MONOTONE AND CONVEX OPERATOR FUNCTIONS(1)

BY

JULIUS BENDAT AND SEYMOUR SHERMAN(2)

1. **Introduction.** Monotone matrix functions of arbitrarily high order were introduced by Charles Loewner in the year 1934 [9](3) while studying real-valued functions which are analytic in their domain of definition, and continued in the complex domain, are regular in the entire open upper half-plane with non-negative imaginary part. The monotone matrix functions of order n defined later in the introduction deal with functions of matrices whose independent and dependent variables are real symmetric matrices of order n; monotone of arbitrarily high order means monotone for each finite integer n. The class corresponding to $n=1$ represents the functions which are monotone in the ordinary sense. Monotone operator functions are precisely the class of monotone matrix functions of arbitrarily high order. Loewner showed that analyticity plus the property of mapping the complex open upper half-plane into itself is characteristic for the class of monotone matrix functions of arbitrarily high order. Literature on this subject and its by-products, convex matrix functions and matrix functions of bounded variation, consists of but four papers, two by Loewner [9; 10] and one each from two of his doctoral students, F. Krauss [8] and O. Dobsch [5].

The principal objective in this work was to discover whether convex operator functions (i.e. convex matrix functions of arbitrarily high order) satisfy properties analogous to those known for monotone operator functions. From the classical case it might be thought if a function is convex, its derivative should be monotone. Does this hold for operator functions? If not, what is true? It was deemed desirable to develop new representations for monotone operator functions.

On physical grounds the interest in monotone operator functions is quite broad. A strong resemblance exists to the mathematical aspect of part of E. P. Wigner's work in quantum-mechanical particle interactions [20], the physical side being discussed by him in [18; 19]. Connections with electrical network theory are mentioned by several writers [3; 4; 12]. The Hamburger moment problem [16; 17] and the theory of typically real functions of order one studied by W. Rogozinski [15] and M. S. Robertson [13; 14] are both allied subjects. Even a slight relation with the special theory of relativity

Presented to the Society, September 2, 1954; received by the editors January 21, 1954.

(1) This is material from the thesis of the first named author in partial fulfillment for the degree of Doctor of Philosophy at the University of Southern California, August, 1953. Now at Northrop Aircraft, Inc., Hawthorne, California.

(2) Now at the University of Pennsylvania.

(3) Numbers in brackets refer to the List of References.

was verbally communicated to the writers by Professor Loewner. The physical importance of convex operator functions is still unknown.

Definition of monotone operator functions and of monotone (matrix) functions of order n: Let $f(x)$ be a bounded real-valued function of a real variable x defined in an interval I (which may be open, half-open, or closed; finite or infinite). We consider the totality of bounded self-adjoint operators K in a Hilbert space H whose spectrum lies in the domain of $f(x)$. If A is in K, then by $f(A)$ we mean the self-adjoint operator which results from A such that $f(A)=\int_{-\infty}^{\infty} f(\lambda)dE_\lambda$ where (E_λ) is the left-continuous spectral resolution corresponding to A [cf. Nagy, 11, p. 60]. For $f(x)=\sum_0^K a_n x^n$, a polynomial, this definition implies, as is well known, $f(A)=\sum_0^K a_n A^n$. $f|K$ is called an operator function.

In the case of n-by-n real symmetric or hermitian matrices X_n, $f(X_n)$ gives the matrix resulting from X_n in which each eigenvector is fixed while the corresponding eigenvalue λ is replaced by $f(\lambda)$. Thus, if $X_n=T'DT$ where T is an orthogonal matrix, T' its transpose matrix, and D a diagonal matrix, then $f(X_n)=T'f(D)T$. The function $f(X_n)$ on all n-by-n matrices X_n with eigenvalues in the domain of $f(x)$ is called a (matrix) function of order n.

When A and B are two bounded self-adjoint operators in H for which the inner product

$$(Ax,\ x) \leqq (Bx,\ x) \text{ for all } x \text{ in } H,$$

then we say A is less than or equal to B, and write $A \leqq B$ or $B \geqq A$.

We now define: An operator function $f|K$ is called monotone if for each two operators A, B in K which stand in the relation $A \leqq B$ to one another, always $f(A) \leqq f(B)$ or $f(A) \geqq f(B)$. In the first case, $f(A)$ is monotone increasing; in the second, monotone decreasing. Throughout this work, the discussion will be always of increasing $f(x)$. When considering n-by-n real symmetric or hermitian matrices, a monotone operator function is called a monotone (matrix) function of order n. A brief summary of each section now follows:

§2, after stating known results about monotone matrix functions, carries out the extension to the case of monotone operator functions. Analyticity is established from results of S. Bernstein [1] and R. P. Boas [2]. A Stieltjes integral representation of monotone operator functions is obtained using the Hamburger moment problem(4). This immediately gives analyticity in the entire complex plane off the real axis with the known mapping property. The representation is shown to be unique even though the solution to the Hamburger moment problem is not in general unique(5). Similarities to the

(4) This result was derived independently by the writers only to learn later by letter Professor Loewner's awareness of it. While nothing of this nature has been published, mention is made by Loewner [15, p. 313] about using the theory of moments.

(5) The writers are indebted to R. S. Phillips for this suggestion.

class of typically real functions of order one [15] are noted.

The final section considers the question of convex operator functions which, in the subdomain of real symmetric matrices of order n, were previously handled by F. Krauss [8]. Convex operator functions are shown to be analytic in their domain of definition and a new criterion for this class is established, namely that a function $f(x) = \sum_1^\infty a_n x^n$ is a convex operator function in $|x| < R$, where R is the radius of convergence, if and only if $f(x)/x$ is a monotone operator function in $|x| < R$. Complex mapping properties of convex operator functions are now presented.

2. **Monotone operator functions.** The fundamental properties of monotone matrix functions of arbitrary order $n > 1$ are given in the following three theorems due to Loewner.

THEOREM 2.1 [Loewner, 9, p. 189]. *In order that a function $f(x)$ be monotonic of order n in the open interval (a, b), it is necessary and sufficient that the determinants*

$$\left| \frac{f(\mu_i) - f(\lambda_k)}{\mu_i - \lambda_k} \right|_{i,k=1}^{m} \geqq 0 \qquad (m = 1, 2, \cdots, n)$$

for arbitrary values μ_i, λ_k $(i, k = 1, 2, \cdots, m)$ in (a, b) provided only that $\lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \cdots < \lambda_m < \mu_m$.

THEOREM 2.2. *A monotone function $f(x)$ of order $n > 1$ is of class $2n-3$ (i.e. $2n-3$ times continuously differentiable) and its $(2n-3)$th derivative is a convex function. The derivatives $f^{(2n-2)}(x)$ and $f^{(2n-1)}(x)$ exist, therefore, almost everywhere.*

THEOREM 2.3. *A function $f(x)$ defined in (a, b) is monotone of arbitrarily high order if and only if it is analytic in (a, b), can analytically be continued into the whole upper half-plane, and represents there an analytic function whose imaginary part is non-negative.*

Simple examples are given by x^μ, $0 \leqq \mu \leqq 1$, and $\log x$ for $x > 0$,

$$\frac{ax + b}{cx + d} \text{ with } ad - bc > 0$$

either in $x > -d/c$ or $x < -d/c$. Examples for non-monotone functions are provided by e^x in arbitrary (a, b) or a simple step function in any interval about the discontinuity.

Dobsch [5, p. 368] transformed the determinant requirement of Theorem 2.1 into a quadratic form requirement obtaining

THEOREM 2.4. *A function $f(x)$ of class C_{2n-1} in (a, b) is monotonic of order n there if and only if the real quadratic form*

$$\sum_{i,k=0}^{n-1} \frac{f^{(i+k+1)}(x)}{(i+k+1)!} \zeta_i \zeta_k \geqq 0$$

for all x in (a, b).

This result will now be extended to monotone operator functions.

LEMMA 2.1. *A monotone operator function in (a, b) is monotone for all finite orders n in (a, b).*

Proof. Trivial.

LEMMA 2.2. *If $f(x)$ is monotone in (a, b) for all finite orders n, then $f(x)$ is a monotone operator function in (a, b).*

Proof. Without loss of generality, assume $f(x)$ defined in interval (a, b) containing the origin. Let A be a bounded self-adjoint operator in a Hilbert space H, with spectrum in (a, b), represented in a special coordinate system $(x_1, x_2, \cdots, x_n, \cdots)$ of H by the matrix $(a_{ik})_{i,k=1}^{\infty}$. The norm of A is denoted by $\|A\|$. Define A_n by the matrix

$$A_n = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Now the strong limit, $\lim_{n\to\infty} A_n = A$, i.e. $\|A_n x - Ax\| \to 0$ for each x in H. Also, for k = integer, $\lim_{n\to\infty} A_n^k = A^k$ in the strong topology since $\|A_n\| \leqq \|A\|$ for all n. Similarly, $\lim_{n\to\infty} (A_n^k + A_n^l) = A^k + A^l$; $\lim_{n\to\infty} cA_n^k = cA^k$ for c = constant. Therefore $\lim_{n\to\infty} p(A_n) = p(A)$ in the strong topology where $p(x) = \sum_{n=0}^{k} c_n x^n$, a polynomial.

Since $f(x)$ is a monotone matrix function for all finite orders n, $f(x)$ is continuous. So, given any $\epsilon > 0$, there exists a polynomial $p_\epsilon(x)$ such that $|p_\epsilon(x) - f(x)| < \epsilon$, uniformly for all x in the closed convex hull of the spectrum of operator A contained in (a, b).

Hence $\lim_{\epsilon\to 0} p_\epsilon(A) = f(A)$ in the uniform operator topology.

CLAIM. $\lim_{n\to\infty} f(A_n) = f(A)$ for fixed $x = (x_1, x_2, \cdots) \in H$.

Proof of claim. Given any $\epsilon > 0$, choose polynomial $p_\epsilon(x)$ such that $\|[f(A) - p_\epsilon(A)]x\| \leqq \epsilon/3$. It follows that $\|[f(A_n) - p_\epsilon(A_n)]x\| \leqq \epsilon/3$ for all n since the end points for the spectrum of A_n are contained within the spectrum of A. Then choose N so that $\|[p_\epsilon(A) - p_\epsilon(A_n)]x\| \leqq \epsilon/3$ for all $n \geqq N$. Now

$$\begin{aligned} \|[f(A) - f(A_n)]x\| &\leqq \|[f(A) - p_\epsilon(A)]x\| + \|[p_\epsilon(A) - p_\epsilon(A_n)]x\| \\ &\quad + \|[p_\epsilon(A_n) - f(A_n)]x\| \leqq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \end{aligned}$$

for $n \geqq N$. Hence $\lim_{n\to\infty} \|[f(A) - f(A_n)]x\| = 0$, proving the claim.

To return to the lemma: Let A and B be two bounded self-adjoint oper-

ators with spectrum in (a, b) such that $A \leqq B$. Then also $A_n \leqq B_n$ and, therefore, $f(A_n) \leqq f(B_n)$ since monotonicity of the nth order is assumed. Now for fixed x in H, letting $n \to \infty$, we obtain $f(A) \leftarrow f(A_n) \leqq f(B_n) \to f(B)$ or $f(A) \leqq f(B)$ since $f(x)$ converges uniformly in the closed convex hull of the combined spectra of A and B contained in (a, b). This proves $f(x)$ is a monotone operator function.

The writers are indebted to Professor Loewner for suggesting this approach of proving Lemma 2.2, the proof of which has never been published.

Combining Lemmas 2.1 and 2.2 gives

THEOREM 2.5. *$f(x)$ is a monotone operator function in (a, b) if and only if $f(x)$ is a monotone matrix function in (a, b) for all finite orders n.*

We now apply the content of Theorems 2.2 and 2.5 to Theorem 2.4 to obtain

THEOREM 2.6. *A function $f(x)$ defined in (a, b) is a monotone operator function there if and only if for all $N=0, 1, 2, \cdots$ the real quadratic form*

$$\sum_{i,k=0}^{N} \frac{f^{(i+k+1)}(x)}{(i+k+1)!} \zeta_i \zeta_k \geqq 0$$

for all x in (a, b).

By means of this quadratic form requirement, the following three lemmas are easily established. Proofs are left to the reader.

LEMMA 2.3. *$f'(x_0)=0$ implies $f^{(n)}(x_0)=0$ for $n=2, 3, \cdots$.*

LEMMA 2.4. *$f^{(2p+1)}(x) \geqq 0$ $(p=0, 1, 2, \cdots)$ for all x in $a<x<b$.*

LEMMA 2.5.

$$\frac{f^{(2n+1)}(x) f^{(2m-1)}(x)}{(2n+1)!(2m-1)!} \geqq \left[\frac{f^{(n+m)}(x)}{(n+m)!}\right]^2 \qquad (m, n = 1, 2, \cdots)$$

for all x in (a, b).

THEOREM 2.7. *A monotone operator function defined in $a<x<b$ is analytic.*

Proof. By Theorem 2.2, $f(x) \in C^\infty$ in (a, b). By Lemma 2.4, $f^{(2p+1)}(x) \geqq 0$ for all x in (a, b) $(p=0, 1, 2, \cdots)$, in words, all odd derivatives of $f(x)$ are of constant sign.

That these two conditions are sufficient to guarantee analyticity (i.e. expansion of $f(x)$ in a power series about each point with a nonzero radius of convergence) now follows from a theorem in analysis due to S. Bernstein [1] and R. P. Boas [2].

The statement of this Theorem 2.7 was first given by Loewner [9, p. 215] but the proof given here is new.

We shall now and from henceforth, unless otherwise stated, without loss of generality, assume $f(x)$ to be a monotone operator function in an interval containing the origin. Also, we shall take $f(0)=0$ for if not, say $f(0)=c$, then $f(x)$ monotone implies $f(x)-c$ monotone and we would work with the latter function. By the previous theorem, $f(x)$ monotone means $f(x)$ analytic; therefore we can express $f(x)$ in a power series $f(x)=\sum_1^\infty a_n x^n$ with a nonzero radius of convergence. All coefficients a_n are real. The first coefficient $a_1>0$, for if not, by Lemmas 3 and 4, all $a_n=0$ giving $f(x)\equiv 0$.

We now seek properties in the large.

THEOREM 2.8. *A monotone operator function $f(x)=\sum_1^\infty a_n x^n$, with radius of convergence R, may be expressed as a Stieltjes integral*

$$f(x) = \int_{-1/R}^{1/R} \frac{x}{1-tx}\, d\psi(t)$$

where $\psi(t)$ is a bounded nondecreasing function, constant for $|t|>1/R$. If the radius of convergence R is infinite, then $f(x)=a_1x$ with $a_1>0$.

Proof. We have $f(x)=\sum_1^\infty a_n x^n$, convergent in $|x|<R$, and

$$\sum_{\lambda,\mu=0}^{N} a_{\lambda+\mu+1}\zeta_\lambda\zeta_\mu \geqq 0$$

where $a_n=(f^{(n)}(0))/n!$, $N=0, 1, 2, \cdots$.

From the Hamburger moment problem [Widder, 31, p. 129], the quadratic condition on the a_n's insures the existence of a bounded nondecreasing function $\psi(t)$ such that

$$a_{n+1} = \int_{-\infty}^{\infty} t^n d\psi(t) \qquad (n = 0, 1, 2, \cdots).$$

This is seen immediately by defining $b_n=a_{n+1}$ $(n=0, 1, 2, \cdots)$. Then $\sum_{\lambda,\mu=0}^N a_{\lambda+\mu+1}\zeta_\lambda\zeta_\mu\geqq 0$ is equivalent with $\sum_{\lambda,\mu=0}^N b_{\lambda+\mu}\zeta_\lambda\zeta_\mu\geqq 0$ and the latter condition on the b_n's gives the above integral. In particular $a_1=\int_{-\infty}^{\infty} d\psi(t)>0$. If desired, the function might be normalized so that $a_1=1$.

But $f(x)=\sum_1^\infty a_n x^n$, convergent in $|x|<R$, means that

$$\limsup_{n\to\infty} |a_n|^{1/n} = 1/R.$$

Hence $\psi(t)$ must be constant for $|t|>1/R$. For if not, say $\psi(t)$ has a positive increase in $t_1\leqq t\leqq t_2$ where $t_1>1/R$. Then

$$a_{2n+1} = \int_{-\infty}^{\infty} t^{2n} d\psi(t) \geqq t_1^{2n}[\psi(t_2)-\psi(t_1)].$$

Now $\limsup_{n\to\infty} |a_n|^{1/n}\geqq t_1>1/R$, giving a contradiction. Hence

$$a_{n+1} = \int_{-1/R}^{1/R} t^n d\psi(t) \qquad (n = 0, 1, 2, \cdots).$$

Therefore $f(x) = \sum_1^\infty a_n x^n = \int_{-1/R}^{1/R} (x/(1-tx)) d\psi(t)$, proving the first part of the theorem.

Letting $R \to \infty$ gives $a_n = 0$ for all $n \neq 1$ since we originally assumed $a_1 > 0$. Thus in this case the series reduces merely to $f(x) = a_1 x$ with $a_1 > 0$.

The corollary below gives a new proof of part of Loewner's Theorem 2.3.

COROLLARY. *A monotone operator function is analytic everywhere off the real axis and maps the open upper half-plane into itself.*

Proof. From the theorem just concluded, for complex $z \neq x$, if R is finite, $f(z) = \int_{-1/R}^{1/R} (z/(1-tz)) d\psi(t)$ where the integrand is continuous in both variables z and t. Hence $f(z)$ is analytic for all z off the real axis. Also the sign of $\operatorname{Im} f(z) = \text{sign of } \operatorname{Im} z/(1-tz) = \text{sign of } \operatorname{Im} [z(1-t\bar{z})] = \text{sign of } z$. Therefore $\operatorname{Im} f(z) \gtrless 0$ if $\operatorname{Im} z \gtrless 0$. If R is infinite, $f(z) = a_1 z$ with $a_1 > 0$ for which the result is immediate.

THEOREM 2.9. *If $\psi(t)$ is a bounded nondecreasing function, constant for $|t| > 1/R$, then $f(x) = \int_{-1/R}^{1/R} (x/(1-tx)) d\psi(t)$ is a monotone operator function in $|x| < R$.*

Proof. The stated result follows from the knowledge that the set of functions $[x/(1-tx) \mid |t| \leqq 1/R]$ are monotone operator functions for $|x| < R$. For then, the

$$\lim_{\delta \to 0} \sum_{i=0}^{n-1} \frac{x}{1 - \zeta_i x} [\psi(t_{i+1}) - \psi(t_i)] = \int_{-1/R}^{1/R} \frac{x}{1 - tx} d\psi(t) = f(x)$$

exists, for each x in $|x| < R$, independent of the manner of subdivision $-1/R = t_0 < t_1 < \cdots < t_n = 1/R$ and of the choice of numbers ζ_i in the subinterval $[t_i, t_{i+1}]$ $(i = 0, 1, \cdots, n-1)$, where $\delta = \sup_i (t_{i+1} - t_i)$, since $x/(1-tx)$ is continuous and $\psi(t)$ is of bounded variation in $[-1/R, 1/R]$. Let $f_\delta(x) = \sum_{i=0}^{n-1} (x/(1-\zeta_i x)) [\psi(t_{i+1}) - \psi(t_i)]$. Then clearly $f_\delta(x)$ as the finite sum of positive multiples of monotone operator functions is one itself. Our function $f(x)$ is the limit of $f_\delta(x)$ in the above described Moore-Smith sense as δ goes to zero, the convergence being uniform in any closed subinterval of $|x| < R$. Take any two operators $A \leqq B$ in the domain of $f(x)$. Then for δ sufficiently small, A and B lie in the domain of $f_\delta(x)$. Hence $f_\delta(A) \leqq f_\delta(B)$ since $f_\delta(x)$ is a monotone operator function. $f_\delta(x)$ converges uniformly to $f(x)$ as δ goes to zero in the closed convex hull of the combined spectra of A and B contained in $|x| < R$. Consequently, the limiting process gives $f(A) \leqq f(B)$ which proves the theorem.

The fact that the set of functions $[(x/(1-tx)) \mid |t| \leqq 1/R]$ are monotone operator functions for $|x| < R$ is known either from Theorem 2.3 or from

Theorem 2.6. We add a direct operator proof of

LEMMA 2.6. *$x/(1-tx)$ for $|t| \leqq 1/R$ is a monotone operator function in $|x| < R$.*

Proof.

CASE 1: $0 < t \leqq 1/R$.

$$\begin{aligned}
x_1 < x_2 &\rightarrow tx_1 < tx_2 \\
&\rightarrow 1 - tx_1 > 1 - tx_2 > 0 \\
&\rightarrow \frac{1}{1 - tx_1} < \frac{1}{1 - tx_2} \\
&\rightarrow \frac{1}{t}\left(\frac{1}{1 - tx_1}\right) < \frac{1}{t}\left(\frac{1}{1 - tx_2}\right) \\
&\rightarrow \frac{-1}{t} + \frac{1}{t(1 - tx_1)} < \frac{-1}{t} + \frac{1}{t(1 - tx_2)} \\
&\rightarrow \frac{-1 + tx_1 + 1}{t(1 - tx_1)} < \frac{-1 + tx_2 + 1}{t(1 - tx_2)} \\
&\rightarrow \frac{x_1}{1 - tx_1} < \frac{x_2}{1 - tx_2}.
\end{aligned}$$

CASE 2: $0 > t > -1/R$.

$$\begin{aligned}
x_1 < x_2 &\rightarrow tx_1 > tx_2 \rightarrow 0 < 1 - tx_1 < 1 - tx_2 \\
&\rightarrow \frac{1}{1 - tx_1} > \frac{1}{1 - tx_2} \rightarrow \frac{1}{t(1 - tx_1)} < \frac{1}{t(1 - tx_2)} \\
&\rightarrow \frac{-1}{t} + \frac{1}{t(1 - tx_1)} < \frac{-1}{t} + \frac{1}{t(1 - tx_2)} \\
&\rightarrow \frac{x_1}{1 - tx_1} < \frac{x_2}{1 - tx_2}.
\end{aligned}$$

CASE 3: For $t = 0$, result is trivial.

REMARK. The class $[f(x) = x/(1-tx) \mid t \text{ arbitrary}]$ is identical with the class of proper projective transformations $[f(x) = (ax+b)/(cx+d) \mid ad - bc > 0$ where $f(0) = 0, f'(0) = 1]$.

THEOREM 2.10. *If $f(x) = \sum_1^\infty a_n x^n$ has radius of convergence R, and if the real quadratic form $\sum_{\lambda,\mu=0}^N (f^{(\lambda+\mu+1)}(0)/(\lambda+\mu+1)!)\zeta_\lambda\zeta_\mu \geqq 0$ for all $N = 0, 1, 2, \cdots$, then $f(x)$ is a monotone operator function in $|x| < R$.*

Proof. We note the hypothesis requires information only at one point $x = 0$ instead of for all x in $|x| < R$ as stated in Theorem 2.6. As carried out in

Theorem 2.8, $f(x)=\int_{-1/R}^{1/R}(x/(1-tx))d\psi(t)$ where $\psi(t)$ is a bounded nondecreasing function, constant for $|t|>1/R$. We have already shown in Theorem 2.9 that $f(x)$ so defined is a monotone operator function in $|x|<R$.

THEOREM 2.11. *Any monotone operator function* $f(x)=\sum_1^\infty a_n x^n$ *with radius of convergence* R $(0<R<\infty)$ *is represented uniquely as a Stieltjes integral*

$$f(x) = \int_{-1/R}^{1/R} \frac{x}{1-tx}\, d\psi(t), \qquad |x| < R,$$

if $\psi(t)$ *is a bounded nondecreasing function, continuous on the left, with* $\psi(-1/R)=0$.

Proof. By Theorem 2.8, plus the usual equivalence of Stieltjes integrals by restricting $\psi(t)$ as above, we know that such a representation exists. We want to show uniqueness. Assume two such formulas exist. Then, for all $|x|<R$,

$$\int_{-1/R}^{1/R} \frac{x}{1-tx}\, d\psi_1(t) = \int_{-1/R}^{1/R} \frac{x}{1-tx}\, d\psi_2(t)$$

where both $\psi_1(t)$ and $\psi_2(t)$ are bounded nondecreasing functions, continuous on the left, with $\psi_1(-1/R)=\psi_2(-1/R)=0$. We shall show that $\psi_1(t)=\psi_2(t)$ in $|t|\leqq 1/R$. First of all, $x/(1-tx)=\sum_0^\infty x^{n+1}t^n$ converges uniformly in $|t|\leqq 1/R$ for fixed $|x|<R$. Hence

$$\sum_0^\infty x^{n+1}\int_{-1/R}^{1/R} t^n d\psi_1(t) = \sum_0^\infty x^{n+1}\int_{-1/R}^{1/R} t^n d\psi_2(t)$$

for all $|x|<R$. The identity theorem for power series now gives

$$\int_{-1/R}^{1/R} t^n d\psi_1(t) = \int_{-1/R}^{1/R} t^n d\psi_2(t) \qquad (n = 0, 1, 2, \cdots).$$

By [17, p. 60, Theorem 6.1], $\psi_1(t)=\psi_2(t)$ for all t such that $|t|\leqq 1/R$.

We remark next on a similarity to the class of typically real functions of order one [13; 14; 15]. A function $f(z)=\sum_1^\infty a_n z^n$ is called typically real of order one if all coefficients a_n are real, $f(z)$ is analytic in $|z|<1$, and $\operatorname{Im} f(z)$ has the same sign as $\operatorname{Im} z$ when z is not real and $|z|<1$. This is exactly what we have shown to hold for monotone operator functions defined in $|x|<1$. It follows that the class of monotone operator functions in $|x|<1$ is contained in the class of typically real functions of order one. One result worth noting from the latter class is that if $f(z)=\sum_1^\infty a_n z^n$ is typically real of order one, then $|a_n|\leqq n|a_1|$ for all n [14, p. 471]. For monotone operator functions, we can make a stronger statement.

LEMMA 2.7. *If* $f(x)=\sum_1^\infty a_n x^n$, *convergent in* $|x|<R$ *where* R *is the radius*

of convergence, is a monotone operator function, then $|a_n| \leqq a_1/R^{n-1}$ *for all* n.

Proof. As developed in Theorem 2.8,

$$a_n = \int_{-1/R}^{1/R} t^{n-1} d\psi(t) \qquad (n = 1, 2, \cdots).$$

Now $|a_n| \leqq \int_{-1/R}^{1/R} |t^{n-1}| d\psi(t) \leqq (1/R^{n-1}) \int_{-1/R}^{1/R} d\psi(t) = a_1/R^{n-1}$. We note, in particular, if $R \geqq 1$, then $|a_n| \leqq a_1$ for all n.

Finally we mention some work of E. Heinz [7, p. 425] in perturbation theory. Heinz uses a function-theoretic result of Herglotz to show that if $f(z)$ is analytic in the z-plane cut along the negative real axis with $\operatorname{Im} f(z) \geqq 0$ for $\operatorname{Im} z > 0$, and if $f(x)$ is continuous and $\geqq 0$ for $0 < x < \infty$, then

$$f(x) = \int_0^\infty \frac{tx + 1}{t - x} d\phi(t) + \alpha x + \beta$$

is a monotone operator function [i.e. $0 \leqq A \leqq B$ implies $0 \leqq f(A) \leqq f(B)$ for self-adjoint, not necessarily bounded, operators A, B]. Here, α and β are two real numbers with $\alpha \geqq 0$ and $\phi(t)$ is a bounded nondecreasing function.

3. **Convex operator functions.** *Definition of a convex operator function.* Let $f(x)$ be a bounded real-valued function of a real variable x defined in an interval I (which may be open, half-open, or closed; finite or infinite). We consider the totality of bounded self-adjoint operators K on a Hilbert space H whose spectrum falls in the domain of $f(x)$. $f|K$ is defined as in §1.

Let A and B be any two bounded self-adjoint operators in H. Consider the connected "segment" of these two operators, i.e. the one-parameter family of operators $C = (1-t)A + tB$ $(0 \leqq t \leqq 1)$. Then the spectrum of C lies in the interval bounded by the minimum and maximum values for the spectrum of A or B.

We now define: An operator function $f|K$ is called convex if for every pair of operators A, B in K,

$$f[(1 - t)A + tB] \leqq (1 - t)f(A) + tf(B), \qquad 0 \leqq t \leqq 1.$$

Geometrically, the value of the function at an interior point lies below or on the chord joining the end points.

When considering n-by-n symmetric matrices, a convex operator function is called a convex (matrix) function of order n.

F. Krauss in [8, p. 41] proved the following theorem:

THEOREM 3.1. *A necessary and sufficient condition that a function* $f(x)$, *defined in an open interval* $a < x < b$, *be convex of order* $n > 1$ *is that it be twice continuously differentiable and all determinants*

$$D_p = \left| [x_0 x_i x_k] \right|_{i,k=1}^{p} \geqq 0 \qquad (p = 1, 2, \cdots, n).$$

$[x_0x_ix_k]$ equals the second divided difference of $f(x)$ at the points x_0, x_i, and x_k. For $p=1, 2, \cdots, n-1$, arbitrary choice of the $p+1$ points $x_0, x_1, \cdots, x_p$ is permitted while $p=n$ requires that x_0 coincide with one of the following x_i values.

Before continuing further, a short discussion of divided differences is advisable. By definition, the first divided difference of $f(x)$ at unequal points x_1, x_2 is

$$[x_1x_2] = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

If $x_1=x_2$ then $[x_1, x_2]=f'(x_1)$. The second divided difference, if $x_1 \neq x_3$, is

$$[x_1x_2x_3] = \frac{[x_2x_3] - [x_1x_2]}{x_3 - x_1}$$

and

$$[x_1, x_2, x_1] = \lim_{x_3 \to x_1} [x_1, x_2, x_3].$$

The following new theorem is basic in the study of convex operator functions.

THEOREM 3.2. *A function $f(x)$ is a convex matrix function of order n in (a, b) (for all positive integers n) if and only if*

$$F_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

is a monotone matrix function of order n in (a, b) (for all positive integers n) for each fixed x_0 in (a, b).

Proof. From Theorem 3.1, we know that convexity of $f(x)$ is characterized by the determinants

$$D_p = \left| [x_0x_ix_k] \right|_{i,k=1}^{p} \geqq 0, \qquad p = 1, 2, \cdots,$$

where x_i, x_k are arbitrary arguments in the domain of $f(x)$. Now,

$$[x_0x_ix_k] = [x_ix_0x_k] = \frac{[x_0x_k] - [x_0x_i]}{x_k - x_i} = \frac{F_{x_0}(x_k) - F_{x_0}(x_i)}{x_k - x_i}$$

where $F_{x_0}(x) = [x_0x] = (f(x) - f(x_0))/(x - x_0)$. Therefore,

$$D_p = \left| \frac{F_{x_0}(x_k) - F_{x_0}(x_i)}{x_k - x_i} \right|_{i,k=1}^{p} \geqq 0, \qquad p = 1, 2, \cdots,$$

which by Theorem 2.1 and its extension [9, p. 189] is both a necessary and sufficient condition that $F_{x_0}(x)$ is a monotone matrix function of order n in (a, b) (for all positive integers n) for each fixed x_0 in (a, b). This completes the proof.

In order to apply Theorem 3.2 to operator functions, we need

LEMMA 3.1. *A function is a convex operator function in (a, b) if and only if it is a convex matrix function in (a, b) for all finite orders n.*

Proof. Follows similar lines as in Lemmas 2.1 and 2.2.

It can now be asserted,

THEOREM 3.3. *A necessary and sufficient condition that $f(x)$ be a convex operator function in (a, b) is that*

$$F_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

be a monotone operator function in (a, b) for each fixed x_0 in (a, b).

Several corollaries are immediate consequences of the above theorem and previously developed properties of monotone operator functions.

COROLLARY 1. *A convex operator function is necessarily analytic in its domain of definition and continued in the complex domain, is analytic for all points off the real axis with*

$$\operatorname{Im}\left(\frac{f(z) - f(x_0)}{z - x_0}\right) \geqq 0$$

when $\operatorname{Im} z \geqq 0$, *independent of the choice of x_0 in its domain.*

Proof. Theorems 3.3 and 2.3.

COROLLARY 2. *If $f(x)$ is a convex operator function in (a, b), then about each value x_0 in (a, b) there exists a power series expansion,*

$$f(x) = \sum_0^\infty a_n(x - x_0)^n$$

with a nonzero radius of convergence R. If R is finite, $f(x)$ can be represented uniquely as a Stieltjes integral,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{-1/R}^{1/R} \frac{(x - x_0)^2}{1 - t(x - x_0)}\, d\psi(t)$$

where $\psi(t)$ is a bounded nondecreasing function, continuous on the left, with $\psi(-1/R) = 0$. For infinite R, $f(x)$ is necessarily the linear function

$$f(x) = f(x_0) + f'(x_0)(x - x_0).$$

Proof. First part follows from Corollary 1 and the integral representation by application to Theorem 2.11.

If $x_0=0$ falls in the domain of $f(x)$, and if we assume without loss of generality that both $f(0)=0, f'(0)=0$, the preceding corollary informs us that $f(x)$ is uniquely representable as

$$f(x) = \int_{-1/R}^{1/R} \frac{x^2}{1-tx}\, d\psi(t)$$

where R is the radius of convergence if $f(x)=\sum_0^\infty a_n x^n$, $\psi(t)$ as defined. We assume R finite in order to have $f(x)\not\equiv 0$.

Conversely, if we merely require $\psi(t)$ to be a bounded nondecreasing function, constant for $|t|>1/R$, then

COROLLARY 3. *$f(x)=\int_{-1/R}^{1/R}(x^2/(1-tx))d\psi(t)$ is a convex operator function in $|x|<R$.*

Proof. This follows readily, as in Theorem 2.9, provided we know that the functions $[x^2/(1-tx)\,|\,|t|\leqq 1/R]$ are convex operator functions in $|x|<R$. A direct use of Theorem 3.2 establishes this fact.

We can now write the following criterion for convex operator functions.

THEOREM 3.4. *A necessary and sufficient condition that $f(x)=\sum_2^\infty a_n x^n$ be a convex operator function in $|x|<R$, where R is the radius of convergence, is*

(1) *$f(x)/x$ be a monotone operator function in $|x|<R$, or*

(2) $\operatorname{Im}[f(z)/z]\geqq 0$ *when* $\operatorname{Im} z\geqq 0$.

By means of this theorem, it is verified that e^x or x^{2n} $(n=2, 3, \cdots)$ are not convex operator functions in any interval. Also, the derivative of the convex operator function $f(x)=x^2/(1-tx)$, namely $f'(x)=(2x-tx^2)/(1-tx)^2$, is *not* a monotone operator function in any interval since $f'(z)$ does not map the upper complex half-plane into itself. Here is a concrete example that the classical situation is false for operator functions.

An examination of quadratic form conditions developed in proofs of Theorems 2.8 and 2.10, together with Corollary 2 of Theorem 3.3, reveals

THEOREM 3.5. *A necessary condition that $f(x)$ be a convex operator function in $|x|<R$, where R is the radius of convergence of its power series expansion about the origin, is that the real quadratic form*

$$\sum_{\lambda,\mu=0}^{N} \frac{f^{(\lambda+\mu+2)}(x)}{(\lambda+\mu+2)!}\zeta_\lambda\zeta_\mu \geqq 0 \qquad \textit{for all } x \textit{ in } |x| < R\ (N=0, 1, 2, \cdots).$$

A sufficient condition is that at the single point $x=0$

$$\sum_{\lambda,\mu=0}^{N} \frac{f^{(\lambda+\mu+2)}(0)}{(\lambda+\mu+2)!}\zeta_\lambda\zeta_\mu \geqq 0 \qquad (N=0, 1, 2, \cdots).$$

Appendix. A second paper now in preparation will discuss certain geometric features of monotone and convex operator functions using the theory of polar cones. Application of all results obtained to some recent work of E. P. Wigner in quantum-mechanical particle interactions has also been carried out and will be the subject of a separate article.

Bibliography

1. S. Bernstein, *Sur la définition et les propriétés des fonctions analytiques d'une variable réele*, Math. Ann. vol. 75 (1914) pp. 449–468.

2. R. P. Boas, Jr., *Functions with positive derivatives*, Duke Math. J. vol. 8 (1941) pp. 163–172.

3. O. Brune, *Synthesis of a finite two-terminal network whose driving point impedance is a prescribed function of frequency*, Journal of Mathematics and Physics vol. 10 (1931) pp. 191–236.

4. W. Cauer, *Untersuchungen über ein Problem das drei positiv definite quadratische Formen mit Streckenkomplexen in Beziehung setzt*, Math. Ann. vol. 105 (1931) pp. 86–131.

5. O. Dobsch, *Matrixfunktionen beschränkter Schwankung*, Math. Zeit. vol. 43 (1937) pp. 353–388.

6. P. R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity*, New York, Chelsea, 1951.

7. E. Heinz, *Beiträge zur Störungstheorie der Spectralzerlegung*, Math. Ann. vol. 123 (1951) pp. 415–438.

8. F. Krauss, *Über konvexe Matrixfunktionen*, Math. Zeit. vol. 41 (1936) pp. 18–42.

9. K. Löwner, *Über monotone Matrixfunktionen*, Math. Zeit. vol. 38 (1934) pp. 177–216.

10. C. Loewner, *Some classes of functions defined by difference or differential inequalities*, Bull. Amer. Math. Soc. vol. 56 (1950) pp. 308–319.

11. B. Nagy, *Spectraldarstellung linearen Transformationen des Hilbertschen Raumes*, Berlin, Springer, 1942.

12. P. I. Richards, *A special class of functions with positive real part in a half-plane*, Duke Math. J. vol. 14 (1947) pp. 777–786.

13. M. S. Robertson, *On the coefficients of a typically real function*, Bull. Amer. Math. Soc. vol. 41 (1935) pp. 565–572.

14. ———, *Analytic functions star-like in one direction*, Amer. J. Math. vol. 58 (1936) pp. 465–472.

15. W. Rogozinski, *Über positive harmonische Entwicklungen und typisch-reele Potenzreihen*, Math. Zeit. vol. 35 (1932) pp. 93–121.

16. J. A. Shohat and J. D. Tamarkin, *The problem of moments*, Mathematical Surveys, No. 1, New York, American Mathematical Society, 1943.

17. D. V. Widder, *The Laplace transform*, Princeton University Press, 1941.

18. E. P. Wigner and L. Eisenbud, *Higher angular momenta and long range interaction in resonance reactions*, Physical Review vol. 72 (1947) pp. 29–41.

19. E. P. Wigner, *Nuclear reactions and level widths*, American Journal of Physics vol. 17 (1949) pp. 99–109.

20. ———, *On a class of analytic functions from the quantum theory of collisions*, Ann. of Math. vol. 53 (1951) pp. 36–67.

University of Southern California,
Los Angeles, Calif.