free paper

Preface

Matrix analysis is a research field of basic interest and has applications in scientific computing, control and systems theory, operations research, mathematical physics, statistics, economics and engineering disciplines. Sometimes it is also needed in other areas of pure mathematics.

A lot of theorems in matrix analysis appear in the form of inequalities. Given any complex-valued function defined on matrices, there are inequalities for it. We may say that matrix inequalities reflect the quantitative aspect of matrix analysis. Thus this book covers such topics as norms, singular values, eigenvalues, the permanent function, and the Löwner partial order.

The main purpose of this monograph is to report on recent developments in the field of matrix inequalities, with emphasis on useful techniques and ingenious ideas. Most of the results and new proofs presented here were obtained in the past eight years. Some results proved earlier are also collected as they are both important and interesting.

Among other results this book contains the affirmative solutions of eight conjectures. Many theorems unify previous inequalities; several are the culmination of work by many people. Besides frequent use of operator-theoretic methods, the reader will also see the power of classical analysis and algebraic arguments, as well as combinatorial considerations.

There are two very nice books on the subject published in the last decade. One is *Topics in Matrix Analysis* by R. A. Horn and C. R. Johnson, Cambridge University Press, 1991; the other is *Matrix Analysis* by R. Bhatia, GTM 169, Springer, 1997. Except a few preliminary results, there is no overlap between this book and the two mentioned above.

At the end of every section I give notes and references to indicate the history of the results and further readings.

This book should be a useful reference for research workers. The prerequisites are linear algebra, real and complex analysis, and some familiarity with Bhatia's and Horn-Johnson's books. It is self-contained in the sense that detailed proofs of all the main theorems and important technical lemmas are given. Thus the book can be read by graduate students and advanced undergraduates. I hope this book will provide them with one more opportunity to appreciate the elegance of mathematics and enjoy the fun of understanding certain phenomena.

I am grateful to Professors T. Ando, R. Bhatia, F. Hiai, R. A. Horn, E. Jiang, M. Wei and D. Zheng for many illuminating conversations and much help of various kinds.

This book was written while I was working at Tohoku University, which was supported by the Japan Society for the Promotion of Science. I thank JSPS for the support. I received warm hospitality at Tohoku University.

Special thanks go to Professor Fumio Hiai, with whom I worked in Japan. I have benefited greatly from his kindness and enthusiasm for mathematics.

I wish to express my gratitude to my son Sailun whose unique character is the source of my happiness.

Sendai, December 2001 *Xingzhi Zhan*

Contents

1. Inequalities in the Löwner Partial Order

Throughout we consider square complex matrices. Since rectangular matrices can be augmented to square ones with zero blocks, all the results on singular values and unitarily invariant norms hold as well for rectangular matrices. Denote by M_n the space of $n \times n$ complex matrices. A matrix $A \in M_n$ is often regarded as a linear operator on $\mathbb{C}^n$ endowed with the usual inner product $\langle x, y \rangle \equiv \sum_j x_j \bar{y}_j$ for $x = (x_j), y = (y_j) \in \mathbb{C}^n$. Then the conjugate transpose A^* is the adjoint of A. The Euclidean norm on $\mathbb{C}^n$ is $\|x\| = \langle x, x \rangle^{1/2}$. A matrix $A \in M_n$ is called *positive semidefinite* if

$$\langle Ax, x \rangle \geq 0 \quad \text{for all} \quad x \in \mathbb{C}^n. \tag{1.1}$$

Thus for a positive semidefinite A, $\langle Ax, x \rangle = \langle x, Ax \rangle$. For any $A \in M_n$ and $x, y \in \mathbb{C}^n$, we have

$$4\langle Ax, y \rangle = \sum_{k=0}^{3} i^k \langle A(x + i^k y), x + i^k y \rangle,$$

$$4\langle x, Ay \rangle = \sum_{k=0}^{3} i^k \langle x + i^k y, A(x + i^k y) \rangle$$

where $i = \sqrt{-1}$. It is clear from these two identities that the condition (1.1) implies $A^* = A$. Therefore a positive semidefinite matrix is necessarily Hermitian.

In the sequel when we talk about matrices $A, B, C, \ldots$ without specifying their orders, we always mean that they are of the same order. For Hermitian matrices G, H we write $G \leq H$ or $H \geq G$ to mean that $H - G$ is positive semidefinite. In particular, $H \geq 0$ indicates that H is positive semidefinite. This is known as the *Löwner partial order*; it is induced in the real space of (complex) Hermitian matrices by the cone of positive semidefinite matrices. If H is positive definite, that is, positive semidefinite and invertible, we write $H > 0$.

Let $f(t)$ be a continuous real-valued function defined on a real interval Ω and H be a Hermitian matrix with eigenvalues in Ω. Let $H = U\text{diag}(\lambda_1, \ldots, \lambda_n)U^*$ be a spectral decomposition with U unitary. Then the *functional calculus* for H is defined as

$$f(H) \equiv U\mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))U^*. \tag{1.2}$$

This is well-defined, that is, $f(H)$ does not depend on particular spectral decompositions of H. To see this, first note that (1.2) coincides with the usual polynomial calculus: If $f(t) = \sum_{j=0}^{k} c_j t^j$ then $f(H) = \sum_{j=0}^{k} c_j H^j$. Second, by the Weierstrass approximation theorem, every continuous function on a finite closed interval Ω is uniformly approximated by a sequence of polynomials. Here we need the notion of a norm on matrices to give a precise meaning of approximation by a sequence of matrices. We denote by $\|A\|_\infty$ the spectral (operator) norm of A: $\|A\|_\infty \equiv \max\{\|Ax\| : \|x\| = 1, x \in \mathbb{C}^n\}$. The spectral norm is submultiplicative: $\|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty$. The positive semidefinite square root $H^{1/2}$ of $H \geq 0$ plays an important role.

Some results in this chapter are the basis of inequalities for eigenvalues, singular values and norms developed in subsequent chapters. We always use capital letters for matrices and small letters for numbers unless otherwise stated.

1.1 The Löwner-Heinz inequality

Denote by I the identity matrix. A matrix C is called *a contraction* if $C^*C \leq I$, or equivalently, $\|C\|_\infty \leq 1$. Let $\rho(A)$ be the spectral radius of A. Then $\rho(A) \leq \|A\|_\infty$. Since AB and BA have the same eigenvalues, $\rho(AB) = \rho(BA)$.

Theorem 1.1 (Löwner-Heinz) *If $A \geq B \geq 0$ and $0 \leq r \leq 1$ then*

$$A^r \geq B^r. \tag{1.3}$$

Proof. The standard continuity argument is that in many cases, e.g., the present situation, to prove some conclusion on positive semidefinite matrices it suffices to show it for positive definite matrices by considering $A+\epsilon I$, $\epsilon \downarrow 0$. Now we assume $A > 0$.

Let Δ be the set of those $r \in [0, 1]$ such that (1.3) holds. Obviously $0, 1 \in \Delta$ and Δ is closed. Next we show that Δ is convex, from which follows $\Delta = [0, 1]$ and the proof will be completed. Suppose $s, t \in \Delta$. Then

$$A^{-s/2}B^sA^{-s/2} \leq I, \quad A^{-t/2}B^tA^{-t/2} \leq I$$

or equivalently $\|B^{s/2}A^{-s/2}\|_\infty \leq 1$, $\|B^{t/2}A^{-t/2}\|_\infty \leq 1$. Therefore

$$\begin{aligned}
\|A^{-(s+t)/4}B^{(s+t)/2}A^{-(s+t)/4}\|_\infty &= \rho(A^{-(s+t)/4}B^{(s+t)/2}A^{-(s+t)/4}) \\
&= \rho(A^{-s/2}B^{(s+t)/2}A^{-t/2}) \\
&= \|A^{-s/2}B^{(s+t)/2}A^{-t/2}\|_\infty \\
&= \|(B^{s/2}A^{-s/2})^*(B^{t/2}A^{-t/2})\|_\infty \\
&\leq \|B^{s/2}A^{-s/2}\|_\infty \|B^{t/2}A^{-t/2}\|_\infty \\
&\leq 1.
\end{aligned}$$

Thus $A^{-(s+t)/4}B^{(s+t)/2}A^{-(s+t)/4} \leq I$ and consequently $B^{(s+t)/2} \leq A^{(s+t)/2}$, i.e., $(s+t)/2 \in \Delta$. This proves the convexity of Δ. □

How about this theorem for $r > 1$? The answer is negative in general. The example

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A^2 - B^2 = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}$$

shows that $A \geq B \geq 0 \not\Rightarrow A^2 \geq B^2$.

The next result gives a conceptual understanding, and this seems a typical way of mathematical thinking.

We will have another occasion in Section 4.6 to mention the notion of a C^*-algebra, but for our purpose it is just M_n. Let $\mathcal{A}$ be a Banach space over $\mathbb{C}$. If $\mathcal{A}$ is also an algebra in which the norm is submultiplicative: $\|AB\| \leq \|A\|\,\|B\|$, then $\mathcal{A}$ is called a *Banach algebra*. An *involution* on $\mathcal{A}$ is a map $A \mapsto A^*$ of $\mathcal{A}$ into itself such that for all $A, B \in \mathcal{A}$ and $\alpha \in \mathbb{C}$

$$\text{(i) } (A^*)^* = A; \quad \text{(ii) } (AB)^* = B^*A^*; \quad \text{(iii) } (\alpha A + B)^* = \bar{\alpha}A^* + B^*.$$

A C^**-algebra* $\mathcal{A}$ is a Banach algebra with involution such that

$$\|A^*A\| = \|A\|^2 \quad \text{for all } A \in \mathcal{A}.$$

An element $A \in \mathcal{A}$ is called *positive* if $A = B^*B$ for some $B \in \mathcal{A}$.

It is clear that M_n with the spectral norm and with conjugate transpose being the involution is a C^*-algebra. Note that the Löwner-Heinz inequality also holds for elements in a C^*-algebra and the same proof works, since every fact used there remains true, for instance, $\rho(AB) = \rho(BA)$.

Every element $T \in \mathcal{A}$ can be written uniquely as $T = A + iB$ with A, B Hermitian. In fact $A = (T + T^*)/2$, $B = (T - T^*)/2i$. This is called the *Cartesian decomposition* of T.

We say that $\mathcal{A}$ is *commutative* if $AB = BA$ for all $A, B \in \mathcal{A}$.

Theorem 1.2 *Let $\mathcal{A}$ be a C^*-algebra and $r > 1$. If $A \geq B \geq 0$, $A, B \in \mathcal{A}$ implies $A^r \geq B^r$, then $\mathcal{A}$ is commutative.*

Proof. Since $r > 1$, there exists a positive integer k such that $r^k > 2$. Suppose $A \geq B \geq 0$. Use the assumption successively k times we get $A^{r^k} \geq B^{r^k}$. Then apply the Löwner-Heinz inequality with the power $2/r^k < 1$ to obtain $A^2 \geq B^2$. Therefore it suffices to prove the theorem for the case $r = 2$.

For any $A, B \geq 0$ and $\epsilon > 0$ we have $A + \epsilon B \geq A$. Hence by assumption, $(A + \epsilon B)^2 \geq A^2$. This yields $AB + BA + \epsilon B^2 \geq 0$ for any $\epsilon > 0$. Thus

$$AB + BA \geq 0 \quad \text{for all } A, B \geq 0. \tag{1.4}$$

Let $AB = G + iH$ with G, H Hermitian. Then (1.4) means $G \geq 0$. Applying this to A, BAB,

$$A(BAB) = G^2 - H^2 + i(GH + HG) \tag{1.5}$$

gives $G^2 \geq H^2$. So the set

$$\Gamma \equiv \{\alpha \geq 1 : G^2 \geq \alpha H^2 \text{ for all } A, B \geq 0 \text{ with } AB = G + iH\}$$

where $G + iH$ is the Cartesian decomposition, is nonempty. Suppose Γ is bounded. Then since Γ is closed, it has a largest element λ. By (1.4) $H^2(G^2 - \lambda H^2) + (G^2 - \lambda H^2)H^2 \geq 0$, i.e.,

$$G^2H^2 + H^2G^2 \geq 2\lambda H^4. \tag{1.6}$$

From (1.5) we have $(G^2 - H^2)^2 \geq \lambda(GH + HG)^2$, i.e.,

$$\begin{aligned} &G^4 + H^4 - (G^2H^2 + H^2G^2) \\ &\geq \lambda[GH^2G + HG^2H + G(HGH) + (HGH)G]. \end{aligned}$$

Combining this inequality, (1.6) and the inequalities $GH^2G \geq 0$, $G(HGH) + (HGH)G \geq 0$ (by (1.4) and $G \geq 0$), $HG^2H \geq \lambda H^4$ (by the definition of λ) we obtain

$$G^4 \geq (\lambda^2 + 2\lambda - 1)H^4.$$

Then applying the Löwner-Heinz inequality again we get

$$G^2 \geq (\lambda^2 + 2\lambda - 1)^{1/2}H^2$$

for all G, H in the Cartesian decomposition $AB = G + iH$ with $A, B \geq 0$. Hence $(\lambda^2 + 2\lambda - 1)^{1/2} \in \Gamma$, which yields $(\lambda^2 + 2\lambda - 1)^{1/2} \leq \lambda$ by definition. Consequently $\lambda \leq 1/2$. This contradicts the assumption that $\lambda \geq 1$. So Γ is unbounded and $G^2 \geq \alpha H^2$ for all $\alpha \geq 1$, which is possible only when $H = 0$. Consequently $AB = BA$ for all positive A, B. Finally by the Cartesian decomposition and the fact that every Hermitian element is a difference of two positive elements we conclude that $XY = YX$ for all $X, Y \in \mathcal{A}$. □

Since M_n is noncommutative when $n \geq 2$, we know that for any $r > 1$ there exist $A \geq B \geq 0$ but $A^r \not\geq B^r$.

Notes and References. The proof of Theorem 1.1 here is given by G. K. Pedersen [79]. Theorem 1.2 is due to T. Ogasawara [77].

1.2 Maps on Matrix Spaces

A real-valued continuous function $f(t)$ defined on a real interval Ω is said to be *operator monotone* if

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

for all such Hermitian matrices A, B of all orders whose eigenvalues are contained in Ω. f is called *operator convex* if for any $0 < \lambda < 1$,

$$f(\lambda A + (1-\lambda)B) \leq \lambda f(A) + (1-\lambda) f(B)$$

holds for all Hermitian matrices A, B of all orders with eigenvalues in Ω. f is called *operator concave* if $-f$ is operator convex.

Thus the Löwner-Heinz inequality says that the function $f(t) = t^r$, $(0 < r \leq 1)$ is operator monotone on $[0, \infty)$. Another example of operator monotone function is $\log t$ on $(0, \infty)$ while an example of operator convex function is $g(t) = t^r$ on $(0, \infty)$ for $-1 \leq r \leq 0$ or $1 \leq r \leq 2$ [17, p.147].

If we know the formula

$$t^r = \frac{\sin r\pi}{\pi} \int_0^\infty \frac{s^{r-1} t}{s+t} ds \quad (0 < r < 1)$$

then Theorem 1.1 becomes quite obvious. In general we have the following useful integral representations for operator monotone and operator convex functions. This is part of Löwner's deep theory [17, p.144 and 147] (see also [32]).

Theorem 1.3 *If f is an operator monotone function on $[0, \infty)$, then there exists a positive measure μ on $[0, \infty)$ such that*

$$f(t) = \alpha + \beta t + \int_0^\infty \frac{st}{s+t} d\mu(s) \tag{1.7}$$

where α is a real number and $\beta \geq 0$. If g is an operator convex function on $[0, \infty)$ then there exists a positive measure μ on $[0, \infty)$ such that

$$g(t) = \alpha + \beta t + \gamma t^2 + \int_0^\infty \frac{st^2}{s+t} d\mu(s) \tag{1.8}$$

where α, β are real numbers and $\gamma \geq 0$.

The three concepts of operator monotone, operator convex and operator concave functions are intimately related. For example, a nonnegative continuous function on $[0, \infty)$ is operator monotone if and only if it is operator concave [17, Theorem V.2.5].

A map $\Phi : M_m \to M_n$ is called *positive* if it maps positive semidefinite matrices to positive semidefinite matrices: $A \geq 0 \Rightarrow \Phi(A) \geq 0$. Denote by I_n the identity matrix in M_n. Φ is called *unital* if $\Phi(I_m) = I_n$.

We will first derive some inequalities involving unital positive linear maps, operator monotone functions and operator convex functions, then use these results to obtain inequalities for matrix Hadamard products.

The following fact is very useful.

Lemma 1.4 *Let $A > 0$. Then*

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$$

if and only if the Schur complement $C - B^*A^{-1}B \geq 0$.

Lemma 1.5 *Let* Φ *be a unital positive linear map from* M_m *to* M_n. *Then*

$$\Phi(A^2) \geq \Phi(A)^2 \quad (A \geq 0), \tag{1.9}$$

$$\Phi(A^{-1}) \geq \Phi(A)^{-1} \quad (A > 0). \tag{1.10}$$

Proof. Let $A = \sum_{j=1}^{m} \lambda_j E_j$ be the spectral decomposition of A, where $\lambda_j \geq 0\,(j = 1, \ldots, m)$ are the eigenvalues and $E_j\,(j = 1, \ldots, m)$ are the corresponding eigenprojections of rank one with $\sum_{j=1}^{m} E_j = I_m$. Then since $A^2 = \sum_{j=1}^{m} \lambda_j^2 E_j$ and by unitality $I_n = \Phi(I_m) = \sum_{j=1}^{m} \Phi(E_j)$, we have

$$\begin{bmatrix} I_n & \Phi(A) \\ \Phi(A) & \Phi(A^2) \end{bmatrix} = \sum_{j=1}^{m} \begin{bmatrix} 1 & \lambda_j \\ \lambda_j & \lambda_j^2 \end{bmatrix} \otimes \Phi(E_j),$$

where $\otimes$ denotes the Kronecker (tensor) product. Since

$$\begin{bmatrix} 1 & \lambda_j \\ \lambda_j & \lambda_j^2 \end{bmatrix} \geq 0$$

and by positivity $\Phi(E_j) \geq 0$ $(j = 1, \ldots, m)$, we have

$$\begin{bmatrix} 1 & \lambda_j \\ \lambda_j & \lambda_j^2 \end{bmatrix} \otimes \Phi(E_j) \geq 0,$$

$j = 1, \ldots, m$. Consequently

$$\begin{bmatrix} I_n & \Phi(A) \\ \Phi(A) & \Phi(A^2) \end{bmatrix} \geq 0$$

which implies (1.9) by Lemma 1.4.

In a similar way, using

$$\begin{bmatrix} \lambda_j & 1 \\ 1 & \lambda_j^{-1} \end{bmatrix} \geq 0$$

we can conclude that

$$\begin{bmatrix} \Phi(A) & I_n \\ I_n & \Phi(A^{-1}) \end{bmatrix} \geq 0$$

which implies (1.10) again by Lemma 1.4. □

Theorem 1.6 *Let* Φ *be a unital positive linear map from* M_m *to* M_n *and* f *an operator monotone function on* $[0, \infty)$. *Then for every* $A \geq 0$,

$$f(\Phi(A)) \geq \Phi(f(A)).$$

Proof. By the integral representation (1.7) it suffices to prove

$$\Phi(A)[sI + \Phi(A)]^{-1} \geq \Phi[A(sI + A)^{-1}], \quad s > 0.$$

Since $A(sI + A)^{-1} = I - s(sI + A)^{-1}$ and similarly for the left side, this is equivalent to

$$[\Phi(sI + A)]^{-1} \leq \Phi[(sI + A)^{-1}]$$

which follows from (1.10). □

Theorem 1.7 *Let Φ be a unital positive linear map from M_m to M_n and g an operator convex function on $[0, \infty)$. Then for every $A \geq 0$,*

$$g(\Phi(A)) \leq \Phi(g(A)).$$

Proof. By the integral representation (1.8) it suffices to show

$$\Phi(A)^2 \leq \Phi(A^2) \tag{1.11}$$

and

$$\Phi(A)^2[sI + \Phi(A)]^{-1} \leq \Phi[A^2(sI + A)^{-1}], \quad s > 0. \tag{1.12}$$

(1.11) is just (1.9). Since

$$A^2(sI + A)^{-1} = A - sI + s^2(sI + A)^{-1},$$

$$\Phi(A)^2[sI + \Phi(A)]^{-1} = \Phi(A) - sI + s^2[sI + \Phi(A)]^{-1},$$

(1.12) follows from (1.10). This completes the proof. □

Since $f_1(t) = t^r$ $(0 < r \leq 1)$ and $f_2(t) = \log t$ are operator monotone functions on $[0, \infty)$ and $(0, \infty)$ respectively, $g(t) = t^r$ is operator convex on $(0, \infty)$ for $-1 \leq r \leq 0$ and $1 \leq r \leq 2$, from Theorems 1.6, 1.7 we get the following corollary.

Corollary 1.8 *Let Φ be a unital positive linear map from M_m to M_n. Then*

$$\Phi(A^r) \leq \Phi(A)^r, \quad A \geq 0, \quad 0 < r \leq 1;$$

$$\Phi(A^r) \geq \Phi(A)^r, \quad A > 0, \quad -1 \leq r \leq 0 \text{ or } 1 \leq r \leq 2;$$

$$\Phi(\log A) \leq \log(\Phi(A)), \quad A > 0.$$

Given $A = (a_{ij}), B = (b_{ij}) \in M_n$, the *Hadamard product* of A and B is defined as the entry-wise product: $A \circ B \equiv (a_{ij}b_{ij}) \in M_n$. For this topic see

[52, Chapter 5]. We denote by $A[\alpha]$ the principal submatrix of A indexed by α. The following simple observation is very useful.

Lemma 1.9 *For any $A, B \in M_n$, $A \circ B = (A \otimes B)[\alpha]$ where $\alpha = \{1, n+2, 2n+3, \ldots, n^2\}$. Consequently there is a unital positive linear map Φ from M_{n^2} to M_n such that $\Phi(A \otimes B) = A \circ B$ for all $A, B \in M_n$.*

As an illustration of the usefulness of this lemma, consider the following reasoning: If $A, B \geq 0$, then evidently $A \otimes B \geq 0$. Since $A \circ B$ is a principal submatrix of $A \otimes B$, $A \circ B \geq 0$. Similarly $A \circ B > 0$ for the case when both A and B are positive definite. In other words, the Hadamard product of positive semidefinite (definite) matrices is positive semidefinite (definite). This important fact is known as the Schur product theorem.

Corollary 1.10

$$A^r \circ B^r \leq (A \circ B)^r, \quad A, B \geq 0, \quad 0 < r \leq 1; \tag{1.13}$$

$$A^r \circ B^r \geq (A \circ B)^r, \quad A, B > 0, \quad -1 \leq r \leq 0 \text{ or } 1 \leq r \leq 2; \tag{1.14}$$

$$(\log A + \log B) \circ I \leq \log(A \circ B), \quad A, B > 0. \tag{1.15}$$

Proof. This is an application of Corollary 1.8 with A there replaced by $A \otimes B$ and Φ being defined in Lemma 1.9.

For (1.13) and (1.14) just use the fact that $(A \otimes B)^t = A^t \otimes B^t$ for real number t. See [52] for properties of the Kronecker product.

For (1.15) we have

$$\begin{aligned} \log(A \otimes B) &= \frac{d}{dt}(A \otimes B)^t|_{t=0} = \frac{d}{dt}(A^t \otimes B^t)|_{t=0} \\ &= (\log A) \otimes I + I \otimes (\log B). \end{aligned}$$

This can also be seen by using the spectral decompositions of A and B. □

We remark that the inequality in (1.14) is also valid for $A, B \geq 0$ in the case $1 \leq r \leq 2$.

Given a positive integer k, let us denote the kth Hadamard power of $A = (a_{ij}) \in M_n$ by $A^{(k)} \equiv (a_{ij}^k) \in M_n$. Here are two interesting consequences of Corollary 1.10: For every positive integer k,

$$(A^r)^{(k)} \leq (A^{(k)})^r, \quad A \geq 0, \quad 0 < r \leq 1;$$

$$(A^r)^{(k)} \geq (A^{(k)})^r, \quad A > 0, \quad -1 \leq r \leq 0 \text{ or } 1 \leq r \leq 2.$$

Corollary 1.11 *For $A, B \geq 0$, the function $f(t) = (A^t \circ B^t)^{1/t}$ is increasing on $[1, \infty)$, i.e.,*

$$(A^s \circ B^s)^{1/s} \leq (A^t \circ B^t)^{1/t}, \quad 1 \leq s < t.$$

Proof. By Corollary 1.10 we have

$$A^s \circ B^s \leq (A^t \circ B^t)^{s/t}.$$

Then applying the Löwner-Heinz inequality with the power $1/s$ yields the conclusion. □

Let P_n be the set of positive semidefinite matrices in M_n. A map Ψ from $P_n \times P_n$ into P_m is called *jointly concave* if

$$\Psi(\lambda A + (1-\lambda)B, \lambda C + (1-\lambda)D) \geq \lambda\Psi(A, C) + (1-\lambda)\Psi(B, D)$$

for all $A, B, C, D \geq 0$ and $0 < \lambda < 1$.

For $A, B > 0$, the *parallel sum* of A and B is defined as

$$A : B = (A^{-1} + B^{-1})^{-1}.$$

Note that $A : B = A - A(A+B)^{-1}A$ and $2(A : B) = \{(A^{-1} + B^{-1})/2\}^{-1}$ is the *harmonic mean* of A, B. Since $A : B$ decreases as A, B decrease, we can define the parallel sum for general $A, B \geq 0$ by

$$A : B = \lim_{\epsilon \downarrow 0} \{(A + \epsilon I)^{-1} + (B + \epsilon I)^{-1}\}^{-1}.$$

Using Lemma 1.4 it is easy to verify that

$$A : B = \max\left\{ X \geq 0 : \begin{bmatrix} A+B & A \\ A & A-X \end{bmatrix} \geq 0 \right\}$$

where the maximum is with respect to the Löwner partial order. From this extremal representation it follows readily that the map $(A, B) \mapsto A : B$ is jointly concave.

Lemma 1.12 *For* $0 < r < 1$ *the map*

$$(A, B) \mapsto A^r \circ B^{1-r}$$

is jointly concave in $A, B \geq 0$.

Proof. It suffices to prove that the map $(A, B) \mapsto A^r \otimes B^{1-r}$ is jointly concave in $A, B \geq 0$, since then the assertion will follow via Lemma 1.9.

We may assume $B > 0$. Using $A^r \otimes B^{1-r} = (A \otimes B^{-1})^r (I \otimes B)$ and the integral representation

$$t^r = \frac{\sin r\pi}{\pi} \int_0^\infty \frac{s^{r-1}t}{s+t} ds \quad (0 < r < 1)$$

we get

$$A^r \otimes B^{1-r} = \frac{\sin r\pi}{\pi} \int_0^\infty s^{r-1}(A \otimes B^{-1})(A \otimes B^{-1} + sI \otimes I)^{-1}(I \otimes B)ds.$$

Since $A \otimes B^{-1}$ and $I \otimes B$ commute, it is easy to see that

$$(A \otimes B^{-1})(A \otimes B^{-1} + sI \otimes I)^{-1}(I \otimes B) = (s^{-1}A \otimes I) : (I \otimes B).$$

We know that the parallel sum is jointly concave. Thus the integrand above is also jointly concave, and so is $A^r \otimes B^{1-r}$. This completes the proof. □

Corollary 1.13 *For $A, B, C, D \geq 0$ and $p, q > 1$ with $1/p + 1/q = 1$,*

$$A \circ B + C \circ D \leq (A^p + C^p)^{1/p} \circ (B^q + D^q)^{1/q}.$$

Proof. This is just the mid-point joint concavity case $\lambda = 1/2$ of Lemma 1.12 with $r = 1/p$. □

Let $f(x)$ be a real-valued differentiable function defined on some real interval. We denote by $\Delta f(x, y) \equiv [f(x) - f(y)]/(x - y)$ the difference quotient where $\Delta f(x, x) \equiv f'(x)$.

Let $H(t) \in M_n$ be a family of Hermitian matrices for t in an open real interval (a, b) and suppose the eigenvalues of $H(t)$ are contained in some open real interval Ω for all $t \in (a, b)$. Let $H(t) = U(t)\Lambda(t)U(t)^*$ be the spectral decomposition with $U(t)$ unitary and $\Lambda(t) = \text{diag}(\lambda_1(t), \ldots, \lambda_n(t))$. Assume that $H(t)$ is continuously differentiable on (a, b) and $f : \Omega \to \mathbb{R}$ is a continuously differentiable function. Then it is known [52, Theorem 6.6.30] that $f(H(t))$ is continuously differentiable and

$$\frac{d}{dt} f(H(t)) = U(t)\{[\Delta f(\lambda_i(t), \lambda_j(t))] \circ [U(t)^* H'(t) U(t)]\} U(t)^*.$$

Theorem 1.14 *For $A, B \geq 0$ and $p, q > 1$ with $1/p + 1/q = 1$,*

$$A \circ B \leq (A^p \circ I)^{1/p}(B^q \circ I)^{1/q}.$$

Proof. Denote

$$C \equiv (A^p \circ I)^{1/p} \equiv \text{diag}(\lambda_1, \ldots, \lambda_n),$$
$$D \equiv (B^q \circ I)^{1/q} \equiv \text{diag}(\mu_1, \ldots, \mu_n).$$

By continuity we may assume that $\lambda_i \neq \lambda_j$ and $\mu_i \neq \mu_j$ for $i \neq j$.

Using the above differential formula we compute

$$\left.\frac{d}{dt}(C^p + tA^p)^{1/p}\right|_{t=0} = X \circ A^p$$

and

$$\frac{d}{dt}(D^q + tB^q)^{1/q}\Big|_{t=0} = Y \circ B^q$$

where $X = (x_{ij})$ and $Y = (y_{ij})$ are defined by

$$x_{ij} = (\lambda_i - \lambda_j)(\lambda_i^p - \lambda_j^p)^{-1} \text{ for } i \neq j \text{ and } x_{ii} = p^{-1}\lambda_i^{1-p},$$

$$y_{ij} = (\mu_i - \mu_j)(\mu_i^q - \mu_j^q)^{-1} \text{ for } i \neq j \text{ and } y_{ii} = q^{-1}\mu_i^{1-q}.$$

By Corollary 1.13

$$C \circ D + tA \circ B \leq (C^p + tA^p)^{1/p} \circ (D^q + tB^q)^{1/q}$$

for any $t \geq 0$. Therefore, via differentiation at $t = 0$ we have

$$\begin{aligned} A \circ B &\leq \frac{d}{dt}(C^p + tA^p)^{1/p} \circ (D^q + tB^q)^{1/q}|_{t=0} \\ &= X \circ A^p \circ D + C \circ Y \circ B^q \\ &= (X \circ I)(A^p \circ I)D + C(Y \circ I)(B^q \circ I) \\ &= p^{-1}C^{1-p}(A^p \circ I)D + q^{-1}CD^{1-q}(B^q \circ I) \\ &= (A^p \circ I)^{1/p}(B^q \circ I)^{1/q}. \end{aligned}$$

This completes the proof. □

We will need the following result in the next section and in Chapter 3. See [17] for a proof.

Theorem 1.15 *Let f be an operator monotone function on $[0, \infty)$, g an operator convex function on $[0, \infty)$ with $g(0) \leq 0$. Then for every contraction C, i.e., $\|C\|_\infty \leq 1$ and every $A \geq 0$,*

$$f(C^*AC) \geq C^*f(A)C, \tag{1.16}$$

$$g(C^*AC) \leq C^*g(A)C. \tag{1.17}$$

Notes and References. As already remarked, Theorem 1.3 is part of the Löwner theory. The inequality (1.16) in Theorem 1.15 is due to F. Hansen [43] while the inequality (1.17) is proved by F. Hansen and G. K. Pedersen [44]. All other results in this section are due to T. Ando [3, 8].

1.3 Inequalities for Matrix Powers

The purpose of this section is to prove the following result.