



Københavns Universitet



Convex and Monotone Matrix Functions and Their Applications in Operator Theory

Hansen, Frank

Publication date:
1983

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Hansen, F. (1983). *Convex and Monotone Matrix Functions and Their Applications in Operator Theory*. Museum Tusulanum. Rapporter, No. 3

KØBENHAVNS UNIVERSITET

MATEMATISK INSTITUT

CONVEX AND MONOTONE MATRIX FUNCTIONS

AND THEIR

APPLICATIONS IN OPERATOR THEORY

Frank Hansen

Oktober 1983

Rapport nr. 3

Afhandlingen

Convex and monotone matrix functions and their applications
in operator theory

er i forbindelse med følgende fire publikationer

- [A] Hansen, F.: An operator inequality. Math. Ann. 246,
249-250 (1980),
- [B] Hansen, F.: Selfadjoint means and operator monotone
functions. Math. Ann. 256, 29-35 (1981),
- [C] Hansen, F.: Means and concave products of positive
semi-definite matrices. Math. Ann. 264, 119-128 (1983),
- [D] Hansen, F., Pedersen, G. K.: Jensen's inequality for
operators and Löwner's theorem. Math. Ann. 258,
229-241 (1982)

indleveret til bedømmelse for den naturvidenskabelige licentiat-
grad.

Det nedsatte bedømmelsesudvalg har i en indstilling til Københavns
Universitets naturvidenskabelige fakultet anbefalet, at cand.scient.
Frank Hansen tildeles licentiatgraden på grundlag af det indleve-
rede materiale.

Offentlig licentiatforelæsning med titlen

Jensens ulighed for operatorer

afholdes i auditorium 10 på H. C. Ørsted Institutet, fredag den
28. oktober 1983 kl. 14.

CONVEX AND MONOTONE MATRIX FUNCTIONS
AND THEIR
APPLICATIONS IN OPERATOR THEORY

Content:

- 1^o Preliminaries.
- 2^o Jensen's Inequality for Operators.
- 3^o Characterizations of Convex and Monotone Matrix Functions of Arbitrary Order.
- 4^o Exponential Ordering Monotone Functions.
- 5^o Means and Concave Products.

Frank Hansen
Matematisk Institut
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark

1° Preliminaries

Let \mathbf{M}_n denote the algebra of complex matrices of order n . The adjoint operation, denoted $*$, is an involution on \mathbf{M}_n . An element $x \in \mathbf{M}_n$ is said to be selfadjoint if $x^* = x$. The selfadjoint elements admit a spectral decomposition

$$x = u^* \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} u ,$$

where u is a unitary and $\lambda_1, \dots, \lambda_n$ are real numbers. The spectrum of x is the set of values taken by the numbers $\lambda_1, \dots, \lambda_n$. If f is a complex function defined on the spectrum of x , then we set

$$f(x) = u^* \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} u .$$

The matrix $f(x)$ is selfadjoint, if f is real.

A selfadjoint element $x \in \mathbf{M}_n$ is said to be positive, if the spectrum is a subset of the positive half-line $[0, \infty[$. This is equivalent to requiring the matrix x to be positive semi-definite. The set of positive elements in \mathbf{M}_n is a convex cone and is denoted by \mathbf{M}_n^+ . The associated order structure on the real vector space of selfadjoint elements in \mathbf{M}_n will play a crucial role in the sequel. It is defined by setting $x \leq y$, if $y-x \in \mathbf{M}_n^+$.

The spectral theorem, the functional calculus and the order structure for selfadjoint matrices, which are sketched above, can be generalized to (even unbounded) operators on a Hilbert space. Likewise they play an important role in the theory of

C*-algebras. We shall occasionally use these much more general settings, but it is important to notice, that almost all of the exhibited results can be appreciated in the context of matrices of arbitrary high order. The generalizations to operators on a Hilbert space are straightforward and do not introduce phenomena, which are not already inherent in the theory for matrices.

We shall say, that a real function f on an interval I (of any type) is operator monotone, if for each n in \mathbb{N} and every pair x, y of selfadjoint elements in \mathbb{M}_n with spectra in I , the condition $x \leq y$ implies $f(x) \leq f(y)$. Likewise we say, that f is operator convex, if for each n in \mathbb{N}

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all selfadjoint x, y in \mathbb{M}_n with spectra in I and every $\lambda \in [0, 1]$. From this definition it is clear, that a pointwise limit of operator monotone (respectively operator convex) functions is again operator monotone (respectively operator convex). The functions $t \rightarrow t(1-\alpha t)^{-1}$ (respectively $t \rightarrow t^2(1-\alpha t)^{-1}$) defined on $] -1, 1[$ are operator monotone (respectively operator convex) for $\alpha \in [0, 1]$.

2^o Jensen's Inequality for Operators

In 1905 Jensen showed [27,28] that a (mid-point) convex, continuous function f on an interval I satisfies

$$(J) \quad f(\sum \lambda_i t_i) \leq \sum \lambda_i f(t_i)$$

for any convex combination $\{\lambda_i\}$ of points $\{t_i\}$ in I . Taking

$$a = \begin{pmatrix} \lambda_1^{\frac{1}{2}} & 0 \dots 0 \\ \vdots & \vdots \\ \lambda_n^{\frac{1}{2}} & 0 \dots 0 \end{pmatrix}, \quad x = \begin{pmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_n \end{pmatrix}$$

and considering f as a function on selfadjoint matrices (with $f(0) \leq 0$), the inequality (J) reads

$$(JO) \quad f(a^*xa) \leq a^*f(x)a.$$

We shall say that a real function f on I satisfies Jensen's Operator Inequality if (JO) holds for any selfadjoint $x \in \mathbf{M}_n$ with spectrum in I and every $a \in \mathbf{M}_n$ with $\|a\| \leq 1, n = 1, 2, \dots$.

A technicality comes in here. Since zero may be an eigenvalue of a (indeed, $a = 0$ is not excluded), zero may belong to the spectrum of a^*xa . Thus the interval I must contain zero in order for (JO) to be meaningful. Moreover, the element a^*xa really is a non-commutative generalization of a degenerate convex combination of x and 0 (i.e. $\sum \lambda_i \leq 1$). The "correct" generalization of a convex combination of x and y would seem to be $a^*xa + b^*yb$, where $a^*a + b^*b = 1$, cf. [25, Theorem 2.1(iii)]. But we shall stick to (JO). Anyway it is clear, that a convex function f will satisfy Jensen's inequality (J) for all degenerate convex combinations, precisely when $0 \in I$ and $f(0) < 0$.

If f is a real function on the half-open interval $[0, \alpha[$ (with $\alpha \leq \infty$), the following conditions are equivalent [25, Theorem 2.1]:

- (i) f is operator convex and $f(0) \leq 0$.
- (ii) $f(a^*xa) \leq a^*f(x)a$ for every selfadjoint $x \in \mathbb{M}_n$ with spectrum in $[0, \alpha[$, all $a \in \mathbb{M}_n$ with $\|a\| \leq 1$ and $n = 1, 2, \dots$.
- (iii) $f(a^*xa + b^*yb) \leq a^*f(x)a + b^*f(y)b$ for all selfadjoint $x, y \in \mathbb{M}_n$ with spectra in $[0, \alpha[$, all $a, b \in \mathbb{M}_n$ with $a^*a + b^*b \leq 1$ and $n = 1, 2, \dots$.
- (iv) $f(pxp) \leq pf(x)p$ for every selfadjoint $x \in \mathbb{M}_n$ with spectrum in $[0, \alpha[$, every projection $p \in \mathbb{M}_n$ and $n = 1, 2, \dots$.

It is apparent from the proof of the implication (i) \Rightarrow (ii), that it is sufficient to assume, that f is mid-point operator convex. That condition again implies operator convexity of f , hence continuity in the open interval $]0, \alpha[$. This is in contrast to the situation for functions, which are only mid-point convex.

Let $\pi: A \rightarrow B(H)$ be a positive, linear contraction on a C^* -algebra A . As a corollary of the result above, we obtain that

$$f(\pi(x)) \leq \pi(f(x))$$

for each continuous, operator convex function f defined on $[0, \alpha[$ with $f(0) \leq 0$, and each selfadjoint x in A with spectrum in $[0, \alpha[$, cf. [25, Corollary 2.2]. The result is also proved in [12, 11] under the additional hypothesis that

$\pi(1) = 1$. However, this hypothesis makes the proof very easy. It implies, that the contraction a in the Stinespring decomposition of π satisfies $a^*a = 1$, cf. the proof of [22, Corollary]. Note that with $f(t) = t^2$, we obtain Kadison's generalization of the Cauchy-Schwarz inequality: $\pi(x)^2 \leq \pi(x^2)$. This is fitting, since Jensen himself first used his inequality to prove Cauchy's inequality, cf. [27,28].

The classes of operator convex and operator monotone functions are closely related. Indeed, a continuous real function f on the interval $]0, \alpha[$ is operator convex with $f(0) \leq 0$, if and only if

(v) $t \rightarrow t^{-1}f(t)$ is operator monotone on $]0, \alpha[$,

cf. [25, Theorem 2.4]. Another relationship is given by the following result: If $f \leq 0$ is a continuous, real function on the half-line $]0, \infty[$, then the conditions (i), (ii), (iii), (iv) and (v) are again equivalent to the condition

(vi) $-f$ is operator monotone,

cf. [22, Theorem] and [25, Theorem 2.5]. The set of strictly positive, operator monotone functions on $]0, \infty[$ plays an important role. The functions are continuous even analytic as will be shown later. The set is left invariant by the two involutions $\#$ and $*$ defined by

$$f^\#(t) = tf(t)^{-1} \quad \text{and} \quad f^*(t) = tf(t^{-1}),$$

cf. [25, Corollary 2.6].

Some of the implications linking the conditions (i) to (vi)

are previously known. Thus the biimplication $(i) \Leftrightarrow (iv)$ is due to Davis [12], and the biimplication $(i) \Leftrightarrow (v)$ to Bendat and Sherman [10]. The implication $(vi) \Rightarrow (ii)$ is proved in [22]. An alternate proof based on the Kubo-Ando theory for means is given by Fujii, [19]. The latter article is received and published first, but note the acknowledgement. Along the same lines a proof for the implication $(i) \Rightarrow (ii)$ can be found in [29]. This proof, due to Kainuma and Nakamura, actually deserves priority, although our proof was obtained independently. However, the proofs given in [19] and [29] both require Löwner's theorem. Our point is to show how Jensen's Operator Inequality can be used as a tool to fit together the conditions (i) , (v) and (vi) with purely algebraic methods, thus serving as the basis for a simplified proof of Löwner's theorem.

3⁰ Characterizations of Convex and Monotone Matrix Functions of Arbitrary Order.

Consider an open interval I in \mathbb{R} and a continuously differentiable function f on I . Fix n and take $x = x^*$ in M_n with $Sp(x) \subset I$. If $\{e_{ij} | 1 \leq i, j \leq n\}$ is a system of matrix units for M_n such that $x = \sum \lambda_i e_{ii}$, we shall denote by $f^{[1]}(x)$ the element in M_n with

$$f^{[1]}(x)_{ij} = (\lambda_i - \lambda_j)^{-1} (f(\lambda_i) - f(\lambda_j)) \quad \text{if } \lambda_i \neq \lambda_j,$$

$$f^{[1]}(x)_{ij} = f'(\lambda_i) \quad \text{if } \lambda_i = \lambda_j.$$

C. Davis proved in [13, III], that

$$\lim_{\epsilon} \epsilon^{-1} (f(x + \epsilon h) - f(x)) = f^{[1]}(x) \cdot \underset{x}{\dot{h}}$$

for every selfadjoint h in M_n , where $\underset{x}{\dot{h}}$ denotes the Hadamard product of matrices in a basis that diagonalizes x , cf. also [25, Lemma 3.1]. The element $f^{[1]}(x)$ in many ways resembles the ordinary differential. The weakness of it is, of course, that the Hadamard product depends on the basis, so that neither the formulation nor the proof of, say, the mean value theorem is obvious. For our purposes the next result replaces the mean value theorem.

If $t \rightarrow x(t)$ is a C^1 -function from $[0, 1]$ to the space of selfadjoint $n \times n$ matrices with spectra in the open interval I , and if $f \in C^1(I)$, then

$$f(x(1)) - f(x(0)) = \int_0^1 f^{[1]}(x(t)) \underset{x(t)}{\dot{x}(t)} x'(t) dt,$$

cf. [25, Lemma 3.3].

As a consequence of the two results above, we obtain that a C^1 -function f on an open interval I is operator monotone, if and only if $f^{[1]}(x) \geq 0$ for every selfadjoint x in M_n with $Sp(x) \subset I$, and every n in \mathbb{N} , cf. [25, Proposition 3.4]. This characterization of operator monotonicity by positive (generalized) differentials is very close to Löwner's original characterization [37, p.183]. The proof used here for the necessity of the condition goes back to Daleckij and Krein; see the discussion in [13, III]. We noticed that the sufficiency of the condition (i.e., $f^{[1]} \geq 0$) can be obtained quite easily integrating the differential. Unlike Löwner's original characterization, the one we use requires the function to be differentiable; and thus it becomes necessary to establish separately, that operator monotone functions are continuously differentiable. For functions defined on a finite open interval we can assume, without loss of generality, that the interval is $] -1, 1[$. First we show, that if $f:] -1, 1[\rightarrow \mathbb{R}$ is continuous and operator monotone, then the function $t \rightarrow (t+\lambda)f(t)$ is operator convex for every $\lambda \in [-1, 1]$, [25, Lemma 3.5]. This result is an almost immediate consequence of the biimplication (i) \Leftrightarrow (v) in Section 2. Next we obtain, that every operator monotone function f on $] -1, 1[$ is continuously differentiable, [25, Theorem 3.6]. The proof uses the same smoothing technique as Nagy in [45], but is simplified by use of the previous result.

If $f \in C^2(]-1, 1[)$ with $f(0) = 0$, and $\{e_{ij} \mid 0 \leq i, j \leq n\}$ is a system of matrix units for M_{n+1} , then with

$$h = \sum_{i=1}^n e_{i0} + e_{0i}, \quad q = 1 - e_{00}, \quad x = \sum_{i=1}^n \lambda_i e_{ii},$$

where $|\lambda_i| < 1$, we have

$$\lim_{\varepsilon} \varepsilon^{-2} (f(x+\varepsilon h) - f(x) - \varepsilon f^{[1]}(x) \dot{x} h) \mathfrak{q} = g^{[1]}(x),$$

where $g(t) = t^{-1}f(t)$ (and $g(0) = f'(0)$), cf. [13,III] and [25, Lemma 3.7]. This technical result together with our characterization of operator monotonicity by positive (generalized) differentials entail the following assertion [25, Lemma 3.8]:

If $f \in C^2(]-1,1[)$ with $f(0) = 0$ and f is operator convex, then the function $g: t \rightarrow t^{-1}f(t)$ is operator monotone. This is an improvement (with regard to our purposes) over the implication (i) \Rightarrow (v) in Section 2, which only gives that the function $t \rightarrow (t+1)^{-1}f(t)$ is operator monotone. Applying the assertion above together with [25, Lemma 3.5] and the same smoothing technique as in the proof of [25, Theorem 3.6], we obtain

(BS) If f is operator monotone on $]-1,1[$ and $f(0) = 0$, then the function $t \rightarrow (1+\lambda t^{-1})f(t)$ is operator monotone for $|\lambda| \leq 1$.

The result is also an immediate consequence of [10, Theorem 2.8] due to Bendat and Sherman. But we have established it as a prelude to the integral representation for operator monotone functions, not a corollary.

From now on we shall denote by K the set of operator monotone functions f on $]-1,1[$, such that $f(0) = 0$ and $f'(0) = 1$. The set K is clearly convex, but since

$$\begin{aligned} f(t) &\leq t(1-t)^{-1} && \text{for } t \geq 0; \\ f(t) &> t(1+t)^{-1} && \text{for } t < 0; \end{aligned}$$

for each $f \in K$, cf. [25, Lemma 4.1], we can also conclude that K is compact in the topology of pointwise convergence, [25, Lemma 4.2]. We furthermore obtain, that for each $f \in K$, the derivative f' is differentiable in 0 with $|f''(0)| \leq 2$, cf. [25, Corollary 3.10 and Lemma 4.10]. The main idea of our approach is to determine the form of the extreme points in K . We have established the property (BS) with this purpose in mind. For any f in K define

$$g_\lambda(t) = (1 + \frac{1}{2}\lambda f''(0))^{-1} ((1 + \lambda t^{-1})f(t) - \lambda)$$

for $|\lambda| < 1$. This definition is meaningful because $|f''(0)| \leq 2$. It follows from the property (BS) that g_λ is operator monotone. In fact the constants have been chosen so that $g_\lambda \in K$. It is immediate that

$$f = \frac{1}{2}(1 + \frac{1}{2}\lambda f''(0))g_\lambda + \frac{1}{2}(1 - \frac{1}{2}\lambda f''(0))g_{-\lambda}.$$

If therefore f is extreme, then $f = g_\lambda$; or

$$f(t) = t(1 - \frac{1}{2}f''(0)t)^{-1}, \text{ [25, Proposition 4.3].}$$

Since K is convex and compact, Krein-Milman's theorem can be applied; and it is quite straightforward to prove, that for each $f \in K$, there is a probability measure μ on $[-1, 1]$ such that

$$f(t) = \int t(1 - \alpha t)^{-1} d\mu(\alpha).$$

From this expression we see, that f has an extension, necessarily unique, to a holomorphic function in $\mathbb{C} \setminus \mathbb{R}$. Moreover, for every continuous function φ on \mathbb{R} with compact support, we have

$$\int \varphi(\alpha) d\mu(\alpha) = \lim_{\varepsilon} \pi^{-1} \int \varphi(s) \operatorname{Im} f((s-i\varepsilon)^{-1}) ds .$$

This proves that μ is unique. We learn thus, a posteriori, that K is a Bauer simplex with $\partial K = \{t(1-\alpha t)^{-1} \mid \alpha \in [-1, 1]\}$, [25, Theorem 4.4]. In general we obtain that to each non-constant operator monotone function f on $] -1, 1[$, there is a unique probability measure μ on $[-1, 1]$ such that

$$f(t) = f(0) + f'(0) \int t(1-\alpha t)^{-1} d\mu(\alpha) .$$

One of Löwner's characterizations of operator monotonicity says that a non-constant function is operator monotone, if and only if it admits an analytic continuation into the upper half-plane $\{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$, which leaves the half-plane invariant, cf. [37, 10]. This characterization comes as an immediate consequence of the result above and [15, Sect. 2, Theorem I]. The property (BS), which is the key to our proof, can also be used to considerably simplify Korányi's proof of Löwner's theorem; see [25, Remark 4.6].

The integral representation for operator monotone functions in conjunction with [25, Lemma 3.8] gives the following representation for operator convex functions ([25, Theorem 4.5]): For each non-linear operator convex function f on $] -1, 1[$ there is a unique probability measure μ on $[-1, 1]$ such that

$$f(t) = f(0) + f'(0)t + \frac{1}{2}f''(0) \int t^2(1-\alpha t)^{-1} d\mu(\alpha) .$$

Next, we turn our attention to operator monotone functions defined on \mathbb{R}_+ . The conform mapping $t \rightarrow \frac{1+t}{1-t} = 1 + \frac{2t}{1-t}$ is operator monotone on $] -1, 1[$ and maps the interval bijectively on \mathbb{R}_+ . The inverse mapping is also operator monotone. We are

thus able to conclude from the previous results, that a non-constant real function defined on \mathbb{R}_+ is operator monotone, if and only if it admits an analytic continuation into the upper half-plane $\{z \in \mathbb{C} \mid \text{Im } z > 0\}$, which maps the half-plane into itself. This shows, for example, that the logarithm is operator monotone on \mathbb{R}_+ , while the exponential function is not. Of special interest is the sub-class of positive functions. We shall normalize them conveniently and put

$$H = \{f: \mathbb{R}_+ \rightarrow \mathbb{R}_+ \mid f(1) = 1, f \text{ is operator monotone}\}.$$

It follows from function-theoretical results [26, Hilfssatz 5], cf. also [10,33], that a function f defined on \mathbb{R}_+ belongs to the class H , if and only if there is a probability measure μ on the extended half-line $[0, \infty]$ such that

$$f(t) = \int \frac{t(1+\lambda)}{t+\lambda} d\mu(\lambda) \quad \forall t \in \mathbb{R}_+,$$

where $\frac{t(1+\lambda)}{t+\lambda}$ is defined as t for $\lambda = \infty$.

Finally, we shall sketch how this result can also be obtained by functional analytical methods. The set H is convex and compact in the topology of pointwise convergence. It is left invariant by the involution $*$ defined in Section 2. For a function $f \in H$, we obtain by calculation that $(f^*)'(1) = 1 - f'(1)$; in particular, $\lambda = f'(1) \in [0,1]$. Instead of the property (BS), we could as well and with the same methods have proved, that H is left invariant by the mapping T defined by

$$(Tf)(t) = \frac{1}{f'(1)} \cdot t \frac{f(t)-1}{t-1}.$$

Take now $f \in H$ and set $\lambda = f'(1)$. The calculation

$$\lambda(Tf)^* + (1-\lambda)Tf^* = f^*$$

shows that if f^* is extreme in H , then it must be a fixpoint under T . But the fixpoints under T are exactly the functions $t \rightarrow \frac{t(t+\lambda)}{t+\lambda}$, where $\lambda \in [0, \infty]$. Note that the involution $*$ maps the fixpoint parametrized by λ on the fixpoint parametrized by λ^{-1} . The statement including unicity of the measure μ now follows by standard arguments, cf. the case of operator monotone functions defined on $] -1, 1[$ treated above. Combining this representation of H with the biimplication (i) \Leftrightarrow (v) in Section 2, we obtain the following result:

A continuous function $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, satisfying

$\lim_{\varepsilon \rightarrow 0} f(\varepsilon) = 0$, is operator convex, if and only if there is a positive, finite measure μ on $[0, \infty]$ such that

$$f(t) = \int \frac{t^2(1+\lambda)}{t+\lambda} d\mu(\lambda).$$

4⁰ Exponential Ordering Monotone Functions.

Let x, y denote selfadjoint elements in M_n , $n = 1, 2, \dots$
We write $x \preccurlyeq y$ if $\exp x \leq \exp y$. It is easily verified that

- (i) $x \preccurlyeq x$,
- (ii) $x \preccurlyeq y$ and $y \preccurlyeq z$ imply $x \preccurlyeq z$,
- (iii) $x \preccurlyeq y$ and $y \preccurlyeq x$ imply $x = y$,

which show that \preccurlyeq is an order relation. Since the logarithm is operator monotone on \mathbb{R}_+ , it follows that \preccurlyeq is stronger than \leq . Indeed, $x \preccurlyeq y$ implies $x \leq y$. Furthermore, since the functions $f(t) = t^\lambda$, $t \in \mathbb{R}_+$ are operator monotone for $\lambda \in [0, 1]$, we obtain

- (iv) $x \preccurlyeq y$ implies $\lambda x \preccurlyeq \lambda y$ for $\lambda \in [0, 1]$.

Finally, using that x and 1 commute, we get

- (v) $x \preccurlyeq y$ implies $x + \lambda \preccurlyeq y + \lambda$ for $\lambda \in \mathbb{R}$.

We shall call the order relation \preccurlyeq for exponential ordering. A function $F:]a, b[\rightarrow \mathbb{R}$ is said to be exponential ordering monotone, if $x \preccurlyeq y$ implies $F(x) \preccurlyeq F(y)$ for all selfadjoint $x, y \in M_n$ with spectra in $]a, b[$ and every n in \mathbb{N} . We denote by E the class of exponential ordering monotone functions defined on \mathbb{R} .

Let $F \in E$. The function

$$\Phi(F)(t) = \exp F(\log t), \quad t \in \mathbb{R}_+$$

maps \mathbb{R}_+ into \mathbb{R}_+ . The mapping Φ is a bijection of E onto the class P of operator monotone functions defined on and with values in \mathbb{R}_+ ; [23, Lemma 2.2]. As a Corollary of Löwner's

theory we obtain:

A non-constant function $F: \mathbb{R} \rightarrow \mathbb{R}$ belongs to the class \mathcal{E} if and only if F admits an analytic continuation into the strip $\{z \in \mathbb{C} \mid 0 < \text{Im } z < \pi\}$ which leaves the strip invariant, [23, Theorem 2.3]. A similar result holds for real functions defined on an open interval of any type.

It does not follow in any easy way from the definition of \mathcal{E} that \mathcal{E} is convex, but this is clear from the characterization given above. The mapping $\Phi: \mathcal{E} \rightarrow \mathcal{P}$ is thus a bijection between two convex sets. We want to derive an integral representation for the functions in \mathcal{E} similar to the expression for the operator monotone functions in \mathcal{P} , cf. Section 3. Since Φ is not affine, this can not be done merely by transport of structure. Let $F \in \mathcal{E}$. The function

$$\varphi(t) = F(\log t): \mathbb{R}_+ \rightarrow \mathbb{R}$$

admits an analytic continuation into the upper half-plane which maps the upper half-plane into the subset $\{z \in \mathbb{C} \mid 0 \leq \text{Im } z < \pi\}$. Hence φ is a Pick function and consequently admits a unique canonical representation as given in [15, Sect. 2, Theorem I]. The boundedness of $\text{Im } \varphi(\zeta)$ entails that the measure μ in the representation is absolutely continuous with respect to Lebesgue measure; and also that the linear part in the representation is vanishing. From this and the characterization of \mathcal{E} given above, we are able to obtain the following result:

Let $F \in \mathcal{E}$. Then F admits a canonical representation

$$F(t) = \beta + \int_{-\infty}^0 \left(\frac{1}{\lambda - \exp t} - \frac{\lambda}{\lambda^2 + 1} \right) h(\lambda) d\lambda \quad \forall t \in \mathbb{R},$$

where $\beta \in \mathbb{R}$, $h:]-\infty, 0] \rightarrow [0, 1]$ is a measurable function and $d\lambda$ denotes Lebesgue measure. The equivalence class containing h is uniquely determined by F . Any function F given on the indicated form belongs to the class \mathcal{E} , [23, Theorem 2.4]. We obtain by computation that $\beta = \operatorname{Re} F(i\frac{\pi}{2})$. An operator monotone function defined on the whole real line is linear. No such phenomenon occurs, as we can see, for exponential ordering monotone functions.

Let F be a function in \mathcal{E} with canonical representation as given above. With standard analytical methods we obtain, that F satisfies $F(-t) = -F(t) \forall t \in \mathbb{R}$, if and only if $\beta = 0$ and $h(\lambda^{-1}) = h(\lambda)$ for almost all $\lambda \in]-\infty, 0[$, [23, Theorem 3.1]. As a corollary we get:

A function $F: \mathbb{R} \rightarrow \mathbb{R}$ is in the class \mathcal{E} and satisfies $F(-t) = -F(t) \forall t \in \mathbb{R}$, if and only if F admits a canonical representation

$$F(t) = \int_{-1}^0 \left(\frac{1}{\lambda - \exp t} + \frac{\exp t}{1 - \lambda \exp t} \right) h(\lambda) d\lambda \quad \forall t \in \mathbb{R},$$

where $h: [-1, 0] \rightarrow [0, 1]$ is a measurable function and $d\lambda$ denotes Lebesgue measure. The equivalence class containing h is uniquely determined by F , [23, Corollary 3.2]. The bijection $\Phi: \mathcal{E} \rightarrow \mathcal{P}$ has the property that $\Phi(F)(t^{-1}) = \Phi(F)(t)^{-1} \forall t \in \mathbb{R}_+$, if and only if $F(-t) = -F(t) \forall t \in \mathbb{R}$. For example, Φ maps the functions $F(t) = \alpha t: \mathbb{R} \rightarrow \mathbb{R}$, $\alpha \in [0, 1]$ onto the functions $f(t) = t^\alpha: \mathbb{R}_+ \rightarrow \mathbb{R}_+$. This property of Φ entails the following result, which is not easy to attack directly:

A function $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is operator monotone and satisfies $f(t)^{-1} = f(t^{-1}) \forall t \in \mathbb{R}_+$, if and only if f admits a ca-

nonical representation

$$f(t) = \exp \int_{-1}^0 \left(\frac{1}{\lambda-t} + \frac{t}{1-\lambda t} \right) h(\lambda) d\lambda \quad \forall t \in \mathbb{R}_+,$$

where $h: [-1,0] \rightarrow [0,1]$ is a measurable function and $d\lambda$ denotes Lebesgue measure. The equivalence class containing h is uniquely determined by f , [23, Theorem 1.1]. In [33, Sect. 4], the question was raised whether the functions $f(t) = t^\alpha$, $\alpha \in [0,1]$ are the only operator monotone functions defined on \mathbb{R}_+ , which satisfy the functional equation $f(t)^{-1} = f(t^{-1}) \forall t \in \mathbb{R}_+$. We can now see, that this is not the case. Choosing h to be a constant $\alpha \in [0,1]$, we indeed obtain the functions $f(t) = t^\alpha$, $t \in \mathbb{R}_+$. But taking $h(\lambda) = \lambda+1$, $\lambda \in [-1,0]$, we obtain a new operator monotone function

$$f(t) = \left(\frac{t}{1+t} \right)^{1+t} (1+t)^{1+\frac{1}{t}}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$$

satisfying the functional equation.

5⁰ Means and Concave Products

Take $A, B \in \mathbb{M}_n$, $n \in \mathbb{N}$. The harmonic mean (or twice the parallel sum) of A and B , denoted $A!B$, can be defined by

$$((A!B)\xi|\xi) = 2 \inf \left\{ (A\xi|\xi) + (B\eta|\eta) \mid \xi + \eta = \xi \right\}$$

for each $\xi \in \mathbb{C}^n$. If A and B commute, then $A!B = 2AB(A+B)^{-1}$. This remains true even for non-commuting matrices A and B under certain interpretations of $2AB(A+B)^{-1}$ (for example, $2(A^{-1}+B^{-1})^{-1}$ if all inverses make sense). See [38, 5,31] for a detailed discussion. From the definition given above, we immediately conclude that the harmonic mean enjoys the following properties:

- (i) $A \leq C$ and $B \leq D$ imply $(A!B) \leq (C!D)$. If $A_k \succ A$ and $B_k \succ B$, then $(A_k!B_k) \succ (A!B)$.
- (ii) $(\lambda A + (1-\lambda)B) ! (\lambda C + (1-\lambda)D) \geq \lambda(A!C) + (1-\lambda)(B!D)$ (concavity) for $\lambda \in [0,1]$.
- (iii) $(\lambda A ! \lambda B) = \lambda(A!B)$, $\lambda \in [0, \infty[$ (homogeneity).
- (iv) $C*(A!B)C \leq (C*AC) ! (C*BC)$ for arbitrary C (the transformer inequality).
- (v) $A ! A = A$.

For $t > 0$, we set

$$A !_t B = \frac{t+1}{2t} (tA!B).$$

It is natural to extend this definition setting

$$A !_t B = A \quad \text{for } t = 0 ,$$

and

$$A !_t B = B \quad \text{for } t = \infty .$$

The so defined binary operations for positive semi-definite matrices $(!_t)_{t \in [0, \infty]}$ all enjoy the same properties as listed for the harmonic mean. Furthermore, for each probability measure μ on the extended half line $[0, \infty]$, the binary operation σ defined by

$$A \sigma B = \int (A !_t B) d\mu(t)$$

again satisfies the same properties. The binary operations defined by an integral representation as given above are the so-called means of positive semi-definite matrices. In their joint work [33], Kubo and Ando characterize them axiomatically as the binary operations σ , which satisfy conditions (i) and (iv) (with $!$ replaced by σ) and are normalized in the sense that $1 \sigma 1 = 1$.

To give a precise meaning to the expression "binary operation for positive semi-definite matrices", we shall define it as a sequence of mappings

$$\sigma_n : \mathbf{M}_n^+ \times \mathbf{M}_n^+ \rightarrow \mathbf{M}_n^+ , \quad n = 1, 2, \dots$$

such that

$$i_n(A \sigma_n B) = i_n(1_n) (i_n(A) \sigma_{n+1} i_n(B)) i_n(1_n)$$

for every $A, B \in \mathbf{M}_n^+$ and n in \mathbf{N} ; where 1_n denotes the unit in \mathbf{M}_n and $i_n : \mathbf{M}_n \rightarrow \mathbf{M}_{n+1}$ is an embedding such that $i_n(1_n) \leq 1_{n+1}$. It readily follows, that we can delete the subscript n in σ_n without introducing any ambiguity, and so we

shall do in the rest of this section.

A binary operation σ for positive semi-definite matrices is called regular, if it satisfies

(I) $U^*(A \sigma B)U = (U^*AU) \sigma (U^*BU)$ for every $A, B \in \mathbb{M}_n^+$ and each unitary $U \in \mathbb{M}_n$, $n = 1, 2, \dots$.

(II) If P is a projection in \mathbb{M}_n , $n = 1, 2, \dots$, and $A, B \in \mathbb{M}_n^+$ commute with P , then $P(A \sigma B)P = P(AP \sigma BP)P$.

If A, B and P are as in the assumptions of (II), then we can set $U = 2P - 1_n$ and use (I) to conclude that $A \sigma B$ (and $AP \sigma BP$) commutes with P . In particular, if $t, s \in [0, \infty[$, then $t1_n \sigma s1_n$ is a scalar multiple of 1_n . Applying (II), we furthermore conclude that the scalar multiple is independent of n . Consequently, $t1_n \sigma s1_n = (t \sigma s)1_n$. This concept of regularity was introduced in [24]. If we for each $t \in [0, \infty[$ define the function

$$g_t(s) = t \sigma ts, \quad s \in [0, \infty[,$$

then regularity of σ entails, that $t \sigma tA = g_t(A)$ for each $A \in \mathbb{M}_n^+$ and n in \mathbb{N} , [24, Lemma 2.1].

Let σ be a regular, binary operation for positive semi-definite matrices. The following statements are equivalent, [24, Theorem 2.2]:

- (i) $(\lambda A + (1-\lambda)B) \sigma (\lambda C + (1-\lambda)D) \geq \lambda(A \sigma C) + (1-\lambda)(B \sigma D)$ (concavity) for every $A, B, C, D \in \mathbb{M}_n^+$, $n = 1, 2, \dots$ and $\lambda \in [0, 1]$.
- (ii) $(\frac{1}{2}A + \frac{1}{2}B) \sigma (\frac{1}{2}C + \frac{1}{2}D) \geq \frac{1}{2}(A \sigma C) + \frac{1}{2}(B \sigma D)$ (mid-point concavity) for every $A, B, C, D \in \mathbb{M}_n^+$ and $n = 1, 2, \dots$.

(iii) $C^*(A \sigma B)C \leq (C^*AC) \sigma (C^*BC)$

for every $A, B \in \mathbb{M}_n^+$, $C \in \mathbb{M}_n$ with $\|C\| \leq 1$ and $n = 1, 2, \dots$.

(iv) $P(A \sigma B)P \leq (PAP) \sigma (PBP)$

for every $A, B \in \mathbb{M}_n^+$, each projection P in \mathbb{M}_n and $n = 1, 2, \dots$.

(v) $A \leq B$ and $C \leq D$ imply $(A \sigma C) \leq (B \sigma D)$

for every $A, B, C, D \in \mathbb{M}_n^+$ and $n = 1, 2, \dots$. Furthermore, if $(A_k)_{k \in \mathbb{N}}$ and $(B_k)_{k \in \mathbb{N}}$ are sequences in \mathbb{M}_n^+ , $n = 1, 2, \dots$, such that $A_k \succ A$ and $B_k \succ B$, then

$$(A_k \sigma B_k) \succ (A \sigma B).$$

We shall, in short, denote a regular, binary operation for positive semi-definite matrices, which satisfies the five equivalent conditions above for a concave product. Furthermore, if $1 \sigma 1 = 1$, we say that σ is normalized. Let now σ be a concave product for positive semi-definite matrices and take $A, B \in \mathbb{M}_n^+$, $n = 1, 2, \dots$, with $A \leq B$. If g_t , $t \in [0, \infty[$, is the function introduced above, we obtain

$$g_t(A) = t \sigma tA \leq t \sigma tB = g_t(B).$$

Hence the function g_t is operator monotone. Furthermore, condition (i) implies that g_t is operator concave, while condition (iii) gives

$$\begin{aligned} C^*g_t(A)C &= C^*(t \sigma tA)C \\ &\leq tC^*C \sigma tC^*AC \leq t \sigma tC^*AC \\ &= g_t(C^*AC), \end{aligned}$$

for every $C \in \mathbb{M}_n$ with $\|C\| \leq 1$. Compare these results with the biimplications linking conditions (i), (ii) and (vi) in Section 2 applied on $-g_t$.

The set of concave products is a convex cone. But it seems difficult to obtain a detailed description of the elements. Instead, we shall uncover one more property of the set itself and then give a necessary and sufficient condition for a concave product to be a mean. If σ is a concave product and A, B are invertible elements in \mathbb{M}_n^+ , $n = 1, 2, \dots$, then the formula

$$A \sigma^* B = A^{\frac{1}{2}} (A^{-1} \sigma A^{-\frac{1}{2}} B A^{-\frac{1}{2}}) A^{\frac{1}{2}}$$

defines a concave product σ^* , [24, Proposition 3.3]. The mapping $\sigma \rightarrow \sigma^*$ is an involution. As a corollary we obtain, that a concave product σ is a mean, if and only if it is normalized and homogeneous for positive numbers t, s , i.e.

$$(\lambda t \sigma \lambda s) = \lambda (t \sigma s) \quad \forall \lambda \in [0, \infty[,$$

cf. [24, Corollary 3.4]. The difficult part is to prove that homogeneity for positive numbers t, s implies homogeneity for arbitrary matrices $A, B \in \mathbb{M}_n^+$, $n = 1, 2, \dots$; i.e.

$$(\lambda A \sigma \lambda B) = \lambda (A \sigma B) \quad \forall \lambda \in [0, \infty[.$$

That this is true is slightly surprising, because the drawn conclusion transcends the case of commuting matrices. The validity of the result is a combined consequence of concavity and the possibility of embedding into matrix algebras of higher dimension. As an application we conclude, that the geometric mean # given by

$$A \# B = A^{\frac{1}{2}} (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^{\frac{1}{2}} A^{\frac{1}{2}}$$

for invertible matrices $A, B \in \mathbf{M}_n^+$, $n = 1, 2, \dots$, is the unique extension of the mapping $(t, s) \rightarrow \sqrt{ts}$ to a concave product, cf. [33, Theorem 3.2]. If σ is a non-trivial mean of positive semi-definite matrices [33], and $f: [0, \infty[\rightarrow [0, \infty[$ is a non-linear operator monotone function, then the mapping $(A, B) \rightarrow A \alpha B = A \sigma f(B)$ is a concave product. It is normalized, if $f(1) = 1$. Such a concave product is neither a mean, nor does it satisfy $A \alpha A = A$ for each A . It is unknown to the author, if a concave product σ , satisfying $A \sigma A = A$ for each A , necessarily must be a mean.

Let finally σ be a regular, binary operation for positive semi-definite matrices. If σ is normalized and satisfies the axiom

$$(III) \quad (A+B) \sigma (C+D) \geq (A \sigma C) + (B \sigma D)$$

for every $A, B, C, D \in \mathbf{M}_n^+$, $n = 1, 2, \dots$,

then σ is a mean, [24, Theorem 3.1]. This result offers an alternative axiomatic setting for the theory of means avoiding the transformer inequality, cf. [33]. The statement is still valid, if for any $\lambda > \frac{1}{2}$, axiom (III) is replaced by the condition

$$(\lambda A + \lambda B) \sigma (\lambda C + \lambda D) \geq \lambda (A \sigma C) + \lambda (B \sigma D)$$

for every $A, B, C, D \in \mathbf{M}_n^+$, $n = 1, 2, \dots$,

cf. [24, Remarks 3.2].

Let σ be a mean for positive semi-definite matrices. The function $f(t) = 1 \sigma t$ is called the representing function for σ . It is operator monotone and normalized and thus belongs to

the class H introduced in Section 3. The associated probability measure μ in the representation for f is the same as the measure occurring in the resolution of σ in terms of the extreme means $(!_{\lambda})_{\lambda \in [0, \infty]}$. The map, $\sigma \rightarrow f$, is an affine order-isomorphism from the class of means onto H , [33, Theorem 3.2]. The mean σ is said to be selfadjoint, if it satisfies

$$(A \sigma B)^{-1} = A^{-1} \sigma B^{-1}$$

for invertible $A, B \in \mathbb{M}_n^+$, $n = 1, 2, \dots$. It readily follows that σ is selfadjoint, if and only if the representing function f satisfies

$$f(t)^{-1} = f(t^{-1}) \quad \forall t \in \mathbb{R}.$$

But these are exactly the functions, we characterized at the end of Section 4. We have consequently determined the class of selfadjoint means.

REFERENCES

1. Anderson, W.N., Jr.: Shorted operators. SIAM J. Appl. Math. 20, 520-525 (1971)
2. Anderson, W.N., Jr., Duffin, R.J.: Series and parallel addition of matrices. J. Math. Anal. Appl. 26, 576-594 (1969)
3. Anderson, W.N., Jr., Morley, D.E., Trapp, G.E.: Characterization of parallel sums. Proc. Natl. Acad. Sci. U.S.A. 76, 3599-3601 (1979)
4. Anderson, W.N., Jr., Schreiber, M.: The infimum of two projections. Acta Sci. Math. 33, 165-168 (1972)
5. Anderson, W.N., Jr., Trapp, G.E.: Shorted operators. II. SIAM J. Appl. Math. 28, 60-71 (1975)
6. Anderson, W.N., Jr., Trapp, G.E.: A class of monotone operator functions related to electrical network theory. Linear Algebra Appl. 15, 53-67 (1976)
7. Ando, T.: Topics on operator inequalities. Lecture notes (mimeographed), Hokkaido Univ., Sapporo 1978
8. Ando, T.: Concavity of certain maps on positive definite matrices and applications to Hadamard products. Linear Algebra Appl. 26, 203-241 (1979)
9. Araki, H.: Relative entropy of states of von Neumann algebras I, II. Publ. RIMS. Kyoto Univ. 11, 809-833 (1975/76), and 13, 173-192 (1977)
10. Bendat, J., Sherman, S.: Monotone and convex operator functions. Trans. Am. Math. Soc. 79, 58-71 (1955)
11. Choi, M.-D.: A Schwarz inequality for positive linear maps on C*-algebras. Ill. J. Math. 18, 565-574 (1974)

12. Davis, C.: A Schwarz inequality for convex operator functions. Proc. Amer. Math. Soc. 8, 42-44 (1957)
13. Davis, C.: Notions generalizing convexity for functions defined on spaces of matrices. Proc. Amer. Math. Soc. Symposia (Convexity) 7, 187-201 (1963)
14. Dobsch, O.: Matrixfunktionen beschränkter Schwankung. Math. Z. 43, 353-388 (1937)
15. Donoghue, W.: Monotone matrix functions and analytic continuation. Berlin, Heidelberg, New York: Springer 1974
16. Epstein, H.: Remarks on two theorems of E. Lieb. Comm. Math. Phys. 31, 317-325 (1973)
17. Fujii, J.I.: Arithmetico-geometric mean of operators. Math. Japon. 23, 667-669 (1979)
18. Fujii, J.I.: On geometric and harmonic means of operators. Math. Japon. 24, 203-207 (1979)
19. Fujii, J.I.: Initial conditions on operator monotone functions. Math. Japon. 24, 459-462 (1979)
20. Fujii, J.I., Fujii, M.: Some remarks on operator means. Math. Japon. 24, 335-339 (1979)
21. Halmos, P.R.: A Hilbert space problem book. American Book 1967
22. Hansen, F.: An operator inequality. Math. Ann. 246, 249-250 (1980)
23. Hansen, F.: Selfadjoint means and operator monotone functions. Math. Ann. 256, 29-35 (1981)
24. Hansen, F.: Means and concave products of positive semi-definite matrices. Math. Ann. 264, 119 - 128 (1983)
25. Hansen, F., Pedersen, G.K.: Jensen's inequality for operators and Löwner's theorem. Math. Ann. 258, 229-241 (1982)

26. Heinz, E.: Beiträge zur Störungstheorie der Spectralzerlegung. Math. Ann. 123, 415-438 (1951)
27. Jensen, J.L.W.V.: Om konvekse Funktioner og Uligheder mellem Middelværdier. Nyt Tidsskr. Math. B 16, 49-68 (1905)
28. Jensen, J.L.W.V.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math. 30, 175-193 (1906)
29. Kainuma, D., Nakamura, M.: Around Jensen's inequality. Math. Japon. 25, 585-588 (1980)
30. Korányi, A.: On a theorem of Löwner and its connections with resolvents of self-adjoint transformations. Acta Sci. Math. (Szeged) 17, 63-70 (1956)
31. Kosaki, H.: Interpolation theory and the Wigner-Yanase-Dyson-Lieb concavity. Comm. Math. Phys. 87, 315-329 (1982)
32. Krauss, F.: Über konvexe Matrixfunktionen. Math. Z. 41, 18-42 (1936)
33. Kubo, I., Ando, T.: Means of positive linear operators. Math. Ann. 246, 205-224 (1980)
34. Lieb, E.: Convex trace functions and the Wigner-Yanase-Dyson conjecture. Ad. in Math. 11, 267-288 (1973)
35. Lieb, E., Ruskai, M.B.: Some operator inequalities of the Schwarz type. Ad. in Math. 12, 269-273 (1974)
36. Loewner, C.: Some classes of functions defined by difference or differential inequalities. Bull. Amer. Math. Soc. 56, 308-319 (1950)
37. Löwner, K.: Über monotone Matrixfunktionen. Math. Z. 38, 177-216 (1934)
38. Nishio, K., Ando, T.: Characterizations of operations derived from network connections. J. Math. Anal. Appl. 53, 539-549 (1976)

39. Pedersen, G.K.: Some operator monotone functions. Proc. Amer. Math. Soc. 36, 309-310 (1972)
40. Pedersen, G.K.: C*-algebras and their automorphism groups. London Math. Soc. Monographs 14. London, New York: Academic Press 1979
41. Pusz, W., Woronowicz, S.L.: Functional calculus for sesquilinear forms and the purification map. Rep. Math. Phys. 8, 159-170 (1975)
42. Pusz, W., Woronowicz, S.L.: Form convex functions and the WYDL and other inequalities. Lett. Math. Phys. 2, 505-512 (1978)
43. Sparr, G.: A new proof of Löwner's theorem on monotone matrix functions. Math. Scand. 47, 266-274 (1980)
44. Stinespring, W.F.: Positive functions on C*-algebras. Proc. Amer. Math. Soc. 6, 211-216 (1955)
45. Sz-Nagy, B.: Remarks to the preceding paper of A. Korányi. Acta Sci. Math. (Szeged) 17, 71-75 (1956)
46. Uhlmann, A.: Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. Comm. Math. Phys. 54, 21-32 (1977)