TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 176, February 1973

ON MONOTONE MATRIX FUNCTIONS OF TWO VARIABLES(1)

BY

HARKRISHAN VASUDEVA

ABSTRACT. The theory of monotone matrix functions has been developed by K. Loewner; he first gives some necessary and sufficient conditions for a function to be a monotone matrix function of order n, and then, as a result of further deep investigations including questions of interpolation he arrives at the following criterion: A real-valued function $f(x)$ defined in (a, b) is monotone of arbitrary high order n if and only if it is analytic in (a, b), can be analytically continued onto the entire upper half-plane, and has there a nonnegative imaginary part. The problem of monotone operator functions of two real variables has recently been considered by A. Koranyi. He has generalized Loewner's theorem on monotone matrix functions of arbitrary high order n to two variables. We seek a theory of monotone matrix functions of two variables analogous to that developed by Loewner and show that a complete analogue to Loewner's theory exists in two dimensions.

1. **Introduction.** The theory of monotone matrix functions was created by Charles Loewner in a celebrated paper published in 1934 [5]. This theory concerns functions of operators on a finite dimensional Hilbert space; the dimension of the space plays an important role in the theory, and the functions in question become more and more special as the dimension increases. A fundamental result of the theory is generally known as Loewner's theorem, and describes those functions which are Pick functions; they are analytic in the plane slit along certain parts of the real axis, and are real and analytic on an interval of that axis, having positive imaginary part in the upper half-plane.

Not long after the publication of Loewner's work it became clear that considerable interest should be attached to the class of monotone operator functions, namely, the corresponding class of functions associated with an infinite dimensional Hilbert space. It was virtually obvious that a monotone operator function should be a monotone matrix function of arbitrarily high order, and therefore, in view of Loewner's theorem, a function in the Pick class. Moreover, it was not difficult to

Received by the editors February 4, 1972.

AMS (MOS) subject classifications (1970). Primary 26A45.

(1) This paper contains the results of the doctoral thesis of the author. This research was in part supported by National Science Foundation Grants GP5436–1328 and the Air Force Office of Scientific Research Grants 1321/67.

Copyright © 1973, American Mathematical Society

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

show that any function in that class was a monotone operator function. Accordingly, the theory of monotone operator functions suffered only one shortcoming: it depended in an essential way on Loewner's theorem, and the proof of this theorem involved a long and difficult study of the Cauchy Interpolation Problem. It therefore seemed likely that the study of monotone operator functions should be essentially simpler than the study of monotone matrix functions, the infinite dimensional case being essentially easier because of the stronger hypotheses. Accordingly, Bendat and Sherman [1] addressed themselves to the problem of finding a direct, self-contained analysis of the monotone operator functions depending on the theory of the Hamburger moment problem. Their success attracted the interest of Sz.-Nagy and Koranyi who further simplified the theory of the operator functions in a sequence of papers ([3], [7]).

Throughout this paper we use the term 'matrix function' and 'operator function' to emphasize that in the one case the dimension of the associated Hilbert space is finite, while in the other it is infinite. While the theory of the monotone operator functions may be regarded as the most important extension of Loewner's work, that work has also been generalized in two different ways.

The first generalization concerns the study of the class of matrix functions of 'bounded variation' corresponding to the monotone functions introduced above. A second paper now in preparation will discuss certain features of matrix functions of bounded variation.

The second extension and generalization of Loewner's work is due to Koranyi [4]. That author considers the class of monotone functions of two variables, introduced by W. H. Young and considered at length in [2]. This leads immediately to a definition of monotone matrix functions of two variables associated with the tensor product of two finite dimensional spaces. Koranyi establishes a complete analogue to Loewner's theorem for functions of two variables. Thus he shows that the functions, monotone of arbitrarily high order, are analytic functions of two variables belonging to a class quite analogous to the Pick functions considered by Loewner.

The essential results of this paper are in the direction of Koranyi's generalization of Loewner's work; that author is concerned throughout with monotone operator functions, whereas we find a theory of monotone matrix functions of two variables analogous to that developed by Loewner.

It should be emphasized that the arguments and results would hold equally well for more than two variables; the statements of the theorems and the formulas would become unnecessarily complicated if we were to insist on stating all our results in terms of n variables.

2. Monotone functions of two variables.

Definition. A real function $f(x, y)$ of two variables defined in an open subset

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

of the plane will be called monotonic there if and only if

(i) it is locally integrable, and

(ii) for every set of points (x, y), $(x + h, y)$, $(x, y + k)$ and $(x + h, y + k)$ in the domain of f where h and k are positive, the quantity

$$f(x + h, y + k) - f(x + h, y) - f(x, y + k) + f(x, y)$$

is nonnegative.

A monotone function determines a distribution whose mixed partial derivative is the distribution limit of the quotients

$$[f(x + t, y + t) - f(x + t, y) - f(x, y + t) + f(x, y)]/t^2$$

and, since these are positive, the distribution $\partial^2 f/\partial x \partial y$ is a positive measure.

Suppose, next, that μ is a positive Radon measure with compact support in the plane, and that $H(x, y)$ is the characteristic function of the positive quadrant $x > 0$, $y > 0$. The convolution

$$f(x, y) = (H * \mu)(x, y) = \iint H(x - t, y - s)\, d\mu(t, s)$$

is simply the μ-measure of $t < x$, $s < y$. This is a nonnegative, bounded function, and in view of Fatou's theorem, this is even a lower semicontinuous function. Since the mixed partial of $H(x, y)$ is the delta distribution, we have $\partial^2 f/\partial x \partial y = \mu$. Moreover, it is pretty easy to see that $f(x, y)$ is monotone, since

$$f(x + h, y + k) - f(x + h) - f(x, y + k) + f(x, y)$$

is exactly the μ-measure of the rectangle determined by the inequalities $x \leq t < x + h$, $y \leq S < y + k$.

More generally, then, if μ is a positive Radon measure, we can write $\mu = \Sigma \mu_i$ where the family μ_i is determined by a suitable partition of unity. We then obtain $f = \Sigma f_i$ in a natural way: f is then a monotone nonnegative lower semicontinuous function whose mixed second partial derivative is the measure μ.

It is well known that any distribution T, solution of the differential equation $\partial^2 T/\partial x \partial y = 0$, is necessarily of the form $T = X + Y$, where the distribution X depends only on the x-coordinate, and Y only on the y-coordinate. It follows, therefore, that the most general function $g(x, y)$ having the mixed second partial $\mu \geq 0$ is necessarily of the form

$$g(x, y) = f(x, y) + X(x) + Y(y)$$

where $f(x, y)$ is lower semicontinuous and the functions $X(x)$ and $Y(y)$ are arbitrary, measurable functions of one variable. It now becomes clear that a canonical

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

determination is possible for the function f having a given positive μ as its mixed partial derivative: we simply require $f(x, 0) = f(0, y) = 0$, that is, that the function be lower semicontinuous and vanish on the coordinate axes. Here we are tacitly supposing that the coordinate axes intersect the domain of definition of the function.

In the case of monotone functions of one variable, the first derivative is a positive distribution; a whole class of monotone functions correspond to that same derivative, and certain normalization conditions must be imposed to make a canonical choice of the function. These conditions are usually twofold: the requirement that the function be lower semicontinuous and the requirement that the function vanishes at the origin. For higher dimensions the rules are almost the same: we need the function to vanish on the coordinate axes and to be lower semicontinuous.

It should be remarked that our definition of monotonicity differs somewhat from that initially taken by W. H. Young in his study of these functions. The reader may refer to Hobson [2]. However, we find that the use of the terminology of distribution theory considerably simplifies the presentation without departing in an essential way from the ideas of Young and others.

3. **Monotone matrix functions of two variables.** Let f be a real-valued Lebesgue measurable function of two real variables x, y in $(-1, 1)$. Let A be a selfadjoint operator in a Hilbert space H_1, with spectrum contained in $(-1, 1)$; let $A = \int_{-1}^{+1} x\, dE_x$ be its spectral resolution. Similarly, let $B = \int_{-1}^{+1} y\, dF_y$ be a selfadjoint operator in another Hilbert space H_2 with spectrum in $(-1, 1)$. The $E_x \otimes F_y$ determines a two parameter spectral family in the tensor product space $H_1 \otimes H_2$. By $f(A, B)$, we understand the operator

$$f(A, B) = \int_{-1}^{+1}\int_{-1}^{+1} f(x, y)\, dE_x \otimes dF_y$$

acting on the space $H_1 \otimes H_2$. If, in particular, H_1 and H_2 are of finite dimensions m and n respectively, and $A = \sum_{i=1}^{m} \lambda_i P_i$, $B = \sum_{i=1}^{n} \mu_i Q_i$, then

$$f(A, B) = \sum_{i=1}^{m} \sum_{j=1}^{n} f(\lambda_i, \mu_j) P_i \otimes Q_j.$$

In this case, f is called a matrix function of two variables.

Definition 3.1. A measurable function f is a monotone operator function, if for any H_1 and H_2, and for any selfadjoint operators A, A' in H_1, B, B' in H_2 whose spectrum is contained in $(-1, 1)$ and for which $A' \geq A$, $B' \geq B$ holds, then

$$f(A', B') - f(A', B) - f(A, B') + f(A, B) \geq 0. \tag{1}$$

If we consider operators on Hilbert spaces of finite dimensions only, a function $f(x, y)$ with property (1) is called a monotone matrix function of two variables.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

The collection of monotone matrix functions of two variables, monotone of order m in the first variable and of order n in the second variable, shall be denoted by $P_{m,n}$.

The following observations are obvious:

1. A Lebesgue measurable function f is in $P_{m,n}$ if and only if

$$f(A', B') - f(A, B) - f(A, B') + f(A, B) \geq 0$$

where $A - A = sP$, $B' - B = tQ$, P and Q are one dimensional projections and $s, t > 0$.

2. $f(x, y)$ is a monotone matrix function of two variables and $g(x)$, $h(y)$ are arbitrary matrix functions. Then

$$K(x, y) = f(x, y) + \alpha g(x) + \beta h(y) + \gamma$$

is a monotone matrix function.

3. Let $f(x)$ and $g(x)$ be monotone matrix functions of one variable in $(-1, 1)$. The product $h(x, y) = f(x)g(y)$ is a monotone in two variables.

4. $P_{m,n}$ is a convex cone, that is $\alpha f + \beta g \in P_{m,n}$ whenever $f, g \in P_{m,n}$ and $\alpha, \beta \geq 0$.

5. $P_{m+1,n} \subseteq P_{m,n}$, $P_{m,n+1} \subseteq P_{m,n}$.

6. $P_{m,n}$ is closed in the topology of pointwise convergence.

Examples. $h(x, y) = x^{\mu}y^{\nu}$ where $0 \leq \mu, \nu \leq 1$ and $h(x, y) = \log x \log y$ where $x, y > 0$ are monotone matrix functions of two variables.

Let f be a real-valued function of two real variables x, y in $(-1, 1)$. Assume (i) $f(x, 0) \equiv f(0, y) \equiv 0$ for all x, y and (ii) the first partial derivatives and the mixed second partial derivatives of f exist and are continuous. For such a function f and for $x_1 < x_2 < \cdots < x_m$; $y_1 < y_2 < \cdots < y_n$; we introduce the matrix

$$M(f) = [f[x_i x_k;\ y_j y_l]]_{i,j;k,l} \tag{2}$$

where

$$f[x_i x_k;\ y_j y_l] = \frac{f(x_i, y_j) - f(x_i, y_l) - f(x_k, y_j) + f(x_k, y_l)}{(x_i - x_k)(y_j - y_l)}$$

denotes the element in the (i, k)th row and (j, l)th column. $f[x_i x_k;\ y_j y_l]$ denotes

$$\frac{\partial f(x_i, y_j)/\partial x - \partial f(x_i, y_l)/\partial x}{y_j - y_l} \quad \text{if } x_i = x_k,$$

$$\frac{\partial f(x_i, y_j)/\partial y - \partial f(x_k, y_j)/\partial y}{x_i - x_k} \quad \text{if } y_j = y_l,$$

$$\frac{\partial^2 f(x_i, y_j)}{\partial x \partial y} \quad \text{if } x_i = x_k \text{ and } x_j = y_l.$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

The higher divided differences are defined inductively:

$$f[x_1x_2\cdots x_{n+1}; y_1y_2\cdots y_m]$$
$$= \frac{f[x_1x_2\cdots x_n; y_1y_2\cdots y_m] - f[x_1x_2\cdots x_{n-1}x_{n+1}; y_1y_2\cdots y_m]}{x_n - x_{n+1}}.$$

Theorem 3.2. *Let f be a real-valued function of two real variables x, y in $(-1, 1)$. Assume* (i) $f(x, 0) = f(0, y) = 0$ *for all x, y, and* (ii) *the first partial derivatives and the mixed second partial derivatives of f exist and are continuous. Then f is a monotone matrix function of two variables of order (m, n) if and only if, for $x_1 < x_2 < \cdots < x_m$; $y_1 < y_2 < \cdots < y_n$, the matrix*

$$M(f) = [f[x_ix_k; y_jy_l]]_{i,j;k,l},$$

where $1 \le i, k \le m$, $1 \le j, l \le n$, is nonnegative definite.

Remark. Since the addition of functions of one variable to f does not change its monotone character, the differentiability hypotheses on f are not automatically satisfied as in one variable case.

Proof. Since the proof of the necessary part of the theorem is exactly the same as the proof of a theorem of Koranyi [4, Theorem 4], we omit it here. Conversely, assume for $x_1 < x_2 < \cdots < x_m$, $y_1 < y_2 < \cdots < y_n$, the matrix

$$M(f) = [f[x_ix_k; y_jy_l]]_{i,k;k,l}, \qquad 1 \le i, k \le m, \quad 1 \le j, l \le n,$$

is nonnegative definite. We shall show that $f \in P_{m,n}$. Let G be the collection of all selfadjoint operators on a finite dimensional Hilbert space H, each having its spectrum in $(-1, 1)$. G is an open subset of the space of selfadjoint operators on H, that is, $\mathfrak{B}(H)$. For if $[\lambda, \mu]$ contains the closed, convex hull of the spectrum of $A \in G$, $d = \inf(|\lambda + 1|, |\mu - 1|)$, then $N_A = \{X \in \mathfrak{B}(H): \|X - A\| < d/2\}$ is an open neighbourhood of A contained in G. Let $G = G_i$ $(i = 1, 2)$. Now f induces a map of $G_1 \times G_2$ into $\mathfrak{B}(H_1 \otimes H_2)$.

Suppose A belonging to $\mathfrak{B}(H_1)$ and $B \in \mathfrak{B}(H_2)$ have simple spectra $\{x_i\}_{i=1}^m$ and $\{y_j\}_{j=1}^n$ respectively. Then the eigenvectors e_i corresponding to the eigenvalues x_i, $i = 1, \cdots, m$, and the eigenvectors f_j corresponding to the eigenvalues y_j, $j = 1, \cdots, n$, constitute bases for H_1 and H_2 respectively. P and Q denote one dimensional projections and their matrices relative to the bases $\{e_i\}_{i=1}^m$ and $\{f_j\}_{j=1}^n$ respectively are $[\sigma_{ik}]$ and $[\tau_{jl}]$. ϵ_1 and ϵ_2 are arbitrary positive real numbers. We now state the following lemma:

Lemma. *f has a mixed differential, that is, there exists an operator $L^{(\epsilon_1 P, \epsilon_2 Q)}$ from $\mathfrak{B}(H_1) \times \mathfrak{B}(H_2) \to \mathfrak{B}(H_1 \otimes H_2)$, such that*

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$$\|f(A+\epsilon_1 P, B+\epsilon_2 Q) - f(A+\epsilon_1 P, B) - f(A, B+\epsilon_2 Q) + f(A, B) - L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\| = o(\epsilon)$$

where the matrix of $L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)$ *relative to the basis* $e_i \otimes f_j$ $(i = 1, \cdots, m,\ j = 1, \cdots, n)$ *is given by the following*:

$$L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)|_{i,j;k,l} = \epsilon_1 \epsilon_2 f[x_i x_k; y_j y_l]\sigma_{ik}\tau_{jl};$$

and $\epsilon = \epsilon_1^{\lambda}\epsilon_2^{\mu}$, $\lambda, \mu > 0$ *and* $\lambda + \mu \geq 3$.

We assume the Lemma and proceed to show that $f \in P_{m,n}$. With A and B the same as in the Lemma, we have

$$f(A+\epsilon_1 P, B+\epsilon_2 Q) - f(A+\epsilon_1 P, B) - f(A, B+\epsilon_2 Q) + f(A, B) = L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B) + O(\epsilon).$$

Thus

$$([f(A+\epsilon_1 P, B+\epsilon_2 Q) - f(A+\epsilon_1 P, B) - f(A, B+\epsilon_2 Q) + f(A, B)]\phi \otimes \psi, \phi \otimes \psi) = (L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\phi \otimes \psi, \phi \otimes \psi) + O(\epsilon),$$

where $\|\phi\| = \|\psi\| = 1$. Define $F(\epsilon_1, \epsilon_2)$ by

$$F(\epsilon_1, \epsilon_2) = (f(A+\epsilon_1 P, B+\epsilon_2 Q)\phi \otimes \psi, \phi \otimes \psi), \qquad \epsilon_1, \epsilon_2 > 0.$$

$F(\epsilon_1, \epsilon_2)$ has a nonnegative, mixed partial derivative w.r.t. ϵ_1 and ϵ_2, that is, $\partial^2 F/\partial\epsilon_1\partial\epsilon_2 \geq 0$. Hence, F is a monotone function of two variables. Therefore $F(\epsilon_1, \epsilon_2) - F(\epsilon_1, 0) - F(0, \epsilon_2) + F(0, 0) \geq 0$. That is, $f \in P_{m,n}$.

In general, when A and B do not have simple spectra, we perturb A and B to obtain A' and B' with simple spectra. Then

$$f(A'+\epsilon_1 P, B'+\epsilon_2 Q) - f(A'+\epsilon_1 P, B') - f(A', B'+\epsilon_2 Q) + f(A', B') \geq 0,$$

and in the limit, we have

$$f(A+\epsilon_1 P, B+\epsilon_2 Q) - f(A+\epsilon_1 P, B) - f(A, B+\epsilon_2 Q) + f(A, B) \geq 0.$$

Proof of the Lemma.

Case I. If f is a polynomial $p(\lambda, \mu)$ of degree k. When $k = 1$, $p(\lambda, \mu) = \alpha\lambda + \beta\mu + \gamma$,

$$\begin{aligned} &p(A+\epsilon_1 P, B+\epsilon_2 Q) - p(A+\epsilon_1 P, B) - p(A, B+\epsilon_2 Q) + p(A, B) \\ &\quad = \alpha(A+\epsilon_1 P) \otimes I + \beta I \otimes (B+\epsilon_2 Q) + \gamma(I \otimes I) - \alpha(A+\epsilon_1 P) \otimes I \\ &\quad\quad - \beta I \otimes B - \gamma(I \otimes I) - \alpha(A \otimes I) - \beta I \otimes (B+\epsilon_2 Q) \\ &\quad\quad - \gamma(I \otimes I) + \beta I \otimes B + \gamma(I \otimes I) + \alpha A \otimes I \\ &\quad = 0; \end{aligned} \tag{3}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$$L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\big|_{i,j;k,l}$$

$$= \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}\,\frac{[\alpha x_i + \beta y_j + \gamma - \alpha x_i - \beta y_l - \gamma - \alpha x_k - \beta y_j - \gamma + \alpha x_k + \beta y_l + \gamma]}{(x_i - x_k)(y_j - y_l)}$$

$$= 0.$$

Thus

$$\|p(A + \epsilon_1 P, B + \epsilon_2 Q) - p(A + \epsilon_1 P, B) - p(A, B + \epsilon_2 Q) + p(A, B) - L^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\| = 0.$$

Suppose the assertion is true for polynomials of degree $k - 1$. We shall show that it is true for polynomials of degree K. We may assume without loss of generality that $p(0, 0) = 0$, for otherwise we may begin with $p(\lambda, \mu) - p(0, 0)$. Then $p(\lambda, \mu) = \lambda R(\lambda, \mu) + \mu S(\lambda, \mu)$.

$$\begin{aligned}
&p(A + \epsilon_1 P, B + \epsilon_2 Q) - p(A + \epsilon_1 P, B) - p(A, B + \epsilon_2 Q) + p(A, B)\\
&\quad = [(A + \epsilon_1 P) \otimes I]R(A + \epsilon_1 P, B + \epsilon_2 Q) + [I \otimes (B + \epsilon_2 Q)]S(A + \epsilon_1 P, B + \epsilon_2 Q)\\
&\qquad - [(A + \epsilon_1 P) \otimes I]R(A + \epsilon_1 P, B) - [I \otimes B]S(A + \epsilon_1 P, B)\\
&\qquad - [A \otimes I]R(A, B + \epsilon_2 Q) - [I \otimes (B + \epsilon_2 Q)]S(A, B + \epsilon_2 Q)\\
&\qquad + [A \otimes I]R(A, B) + [I \otimes B]S(A, B)\\
&\quad = (A \otimes I)[R(A + \epsilon_1 P, B + \epsilon_2 Q) - R(A + \epsilon_1 P, B) - R(A, B + \epsilon_2 Q) + R(A, B)]\\
&\qquad + (I \otimes B)[S(A + \epsilon_1 P, B + \epsilon_2 Q) - S(A + \epsilon_1 P, B) - S(A, B + \epsilon_2 Q) + S(A, B)]\\
&\qquad + (\epsilon_1 P \otimes I)[R(A + \epsilon_1 P, B + \epsilon_2 Q) - R(A + \epsilon_1 P, B) - R(A, B + \epsilon_2 Q) + R(A, B)]\\
&\qquad + (I \otimes \epsilon_2 Q)[S(A + \epsilon_1 P, B + \epsilon_2 Q) - S(A, B + \epsilon_2 Q) - S(A + \epsilon_1 P, B) + S(A, B)]\\
&\qquad + (\epsilon_1 P \otimes I)[R(A, B + \epsilon_2 Q) - R(A, B)] + (I \otimes \epsilon_2 Q)[S(A + \epsilon_1 P, B) - S(A, B)]\\
&\quad = (A \otimes I)L_R^{(\epsilon_1 P, \epsilon_2 Q)}(A, B) + I \otimes BL_S^{(\epsilon_1 P, \epsilon_2 Q)}(A, B) + (\epsilon_1 P \otimes I)[R(A, B + \epsilon_2 Q) - R(A, B)]\\
&\qquad + (I \otimes \epsilon_2 Q)[S(A + \epsilon_1 P, B) - S(A, B)] + (\epsilon_1 P \otimes I)L_R^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\\
&\qquad + (I \otimes \epsilon_2 Q)L_S^{(\epsilon_1 P, \epsilon_2 Q)}(A, B).
\end{aligned} \tag{4}$$

The last two terms on the right side may be neglected, since they are of higher order in ϵ. Moreover,

$$(A \otimes I)L_R^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\Big|_{i,j;k,l} = \sum x_i\delta_{im}\,\delta_{jn}R[x_m x_k;\, y_n y_l]\sigma_{mk}\tau_{nl}$$

$$= \epsilon_1\epsilon_2 x_i R[x_i x_k;\, y_j y_1]\sigma_{ik}\tau_{jl}$$

and

$$(I \otimes B)L_S^{(\epsilon_1 P, \epsilon_2 Q)}(A, B)\Big|_{i,j;k,l} = \epsilon_1\epsilon_2 y_j S[x_i x_k;\, y_j y_l]\sigma_{ik}\tau_{jl}.$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

We assert that

$$(5)\qquad (\epsilon_1 P \otimes I)[R(A, B+\epsilon_2 Q) - R(A, B)]|_{i,j;k,l} = \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}\frac{R(x_k, y_j) - R(x_k, y_l)}{y_j - y_l}$$

and

$$(I \otimes \epsilon_2 Q)[S(A+\epsilon_1 P, B) - S(A, B)]|_{i,j;k,l} = \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}\frac{S(x_i, y_l) - S(x_k, y_l)}{x_i - x_k}.$$

Then the $(i, j; k, l)$th element of the right-hand side of (4)

$$\begin{aligned}
&= \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}\Big\{x_i R[x_i x_k; y_j y_l] + y_j S[x_i x_k; y_j y_l] \\
&\qquad + \frac{R(x_k, y_j) - R(x_k, y_l)}{y_j - y_l} + \frac{S(x_i, y_l) - S(x_k, y_l)}{x_i - x_k}\Big\} \\
&= \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}p[x_i x_k; y_j y_l],
\end{aligned}$$

which is the $(i, j; k, l)$th element of the left-hand side of (4). In order to complete the proof of assertion (3), when f is a polynomial of degree $k > 1$, we need to prove assertion (5). We proceed to show that

$$(\epsilon_1 P \otimes I)[R(A, B+\epsilon_2 Q) - R(A, B)]|_{i,j;k,l} = \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}\frac{R(x_k, y_j) - R(x_k, y_l)}{y_j - y_l},$$

where R is a polynomial

$$R(\lambda, \mu) = \sum_{\alpha,\beta} A_{\alpha\beta}\lambda^\alpha\mu^\beta.$$

Writing out the $(i, j; k, l)$th element of $\lambda^\alpha\mu^\beta$, and neglecting the terms of higher order in ϵ,

$$\begin{aligned}
&(\epsilon_1 P \otimes I)[A^\alpha \otimes (B+\epsilon_2 Q)^\beta - A^\alpha \otimes B^\beta]_{i,j;k,l} \\
&= (\epsilon_1 P \otimes I)A^\alpha \otimes (\epsilon_2 B^{\beta-1}Q + \epsilon_2 B^{\beta-2}QB + \epsilon_2 B^{\beta-3}QB^2 + \cdots + \epsilon_2 QB^{\beta-1})|_{i,j;k,l} \\
&= \epsilon_1\epsilon_2[(PA^\alpha \otimes B^{\beta-1}Q) + (PA^\alpha \otimes B^{\beta-2}QB) + \cdots + (PA^\alpha \otimes QB^{\beta-1})]|_{i,j;k,l} \\
&= \epsilon_1\epsilon_2\left[\left(\sum \sigma_{im}x_m^\alpha\delta_{mk}\right)\left(\sum y_j^{\beta-1}\delta_{jn}\tau_{nl}\right) + \cdots\right] \\
&= \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}x_k^\alpha(y_j^{\beta-1} + y_j^{\beta-2}y_l + \cdots + y_l^{\beta-1}) \\
&= \epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}x_k^\alpha(y_j^\beta - y_l^\beta)/(y_j - y_l).
\end{aligned}$$

Therefore, the $(i, j; k, l)$th element of $(\epsilon_1 P \otimes I)[R(A, B+\epsilon_2 Q) - R(A, B)]$ is

$$\epsilon_1\epsilon_2\sigma_{ik}\tau_{jl}(R(x_k, y_j) - R(x_k, y_l))/(y_j - y_l).$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

The proof the second part of the assertion (5) is similar.

Case II. To prove the lemma in general, we shall construct a polynomial $p(\lambda, \mu)$ satisfying the following conditions:

(i) $p(x_i, y_j) = f(x_i, y_j)$;

(ii) $\partial p(x_i, y_j)/\partial x = \partial f(x_i, y_j)/\partial x$, $\partial p(x_i, y_j)/\partial y = \partial f(x_i, y_j)/\partial y$, $i = 1, \cdots, m$, $j = 1, \cdots, n$; and

(iii) $\partial^2 p(x_i, y_j)/\partial x \partial y = \partial^2 f(x_i, y_j)/\partial x \partial y$.

Show that f and p have the same mixed differential. Let $\psi = f - p$. For ψ, the first partial derivative, mixed second partial derivatives exist, are continuous and vanish at (x_i, y_j) $(i = 1, \cdots, m;\ j = 1, \cdots, n)$. Let x_i' $(i = 1, \cdots, m)$ and y_i' $(i = 1, \cdots, n)$ be the eigenvalues of $A + \epsilon_1 P$, and $B + \epsilon_2 Q$ respectively. Then

$$|x_i' - x_i| \le \epsilon_1 \|P\| \ (i = 1, \cdots, m) \quad \text{and} \quad |y_i' - y_i| \le \epsilon_2 \|Q\| \quad (i = 1, \cdots, n).$$

Now,

$$\begin{aligned} &\|\psi(A + \epsilon_1 P, B + \epsilon_2 Q) - \psi(A + \epsilon_1 P, B) - \psi(A, B + \epsilon_2 Q) + \psi(A, B)\| \\ &= \left\| \frac{\psi(x_i', y_j') - \psi(x_i', y_l) - \psi(x_k, y_j') + \psi(x_k, y_l)}{(x_i' - x_k)(y_j' - y_l)} \right\|. \end{aligned} \tag{6}$$

Each term in the matrix on the right-hand side of (6) is arbitrarily small for small ϵ_1, ϵ_2. Thus the mixed differential of ψ is zero, that is, f and p have the same mixed differential. This completes the proof.

Remark 3.3. Let $\phi(x, y)$ be a positive, C^∞-function of two variables which vanishes outside the disc $x^2 + y^2 \le 1$ and which is also normalised that

$$\iint \phi(x, y)\, dx\, dy = 1.$$

Given such a function, we form for $\epsilon > 0$, the family of functions $\phi_\epsilon(x, y) = (1/\epsilon^2)\phi(x/\epsilon, y/\epsilon)$; each ϕ_ϵ is a positive, C^∞-function vanishing outside the disc of radius ϵ, and $\iint \phi_\epsilon(x, y)\, dx\, dy = 1$. Let f be a monotone matrix function of order (m, n). We form the regularisation of f,

$$\begin{aligned} f_\epsilon(x, y) &= \iint \phi_\epsilon(x - u, y - v) f(u, v)\, du\, dv \\ &= \frac{1}{\epsilon^2} \iint \phi\left(\frac{x - u}{\epsilon}, \frac{y - v}{\epsilon}\right) f(u, v)\, du\, dv. \end{aligned} \tag{7}$$

Since a function representable in the form (7) is, on every compact subset, a uniform limit of positive linear combinations of monotone matrix functions, it follows that $f_\epsilon(x, y)$ is also a monotone matrix function.

Theorem 3.4. *If $g(x, y) = f_\epsilon(x, y)$ is the regularisation of $f(x, y)$ in $P_{m,n}$ then the matrix $D(g(x, y))$, where*

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$$D(g(x, y)) = \left(\frac{1}{(i+k-1)!} \frac{1}{(j+l-1)!} \frac{\partial^{i+j+k+l-2} f}{\partial^{i+k-1} x \partial^{j+l-1} y} \right)_{i,j;k,l}$$

$$i, k = 1, \cdots, m, \quad j, l = 1, \cdots, n,$$

is nonnegative definite.

Proof. Since $g(x, y)$ is a monotone matrix function of order (m, n), it follows from the proof of Theorem 3.2 that for

$$x_1 < x_1' < x_2 < x_2' < \cdots < x_m < x_m'; \; y_1 < y_1' < y_2 < y_2' < \cdots < y_n < y_n',$$

the matrix

$$[g[x_i' x_k; y_j' y_l]] \geq 0. \tag{8}$$

We divide the determinant of (8) by

$$L = \prod_{i,k=1;i>k}^{m} (x_i - x_k)^n (x_i' - x_k')^n \prod_{j,l=1;j>l}^{n} (y_j - y_l)^m (y_j' - y_l')^m,$$

and proceed with the quotient

$$(g) = \frac{1}{L} \begin{vmatrix} g[x_1' x_1; y_1' y_1] & \cdots & g[x_1' x_1; y_1' y_n] & \cdots & g[x_1' x_m; y_1' y_1] & \cdots & g[x_1' x_m; y_1' y_n] \\ \vdots & & & & & & \\ g[x_1' x_1; y_n' y_1] & \cdots & g[x_1' x_1; y_n' y_n] & \cdots & g[x_1' x_m; y_n' y_1] & \cdots & g[x_1' x_m; y_n' y_n] \\ \vdots & & & & & & \\ g[x_m' x_1; y_1' y_1] & \cdots & g[x_m' x_1; y_1' y_n] & \cdots & g[x_m' x_m; y_1' y_1] & \cdots & g[x_m' x_m; y_1' y_n] \\ \vdots & & & & & & \\ g[x_m' x_1; y_n' y_1] & \cdots & g[x_m' x_1; y_n' y_n] & \cdots & g[x_m' x_m; y_n' y_1] & \cdots & g[x_m' x_m; y_n' y_n] \end{vmatrix} \tag{9}$$

as follows; subtract the $(n-1)$th column from the nth, the $(n-2)$th column from the $(n-1)$th, $\cdots$, the 1st from the 2nd and take in each time the corresponding denominator $(y_n - y_{n-1}), \cdots, (y_2 - y_1)$. Repeat this process starting with the sub-

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

traction of the $(n-2)$th column from the nth. During this process we divide by $(y_n - y_{n-2}), \cdots, (y_3 - y_1)$. Continuing this way, at the $(n-1)$th step we are left with the subtraction of the 1st column from the nth column and division by $(y_n - y_1)$. Repeat this process for each of the blocks from the $(n+1)$th column to the $2n$th column, $\cdots$, the $((m-1)n+1)$th column to the (mn)th column. Thus all terms $\Pi^n_{j,l=1;j>l}(y_j - y_l)^m$ are used up. Subtract the kth column of the $(m-2)$th block, $\cdots$, the kth column of the 1st block from the kth column of the 2nd block, and take in each time the corresponding denominator $(x_m - x_{m-1}), \cdots, (x_2 - x_1)$, for $k = 1, \cdots, n$. Repeat the above process starting with the subtraction of the kth column of the $(m-2)$th block from the kth column of the mth block. During this process we divide by $(x_m - x_{m-2}), \cdots, (x_3 - x_1)$. Continuing this way at the $(n-1)$th step, we are left with the subtraction of the kth column of the 1st block from the kth column of the mth block and division by $(x_m - x_1)$, $k = 1, \cdots, n$. Thus all terms $\Pi^m_{i,k=1;i>k}(x_1 - x_k)^n$ are used up. If one carries out this process with rows, and allows $x_i, x_i' \to x$ $(i = 1, 2, \cdots, m)$, $y_i', y_i \to y$ $(i = 1, 2, \cdots, n)$, then the determinant reduces to a determinant which has

$$\frac{1}{(i+k-1)!}\,\frac{1}{(j+l-1)!}\,\frac{\partial^{i+j+k+l-2} f}{\partial^{i+k-1}x\,\partial^{j+l-1}y} \qquad (i,\ k = 1, \cdots, m; \quad j,\ l = 1, \cdots, n)$$

as an element in the (i, k)th row and the (j, l)th column. Thus the new determinant and all determinants of lower order on the main diagonal are nonnegative. From this it follows by the usual algebraic reasoning that the matrix $D(g(x, y))$ is nonnegative definite. This completes the proof.

Lemma. *Let $f_n(x, y)$ be a sequence of nonnegative convex functions defined on $(-1, 1) \times (-1, 1)$ which converges as a sequence of distributions to some distribution T. Then there exists a subsequence $f_{n_k}(x, y)$ which converges uniformly on compacts to some convex function $f_0(x, y)$, and the distribution $f_0(x, y)$ coincides with T.*

Proof. We show that the sequence is equicontinuous on compact sets. It follows from the Ascoli-Arzela theorem that there is a subsequence converging uniformly on compact subsets, and it is evident that the limiting function is convex and represents the limiting distribution T.

Let $M(x, y) = \sup_n f_n(x, y)$. We shall show that $M(x, y)$ is finite in $(-1, 1) \times (-1, 1)$. If not, let (x_0, y_0) be a point in the square where $M(x_0, y_0)$ is infinite. Then there exists a subsequence $(f_{n_k}(x, y))$ such that $f_{n_k}(x_0, y_0)$ approaches $+\infty$ as n_k approaches $+\infty$. Consider a circle $S_r(x_0, y_0)$ wholly contained in $(-1, 1) \times (-1, 1)$. Since f_{n_k} is convex, thus subharmonic for each n_k, we have

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

$$f_{n_k}(x_0, y_0) \leq \frac{1}{|S_r(x_0, y_0)|} \iint_{S_r(x_0,y_0)} f_{n_k}(x, y)\, dx\, dy, \qquad |S_r(x_0, y_0)| = \pi r^2. \tag{10}$$

The left-hand side of (10) approaches $+\infty$ as n_k approaches $+\infty$. Thus $\iint_{S_r(x_0,y_0)} f_{n_k}(x, y)\, dx\, dy$ is arbitrarily large for large n_k. Let ϕ be a positive test function which is ≥ 1 on S_r and vanishes outside $S_{r+\epsilon}$. Then

$$\iint f_{n_k}(x, y)\phi(x, y)\, dx\, dy \geq \iint_{S_r(x_0,y_0)} f_{n_k}(x, y)\phi(x, y)\, dx\, dy, \tag{11}$$

and the right-hand side of (11) is arbitrarily large for large n_k. This contradicts the fact that f_{n_k} converges as a distribution. Thus $M(x, y)$ is finite and hence convex.

We shall show that $f_n(x, y)$ is uniformly Lipschitzian on compacts.

Let K be a compact subset in $(-1, 1) \times (-1, 1)$. Surround K by an arc C lying completely within $(-1, 1) \times (-1, 1)$. Let d denote the distance from K to C, and note that for any $f_n(x, y)$ in the sequence and $(x, y) \in K$,

$$|\partial f_n(x, y)/\partial x| \leq \max\,(|\partial f_n(\alpha, y)/\partial x|\, |\partial f_n(\beta, y)/\partial x|)$$

where (α, y) and (β, y) are the points in which the horizontal line through (x, y) meets the boundary of K and $\alpha < \beta$. But

$$\begin{aligned} \partial f_n(\beta, y)/\partial x &\leq (f_n(\beta + d, y) - f_n(\beta, y))/d \\ &\leq (f_n(\beta + d, y) + f_n(\beta, y))/d \leq 2M/d \end{aligned}$$

where M denotes the maximum of $M(x, y)$ over the curve C. Also $-2M/d \leq \partial f_n(\alpha, y)/\partial x$; since $f_n(x, y)$ is convex, $-2M/d \leq \partial f_n(\alpha, y)/\partial x \leq \partial f_n(\beta, y)/\partial x \leq 2M/d$, that is, $|\partial f_n(x, y)/\partial x| \leq 2M/d$. Obviously, the same argument serves to show that $|\partial f_n(x, y)/\partial y| \leq 2M/d$, and therefore $4M/d$ is a uniform bound for $|\text{grad } f_n|$ over K. It is now clear that the functions of the sequence are uniformly Lipschitzian, hence equicontinuous.

Theorem 3.5. *If $f(x, y)$ is a monotone matrix function of order (m, n), $m > n > 1$, then the distribution derivatives $\partial^{2p+2q-6} f/\partial^{2p-3}x\partial^{2q-3}y$ $(p = 2, \cdots, m;$ $q = 2, \cdots, n)$ are convex and positive.*

Proof. We form the regularisation $g(x, y) = f_\epsilon(x, y)$ of $f(x, y)$. By Theorem 3.4, $D(g(x, y))$ is a nonnegative definite matrix. Hence all its diagonal terms, namely,

$$\frac{\partial^2 g}{\partial x \partial y}, \cdots, \frac{\partial^{2n} g}{\partial x \partial^{2n-1} y}, \cdots, \frac{\partial^{2m} g}{\partial^{2m-1} x \partial y}, \cdots, \frac{\partial^{2m+2n-2} g}{\partial^{2m-1} x \partial^{2n-1} y}$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

are nonnegative. We shall show that

$$\frac{\partial^2 g}{\partial x \partial y}, \ldots, \frac{\partial^{2n-2} g}{\partial x \partial^{2n-3} y}, \ldots, \frac{\partial^{2m-2} g}{\partial^{2m-3} x \partial y}, \ldots, \frac{\partial^{2m+2n-6} g}{\partial^{2m-3} x \partial^{2n-3} y}$$

are convex. We consider for $(1 < p \leq m;\ 1 < q \leq n)$ the 2×2 matrix

$$(12) \quad \begin{bmatrix} \dfrac{1}{(2p-3)!\,(2q-1)!} \dfrac{\partial^{2p+2q-4} g}{\partial^{2p-3} x \partial^{2q-1} y} & \dfrac{1}{(2p-2)!\,(2q-2)!} \dfrac{\partial^{2p+2q-4} g}{\partial^{2p-2} x \partial^{2q-2} y} \\ \dfrac{1}{(2p-2)!(2q-2)!} \dfrac{\partial^{2p+2q-4} g}{\partial^{2p-2} x \partial^{2q-2} y} & \dfrac{1}{(2p-1)!(2q-3)!} \dfrac{\partial^{2p+2q-4} g}{\partial^{2p-1} x \partial^{2q-3} y} \end{bmatrix}$$

From the positivity of the matrix (12), it follows that

$$\frac{\partial^{2p+2q-4} g}{\partial^{2p-3} x \partial^{2q-1} y} \frac{\partial^{2p+2q-4} g}{\partial^{2p-1} x \partial^{2q-3} y} - \left(\frac{\partial^{2p+2q-4} g}{\partial^{2p-2} x \partial^{2q-2} y} \right)^2$$

$$\geq \frac{\partial^{2p+2q-4} g}{\partial^{2p-3} x \partial^{2q-1} y} \frac{\partial^{2p+2q-4} g}{\partial^{2p-1} x \partial^{2q-3} y} - \frac{(2p-1)(2q-1)}{(2p-2)(2q-2)} \left(\frac{\partial^{2p+2q-4} g}{\partial^{2p-2} x \partial^{2q-2} y} \right)^2 \geq 0,$$

that is $\partial^{2p+2q-6} g / \partial^{2p-3} x \partial^{2q-3} y$ $(1 < p \leq m;\ 1 < q \leq n)$ is convex. Since $\partial^{2p+2q-6} f_\epsilon / \partial^{2p-3} x \partial^{2q-3} y$ $(1 < p \leq m;\ 1 < q \geq n)$ is convex for every ϵ, and $\partial^{2p+2q-6} f_\epsilon / \partial^{2p-3} x \partial^{2q-3} y$ converges to $\partial^{2p+2q-6} f / \partial^{2p-3} x \partial^{2q-3} y$ as a distribution, by the lemma there exists an infinite sequence which converges to $\partial^{2p+2q-6} f / \partial^{2p-3} x \partial^{2q-3} y$ uniformly on compacts. $\partial^{2p+2q-6} f / \partial^{2p-3} x \partial^{2q-3} y$ $(1 < p \leq m;\ 1 < q < n)$, being the limit of a sequence of convex functions, is itself convex. This completes the proof.

Acknowledgement. The author is indebted to Professor W. F. Donoghue, Jr., for his continued interest and expert supervision.

REFERENCES

1. J. S. Bendat and S. Sherman, *Monotone and convex operator functions*, Trans. Amer. Math. Soc. 79 (1955), 58–71. MR 18, 588.

2. E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier's series*. Vol. 1, Dover, New York, 1958. MR 19, 1166.

3. A. Korányi, *On a theorem of Löewner and its connections with resolvents of self-adjoint transformations*, Acta Sci. Math. Szeged 17 (1956), 63–70. MR 18, 588.

4. ———, *On some classes of analytic functions of several variables*, Trans. Amer. Math. Soc. 101 (1961), 520–554. MR 25 #226.

5. K. Loewner, *Über monotone Matrixfunctionen*, Math. Z. 38 (1934), 177–216.

6. ———, *Advanced matrix theory*, Mimeographed Notes, Stanford University, Stanford, Calif., 1957.

7. Béla Sz.-Nagy, *Remarks to the preceding paper of A. Korányi*, Acta Sci. Math. Szeged 17 (1956), 71–75. MR 18, 588.

DEPARTMENT OF MATHEMATICS, PANJAB UNIVERSITY, CHANDIGARH-14, INDIA

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use