THE THEORY OF INFINITELY DIVISIBLE MATRICES AND KERNELS

BY

ROGER A. HORN[1]

Introduction and notation. Kernels which generate nonnegative definite or semi-definite quadratic forms play an important role in many branches of mathematics, but general kernels whose fractional powers *all* have this same property have only recently been studied. A special family of such kernels has long been of importance in probability theory, and it is this family which inspires the name *infinitely divisible kernels* for the general class.

It was work of C. Loewner concerning semigroups of conformal mappings which first drew attention to these kernels, and interest in them could only increase when he showed recently that the Green's function for Laplace's equation is, under certain conditions, an infinitely divisible kernel. In this paper we shall develop a general theory of infinitely divisible kernels and indicate briefly a few of its applications.

It is convenient to begin by examining the discrete analogues of the infinitely divisible kernels, i.e., infinitely divisible matrices. In §1 we characterize this class of matrices and show that the occurrence of zeros among their entries is subject to a certain condition which is most easily stated in the language of graph theory. We then use the matrical theory to obtain corresponding results for continuous infinitely divisible kernels, and we develop the technical tools for dealing with these kernels which are needed later in the applications.

We shall use R^m and C^m to denote m-dimensional real and complex space, respectively, with $R \equiv R^1$, $C \equiv C^1$, and $R^+ \equiv \{x \in R \mid x > 0\}$. If $S \subseteq R^m$ is a set with appropriate regularity properties, $m \geq 1$, and if $0 \leq k < \infty$, we define $C^k(S) \equiv \{f : S \to C \mid f$ and all its derivatives through the kth order are continuous on $S\}$, $C^\infty(S) \equiv \bigcap_{k=1}^\infty C^k(S)$, $C_0^k(S) \equiv \{f \in C^k(S) \mid f$ has compact support contained in $S\}$, $C_0^\infty(S) \equiv \bigcap_{k=1}^\infty C_0^k(S)$, $C(S) \equiv C^0(S)$, $C_0(S) \equiv C_0^0(S)$, and

$$L_0(S) = \left\{ f \in C_0(S) \;\middle|\; \int_S f(P)\, dP = 0 \right\}.$$

Finally, $C_0^k(S, C^n) \equiv \{F(x) \mid F(x) = (F_i(x))_{i=1}^n$ and $F_i(x) \in C_0^k(S)\}$.

1. Matrices. If $A = (a_{ij})_{i,j=1}^n$, $a_{ij} \in C$, is an $n \times n$ matrix, we shall say $A \succeq 0$ on C^n (or sometimes just $A \succeq 0$) if A generates a nonnegative definite or semidefinite

quadratic form on C^n, i.e., if $\langle x, Ax \rangle \equiv \sum_{i,j=1}^n x_i a_{ij} \bar{x}_j \geqq 0$ for all $x = (x_i)_{i=1}^n \in C^n$. More generally, if $T \subseteq C^n$ is any subspace, we shall write $A \succeq 0$ on T if and only if $\langle x, Ax \rangle \geqq 0$ for all $x \in T$. By $A \gg 0$ (strictly positive definite) we shall mean $\langle x, Ax \rangle > 0$ for all $x \in C^n$, $x \neq 0$.

Define the subspace $L^n \equiv \{x \in C^n \mid \sum_{i=1}^n x_i = 0\}$ so that we have the corresponding notion of $A \succeq 0$ on L^n. We shall always mean by $A \geqq 0$ or $A > 0$ that all $a_{ij} \geqq 0$ or $a_{ij} > 0$, respectively. We write $A^* \equiv (\bar{a}_{ji})_{i,j=1}^n$ for the conjugate transposed matrix, and use δ_{ij} only for the Kronecker symbol, i.e., $\delta_{ij} \equiv 1$ if $i = j$, $\delta_{ij} \equiv 0$ if $i \neq j$.

We observe that if the quadratic form $\langle x, Ax \rangle$ assumes only real values for $x \in C^n$, then A is necessarily Hermitian ($A = A^*$). In particular, if $A \succeq 0$, then A must be Hermitian. All functions and products of matrices will be taken in the pointwise sense, i.e., $f(A) \equiv (f(a_{ij}))_{i,j=1}^n$ and $A \circ B \equiv (a_{ij} b_{ij})_{i,j=1}^n$, and we recall the *Schur product theorem* (vid. [14, p. 14], or [1, p. 94]):

THEOREM 1.1. *If $A \succeq 0$ and $B \succeq 0$, then $A \circ B \succeq 0$. Furthermore, if $A \gg 0$ and $B \gg 0$, then $A \circ B \gg 0$.*

This theorem shows that if $f(t) \equiv t^n$, $n = 1, 2, \ldots$, then $f(A) \succeq 0$ whenever $A \succeq 0$, so it is not unnatural to ask what other functions share this property of leaving invariant the convex cone of nonnegative quadratic forms. C. Loewner has found (oral communication) certain necessary conditions, and we have the

THEOREM 1.2. *Let $f \in C(R^+)$, let $n \geqq 1$ be an integer, and suppose that $f(A) \succeq 0$ for every $n \times n$ matrix A such that $A > 0$ and $A \succeq 0$. Then $f \in C^{n-3}(R^+)$, $f^{(k)}(x) \geqq 0$ for all $x \in R^+$ and all $k = 0, 1, 2, \ldots, n-3$, and $f^{(n-3)}$ is a convex and monotone nondecreasing function on R^+. In particular, if $f \in C^{n-1}(R^+)$, then $f^{(k)}(x) \geqq 0$ for all $x \in R^+$ and all $k = 0, 1, 2, \ldots, n-1$.*

Proof. Assume first that $f \in C^\infty(R^+)$. It is trivial that for $n = 1$ we have $f(x) \geqq 0$ for all $x \in R^+$, so we may assume that $n \geqq 2$ and proceed by induction, observing that the hypotheses guarantee that $f(B) \succeq 0$ whenever $B > 0$, $B \succeq 0$, is an $m \times m$ matrix, $1 \leqq m \leqq n$. Thus, we may assume that $f^{(k)}(x) \geqq 0$ for all $x \in R^+$ and all $k = 0, 1, \ldots, n-2$, and we must prove that $f^{(n-1)}(x) \geqq 0$ for all $x \in R^+$.

Let $a > 0$ and $\alpha = (\alpha_i)_{i=1}^n \in R^n$, so that for all sufficiently small $t \in R$, the matrix $A(t) \equiv (a + t\alpha_i\alpha_j)_{i,j=1}^n \succeq 0$ and $A(t) > 0$. But then by hypothesis

$$f(A(t)) = (f(a + t\alpha_i\alpha_j))_{i,j=1}^n \succeq 0$$

so that $\Delta(t) \equiv \det f(A(t)) \geqq 0$. Because $f(A(0))$ has all columns identical, $\Delta(0) = 0$. Furthermore,

$$(d^k/dt^k)\, \Delta(t)|_{t=0} \equiv \Delta^{(k)}(0) = 0$$

for all $k = 0, 1, 2, \ldots, n(n-1)/2 - 1$ because in each case there is at least one pair of proportional columns in each of the n^k determinant terms arising from the

k-fold differentiation. Let $N \equiv n(n-1)/2$ and conclude from an explicit calculation that

$$0 \leqq \lim_{t \to 0} \frac{\Delta(t)}{t^N} = \Delta^{(N)}(0)$$

$$= \binom{N}{1}\binom{N-1}{2}\binom{N-3}{3}\binom{N-6}{4}\cdots\binom{2n-3}{n-2} V^2(\alpha) f(a) f'(a)\cdots f^{(n-1)}(a),$$

where

$$\binom{m}{l} \equiv \frac{m!}{l!(m-l)!} > 0$$

is the binomial coefficient and

$$V(\alpha) \equiv \det \begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \cdots & \alpha_n^{n-1} \end{pmatrix} = \prod_{1 \leqq i < j \leqq n} (\alpha_i - \alpha_j)$$

is the van der Monde determinant. $V^2(\alpha) > 0$ if all $\alpha_i \neq \alpha_j$.

Therefore, we find in general that $f(a)f'(a)\cdots f^{(n-1)}(a) \geqq 0$ is a necessary condition, and our induction hypothesis is that $f(a), f'(a), \ldots, f^{(n-2)}(a)$ are all nonnegative. Since we cannot conclude directly from this that $f^{(n-1)}(a) \geqq 0$, recall that $g(x) \equiv x^n$ leaves invariant the $n \times n$ nonnegative quadratic forms (the Schur product theorem) and so the function $f(x) + \tau g(x)$ has the same property for all $\tau \geqq 0$. But then we have just shown that for $\tau \geqq 0$, necessarily

$$p(\tau) \equiv (f(a) + \tau a^n)(f'(a) + \tau n a^{n-1})\cdots(f^{(n-1)}(a) + \tau n!) \geqq 0.$$

Now, $p(\tau)$ is a polynomial with at most n real zeros, so for some $\varepsilon > 0$, $p(\tau) > 0$ when $0 < \tau < \varepsilon$. Thus, because we know $f(a)$, $f'(a)$, $\ldots$, and $f^{(n-2)}(a)$ are all nonnegative, we find $(f^{(n-1)}(a) + \tau n!) > 0$ for all sufficiently small $\tau > 0$ and so $f^{(n-1)}(a) \geqq 0$ for all $a > 0$.

If we remove the assumption that $f \in C^\infty(R^+)$ and suppose only that $f \in C(R^+)$, we may proceed as follows. Let $\phi(x)$ be any function such that $\phi \in C_0^\infty(-1, 0)$, $\phi(x) \geqq 0$, and $\int_{-\infty}^\infty \phi(x)\,dx = 1$, and let $\phi_\varepsilon(x) \equiv \phi(x\varepsilon^{-1})$. Then the function $f_\varepsilon(x) \equiv \varepsilon^{-1}\int_{-\infty}^\infty f(t)\phi_\varepsilon(x-t)\,dt \in C^\infty(R^+)$ satisfies the hypotheses of the theorem if f does and $f_\varepsilon(x) \to f(x)$ as $\varepsilon \to 0$. We have just seen that necessarily $f_\varepsilon^{(k)}(x) \geqq 0$ for all $k = 0, 1, \ldots, n-1$, so all the finite differences of f_ε up to order $n-1$ are nonnegative. Now let $\varepsilon \to 0$ and conclude that all the finite differences of f up to order $n-1$ must be nonnegative. But then it is well known [2, p. 497] that necessarily $f \in C^{n-3}(R^+)$ and the theorem follows since all $f^{(k)}$ exist and are nonnegative on R^+, and since $f^{(n-3)}$ is convex and monotone nondecreasing. Q.E.D.

Of particular interest to us is the immediate

COROLLARY 1.3. *Let $0 < \alpha < n-2$, α not an integer. Then there is some $n \times n$ matrix $A > 0$, $A \succeq 0$ for which it is not true that $A^\alpha = (a_{ij}^\alpha)_{i,j=1}^n \succeq 0$.*

This shows that the classical Schur product theorem cannot be generalized to fractional powers even for matrices with positive entries, and shows that those matrices $A \succeq 0$ for which $A^\alpha \succeq 0$ for all $\alpha \geq 0$ are an essential subclass of all the nonnegative semidefinite matrices. In what follows, we wish to be able to treat Hermitian matrices with complex entries as well as positive matrices, and we make the

DEFINITION. Let A be an $n \times n$ complex matrix. If the arguments of the entries can be and have been assigned in such a way that $\arg a_{ij} = -\arg a_{ji}$ for all $1 \leq i$, $j \leq n$, we shall say that the arguments of A have been chosen *consistently*.

Notice that for a consistent choice of arguments of A to be possible, it is necessary that all main diagonal entries a_{ii} be nonnegative. If A is Hermitian and all $a_{ii} \geq 0$, it is always possible to make a consistent choice of the arguments, and if we define $\arg a_{ij}^\alpha \equiv \alpha \arg a_{ij}$, we see that all A^α are again Hermitian with consistent arguments. In the sequel we shall speak about the fractional powers or logarithm of a matrix *only when some consistent choice of arguments can be and has been made*; it is this choice of arguments which is used to define these functions. Thus, we often consider not just complex matrices, but *a matrix together with a certain choice of arguments*. We shall forgo the introduction of a special notation for this extended notion of a matrix since it should always be clear from the context when this notion is required.

We observe that if $A \succeq 0$, then there is always a consistent choice of arguments for A since A is Hermitian and all main diagonal entries a_{ii} are nonnegative. Furthermore, if some $a_{jj} = 0$, then all $a_{jk} = a_{kj} = 0$, $1 \leq k \leq n$. There will be, then, no loss of generality in the sequel when we assume that all $a_{ii} > 0$, since if some $a_{ii} = 0$, A is equivalent (as a quadratic form) to a matrix of lower order with no zero entries on the main diagonal.

DEFINITION. Let A be an $n \times n$ complex matrix. If the arguments of A may be chosen consistently, and if with some (fixed) consistent choice of arguments $A^\alpha \succeq 0$ for all $\alpha \geq 0$, we shall say that, with this choice of arguments, A is *infinitely divisible*.

Observe that if A is infinitely divisible, then in particular $A^1 = A \succeq 0$. In order to state our basic theorem, we shall also require three additional notions:

DEFINITION. If A is an $n \times n$ matrix, then the *incidence matrix of A*, $M(A)$ $\equiv (\mu_{ij})_{i,j=1}^n$, is defined by

$$\mu_{ij} \equiv 0 \quad \text{if } a_{ij} = 0,$$
$$\equiv 1 \quad \text{otherwise.}$$

DEFINITION. $L(M(A)) \equiv \{x \in C^n \mid \langle x, M(A)x \rangle = \sum_{i,j=1}^n x_i \mu_{ij} \bar{x}_j = 0\}$.

DEFINITION. If A has a consistent choice of arguments, then we define $\log^\# A$ $\equiv M(A) \circ \log A = (\mu_{ij} \log a_{ij})_{i,j=1}^n$, where we agree that $0 \log 0 \equiv 0$.

Notice that if all $a_{ij} \neq 0$, then $L(M(A)) = L^n = \{x \in C^n \mid \sum_{i=1}^n x_i = 0\}$. If we understand by a *cone* any subset of a vector space which is closed under multiplication by scalars, then in general $L(M(A)) \subseteq C^n$ is just a cone. But if $M(A) \succeq 0$, then $L(M(A))$

is the null-space of $M(A)$ and so it is actually a subspace of C^n. We can now state the basic

THEOREM 1.4. *Let A be an $n \times n$ matrix. Then A is infinitely divisible if and only if (a) A is Hermitian with $a_{ii} \geqq 0$, (b) $M(A) \succeq 0$, and (c) $\log^{\#} A \succeq 0$ on $L(M(A))$.*

Proof. If A is infinitely divisible, then $A \succeq 0$, so A is Hermitian, $a_{ii} \geqq 0$, and A has a consistent choice of arguments with which each $A^{\alpha} \succeq 0$. Now, $M(A) = \lim_{\alpha \to 0} A^{\alpha}$, so it is clear that $M(A) \succeq 0$. Finally, if $x \in L(M(A))$, and $\alpha > 0$, then

$$0 \leqq \sum_{i,j=1}^{n} x_i \alpha^{-1} a_{ij}^{\alpha} \bar{x}_j = \sum_{i,j=1}^{n} x_i \alpha^{-1} (a_{ij}^{\alpha} - \mu_{ij}) \bar{x}_j \to \sum_{i,j=1}^{n} x_i (\mu_{ij} \log a_{ij}) \bar{x}_j$$

as $\alpha \to 0$, so $\log^{\#} A \succeq 0$ on $L(M(A))$.

Conversely, if A is Hermitian and $a_{ii} \geqq 0$, there is a consistent choice of arguments for A and so $\log^{\#} A$ is well defined. On the unit ball $B_n \equiv \{x \in C^n \mid \langle x, x \rangle = \sum_{i=1}^{n} |x_i|^2 = 1\}$ of C^n, define the two continuous functions $g(x) \equiv \langle x, M(A)x \rangle = \sum_{i,j=1}^{n} x_i \mu_{ij} \bar{x}_j$ and $f_{\alpha,\varepsilon}(x) \equiv \langle x, (\varepsilon + \alpha \log^{\#} A)x \rangle = \sum_{i,j=1}^{n} x_i (\alpha \mu_{ij} \log a_{ij}) \bar{x}_j + \varepsilon$ for any $\alpha, \varepsilon > 0$. By hypothesis, $g(x) \geqq 0$ on the compact set B_n and whenever $g(x) = 0$, $f_{\alpha,\varepsilon}(x) \geqq \varepsilon > 0$. We now recall the elementary

LEMMA 1.5 [5, p. 262]. *Let K be a compact set, and let f, g be two real-valued continuous functions on K. Suppose that $g \geqq 0$ on K and suppose that $f(x) > 0$ if $g(x) = 0$. Then there is an $N_0 \in R^+$ such that $(g + N^{-1}f) > 0$ on K for all $N \geqq N_0$.*

Using this lemma, we see that there is some $N_0 > 0$ such that for all integers $N > N_0$ we have

$$g(x) + N^{-1} f_{\alpha,\varepsilon}(x) = \sum_{i,j=1}^{n} x_i \{\mu_{ij} + N^{-1}(\varepsilon + \alpha \mu_{ij} \log a_{ij})\} \bar{x}_j > 0$$

for all $x \in B_n$. But then the quadratic form generated by $(\mu_{ij} + N^{-1}(\varepsilon + \alpha \mu_{ij} \log a_{ij}))_{i,j=1}^{n}$ is strictly positive definite, so by the Schur product theorem,

$$(\{\mu_{ij} + N^{-1}(\varepsilon + \alpha \mu_{ij} \log a_{ij})\}^N)_{i,j=1}^{n} \gg 0$$

for all integers $N > N_0$. Letting $N \to \infty$, we find

$$(\mu_{ij} \exp \{\varepsilon + \mu_{ij} \alpha \log a_{ij}\})_{i,j=1}^{n} = e^{\varepsilon} M(A) \circ A^{\alpha} = e^{\varepsilon} A^{\alpha} \succeq 0.$$

We conclude that A is infinitely divisible.   Q.E.D.

For completeness, we include the proof of the lemma.

Proof. Let $C \equiv \{x \in K \mid f(x) \leqq 0\} \subset K$. Then C is closed, hence compact, so $0 \geqq -M \equiv \inf \{f(x) \mid x \in C\} > -\infty$ and $m \equiv \inf \{g(x) \mid x \in C\} > 0$. We assume that C is not void, for if it is there is nothing to prove. Let $M' > M$ and take $N > N_0 \equiv M'/m > 0$, so that if $x \notin C$, $(g + N^{-1}f)(x) \geqq N^{-1}f(x) > 0$. If $x \in C$, $(g + N^{-1}f)(x) \geqq m + N^{-1}f(x) \geqq m - MN_0^{-1} = m[1 - M/M'] > 0$.   Q.E.D.

This theorem takes its simplest form in the special case that A has no zeros among its entries, in particular if $A > 0$. Under this condition, the theorem is due to C. Loewner [12, p. 324] and we state it as the

COROLLARY 1.6. *Let A be an $n \times n$ matrix with all $a_{ij} \neq 0$. Then A is infinitely divisible if and only if A is Hermitian, $a_{ii} > 0$, and $\log A \succeq 0$ on L^n.*

REMARK. At the opposite extreme from those matrices $A \succeq 0$ with all $a_{ij} \neq 0$ are those which are strictly diagonal, i.e., $a_{ij} = \delta_{ij}\lambda_j$, $\lambda_j \neq 0$. One sees easily that $M(A)$ is nonsingular if and only if A is strictly diagonal, and in this event $L(M(A)) = \{0\}$. In this case, our theorem is just the obvious fact that if A is diagonal, then $A \succeq 0$ if and only if A is infinitely divisible.

If we refer to any symmetric matrix of zeros and ones as an *incidence matrix* and agree that $M \subseteq M'$ is to mean that the two incidence matrices M and M' satisfy the relation $(M' - M) \geqq 0$, then we can discuss situations intermediate to the two extremes above. If $A \succeq 0$ and if we have the strict inclusion $\{0\} \subset L(M(A)) \subset L^n$, it may be that A fails to be infinitely divisible only because it has zeros in the "wrong places". We make this precise in the

THEOREM 1.7. *Let A be an $n \times n$ Hermitian matrix and let $M' = (\mu'_{ij})_{i,j=1}^n$ be any incidence matrix satisfying $M(A) \subseteq M'$. If $M' \succeq 0$ and if $\log^{\#} A \succeq 0$ on $L(M') = $ null-space of M', then $A' \equiv (\mu'_{ij} \exp(\mu_{ij} \log a_{ij}))_{i,j=1}^n = A + (M' - M(A))$ is infinitely divisible.*

Proof. One proceeds just as in the second part of the proof of Theorem 1.4.

It will be convenient to record the following simple facts also.

PROPOSITION 1.8. *The $n \times n$ matrix A is infinitely divisible if and only if it has a consistent choice of arguments and $A^{\alpha_k} \succeq 0$ for all $k = 1, 2, 3, \ldots$ for some sequence $\{\alpha_k\}_{k=1}^\infty$ such that all $\alpha_k > 0$ and $\alpha_k \to 0$ as $k \to \infty$.*

Proof. The Schur product theorem and a simple limiting argument.

In particular, A is infinitely divisible if and only if $A^{1/k} \succeq 0$ for all $k = 1, 2, 3, \ldots$, with a fixed consistent choice of arguments.

PROPOSITION 1.9. *If A and B are infinitely divisible $n \times n$ matrices, then so are $A \circ B$ and A^γ, $\gamma \geqq 0$.*

DEFINITION. If A is an $n \times n$ matrix, then ΔA is defined to be the $(n-1) \times (n-1)$ matrix $\Delta A \equiv (a_{ij} + a_{i+1,j+1} - a_{i,j+1} - a_{i+1,j})_{i,j=1}^{n-1}$.

Since it is difficult to deal directly with the condition $A \succeq 0$ on L^n, it is helpful to have the

THEOREM 1.10. *Let A be an $n \times n$ matrix. Then $A \succeq 0$ on L^n if and only if $\Delta A \succeq 0$ on V^{n-1}, i.e., $\Delta A \succeq 0$.*

Proof. Partial summation twice leads to the identity

$$\sum_{i,j=1}^n x_i a_{ij} \bar{x}_j = \sum_{i,j=1}^{n-1} X_i (\Delta a)_{ij} \bar{X}_j + 2 \operatorname{Re}\left\{ \bar{X}_n \sum_{i=1}^{n-1} X_i (a_{i,n} - a_{i+1,n}) \right\} + a_{nn} |X_n|^2$$

where $X_k \equiv \sum_{i=1}^k x_i$. But L^n is $(n-1)$-dimensional and $x \in L^n$ if and only if $X_n = 0$. Q.E.D.

Because of condition (b) in Theorem 1.4, we shall find it convenient to have a simple criterion for deciding when a symmetric matrix of zeros and ones, which is the incidence matrix of a nonnegative semidefinite matrix, generates a nonnegative quadratic form. To this end, if $A \geq 0$ is an $n \times n$ matrix, we define *the graph* $\Gamma(M(A))$ *associated with* $M(A) = (\mu_{ij})_{i,j=1}^{n}$ to be the undirected graph of n nodes $\{p_i\}_{i=1}^{n}$ such that there is an arc between p_i and p_j if and only if $\mu_{ij} = 1$. The *irreducible components* of $\Gamma(M(A))$ are the topologically connected components, and an irreducible component is said to be *complete* if each pair of nodes in the component is connected with an arc. A *loop* is an arc which joins a node to itself.

Because $A \geq 0$, we may adopt consistently the convention that the node p_k is missing from the graph $\Gamma(M(A))$ if $a_{kk} = 0$; otherwise, if $a_{kk} \neq 0$, there is a loop at p_k. There will then be no loss of generality if in the sequel we assume that *at each node p_i there is a loop*, for this is equivalent to the assumption that all $a_{ii} > 0$. It is trivial that each graph Γ of this type is of the form $\Gamma = \Gamma(M(A))$ for some (not unique) $A \geq 0$, and the (unique) incidence matrix M for which $\Gamma = \Gamma(M)$ will be called the *incidence matrix of the graph* Γ. We shall say that the nodes p_j and p_k of Γ can be *connected by a path of length l* if there is a sequence $\{p_{i_1} = p_j, p_{i_2}, \ldots, p_{i_l} = p_k\}$ of nodes (not necessarily distinct) such that there is an arc of Γ between p_{i_m} and $p_{i_{m+1}}$ for all $1 \leq m \leq l - 1$.

Now, if $A \geq 0$, we wish to interpret the condition $M(A) \geq 0$ graph-theoretically, so we must observe that if $\Gamma = \Gamma(M(A))$ and if Γ' is the same topological graph as Γ but with the labels (indexing) of the nodes permuted, then M', the incidence matrix of Γ', satisfies $M' \geq 0$ if and only if $M(A) \geq 0$. The truth of this statement is obvious from the relation $M' = PMP^t$, where P is the permutation matrix permuting the labelling of Γ to that of Γ'.

Whether or not $M(A)$ generates a nonnegative quadratic form is therefore a *graph-theoretical property*, i.e., a property only of the topological structure of the graph $\Gamma(M(A))$, and in order to expose this property we introduce the

DEFINITION. A graph Γ has the *three point property* if and only if each pair of nodes which can be connected by a path of length two can in fact be connected by a path of length one.

Let us agree to abbreviate by $p_i R p_j$ the following statement: The nodes p_i and p_j of the graph Γ can be joined by an arc of Γ. Then the three point property is just the assertion that if $p_i R p_j$ and $p_j R p_k$, then $p_i R p_k$.

LEMMA 1.11. *A graph Γ has the three point property if and only if each irreducible component of Γ is complete.*

Proof. If Γ has fewer than three nodes, the lemma is trivial. If p, q, r are distinct nodes of Γ, then our conventions on Γ guarantee quite generally the truth of the two statements

(a) pRp for all $p \in \Gamma$,

(b) pRq implies qRp.

But the three point property is precisely the third statement

(c) pRq and qRr implies pRr,

so R is an equivalence relation on the nodes of any graph Γ with the three point property. The equivalence classes of nodes under R are just the irreducible components of Γ and they are clearly complete. The converse is immediate.   Q.E.D.

LEMMA 1.12. *Let Γ be a graph and let M be its incidence matrix. Then Γ has the three point property if and only if each third order principal submatrix of M generates a nonnegative quadratic form.*

Proof. The only possible third order submatrices which can occur are

$$T_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_1' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad T_1'' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

$$T_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad T_2' = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad T_2'' = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

These submatrices are associated with the respective subgraphs in Figure 1; the matrices T_1, T_1', T_1'', and T_2, T_2', T_2'' are of course equivalent under a permutation. One computes easily that only T_2, T_2', T_2'' do not generate nonnegative quadratic forms, and clearly Γ_2 is the only subgraph without the three point property. Some
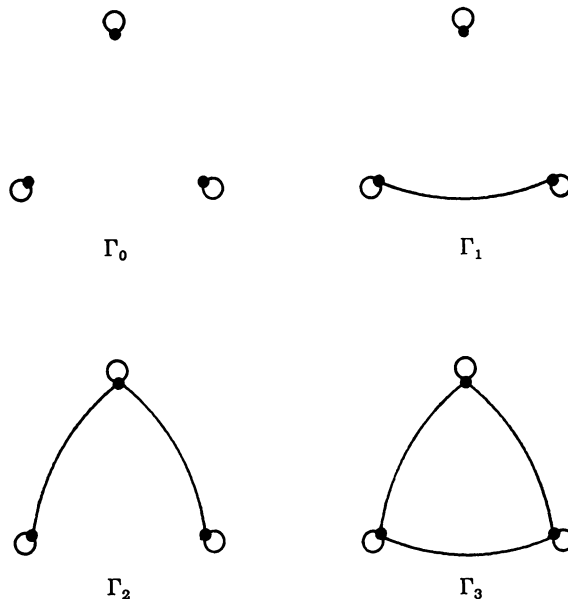

FIGURE 1

third order principal submatrix of the incidence matrix does not generate a non-negative quadratic form if and only if T_2, T_2', or T_2'' occurs. This is the case if and only if the subgraph Γ_2 occurs, while this is equivalent to Γ not having the three point property.   Q.E.D.

We can now prove easily the

THEOREM 1.13. *Let Γ be a graph and let M be its incidence matrix. Then the following four statements are equivalent*:

(a) $M \succeq 0$.

(b) *All third order principal submatrices of M generate nonnegative quadratic forms.*

(c) *Γ has the three point property.*

(d) *Each irreducible component of Γ is complete.*

Proof. That (a) implies (b) is obvious, while (b) implies (c) is half of Lemma 1.12 and (c) implies (d) is half of Lemma 1.11. Finally, if each irreducible component of Γ is complete, then after a permutation M has the form of a direct product of square matrices composed entirely of ones and so $M \succeq 0$.   Q.E.D.

This, then, is the desired graphical property: if $A \succeq 0$ and if one wants to know whether $M(A) \succeq 0$, just construct $\Gamma(M(A))$ and see whether all irreducible components are complete. We see that it also suffices to check *only* the principal third order submatrices of $M(A)$, so that a simple algebraic condition comes out of these graph-theoretical considerations as well. We state these facts for reference as

THEOREM 1.14. *Let A be an infinitely divisible $n \times n$ matrix. Then every irreducible component of $\Gamma(M(A))$ is complete and every third order principal submatrix of $M(A)$ generates a nonnegative quadratic form.*

COROLLARY 1.15. *An $n \times n$ matrix A is infinitely divisible if and only if there exists an $n \times n$ permutation matrix P such that PAP^t is a block diagonal matrix whose diagonal blocks are infinitely divisible matrices with no zeros among their entries.*

REMARKS. Our definition of the three point property may be paraphrased by saying that the graph Γ is *one-connected*. In a natural way, we may define a *k-point property* for each $k \geq 3$ which is equivalent to the graph Γ being $(k-2)$-*connected* and to a natural higher order incidence matrix of Γ being nonnegative semidefinite. Thus, we may compute the *diameter* and other higher order topological properties of a graph in terms of the nonnegativity of certain quadratic forms. We shall not discuss these matters further here.

Let $A \geq 0$, $A \succeq 0$ be an $n \times n$ matrix. One can show by direct calculation that if $n = 2$, then A is always infinitely divisible and that if $n = 3$, then $A^\alpha \succeq 0$ for all $\alpha \geq 1$. Thus, one could conjecture from Corollary 1.3 that in general $A^\alpha \succeq 0$ for all $\alpha \geq n-2$, i.e., the generalization of the Schur product theorem to fractional powers might be true for all $\alpha \geq n-2$. It is easy to prove that if $A \succ 0$, then $A^\alpha \succ 0$ for all *sufficiently large* α (depending on A), so that the conjecture is a sharpened form of this statement with a precise lower bound.

Finally, we remark that infinite divisibility of a matrix can be defined in such a way that the problem of choosing arguments is avoided and the analogy with probability is made more transparent. One can show easily that A is infinitely divisible if and only if for some sequence of integers $\{n_k\}_{k=0}^\infty$ there exist matrices $A_k \succeq 0$ such that $A_k^{n_k} = A$ for all $k = 1, 2, 3, \ldots$ and $A_k \to M(A)$ as $k \to \infty$.

2. **Continuous kernels.** We shall now develop the continuous analogues of the preceding statements about discrete quadratic forms. Let $K(P, Q)$ be a complex valued function defined on $S \times S$ for some set S. We shall say that, over S, $K(P, Q) \succeq 0$ on C^n, $K(P, Q) \succeq 0$ on L^n, or $K(P, Q)$ is infinitely divisible on C^n if and only if for every choice of n points $\{P_i\}_{i=1}^n \subseteq S$ the matrix $(K(P_i, P_j))_{i,j=1}^n$ has the respective property. In particular, if $K(P, Q) \succeq 0$ on C^2, then $K(P, P) \geqq 0$ and $K(P, Q)$ is *Hermitian*, i.e., $K(P, Q) = \bar{K}(Q, P)$.

Now suppose $K(P, Q) \in C(S \times S)$ where *we shall always assume hereafter* that $S \subseteq R^m$ for some $m = 1, 2, \ldots$ is as smooth as necessary (usually, just open or even measurable) and that S is connected, although it would suffice for nearly all our considerations to let S be a countably compact measurable connected Hausdorff space carrying a regular Borel measure dP which assigns positive measure to every nonempty open set. We shall say $K(P, Q) \succeq 0$ on $C_0(S)$ (respectively, *on* $L_0(S)$) if and only if $K(P, Q) \succeq 0$ on C^n (respectively, L^n) for all $n = 1, 2, \ldots$, and we shall always understand $K(P, Q) \succeq 0$ to mean $K(P, Q) \succeq 0$ *on* $C_0(S)$.

If $K(P, Q) \in C(S \times S)$ and if a continuous branch of the argument function arg $K(P, Q)$ can be and has been defined on $S \times S$ with arg $K(P, Q) = -$ arg $K(Q, P)$, we shall say that the argument of $K(P, Q)$ has been chosen *consistently*. If with some (fixed) consistent choice of the argument $K(P, Q)$ is infinitely divisible on C^n for all $n = 1, 2, 3, \ldots$, then we shall say that with this choice of argument $K(P, Q)$ *is infinitely divisible* (*on* $C_0(S)$). When these strong hypotheses on arg $K(P, Q)$ are satisfied, the functions $K^\alpha(P, Q)$ and log $K(P, Q)$ will be defined correspondingly. We observe that $K(P, Q)$ always has a consistent argument if $K(P, Q) \geqq 0$ on $S \times S$, but that in general it is *not* sufficient for the existence of a consistent argument that $K(P, Q)$ be Hermitian with $K(P, P) \geqq 0$ for all $P \in S$.

Many theorems about continuous kernels are obtained by limiting processes from the corresponding matrix theorems, and we state for easy reference the

THEOREM 2.1. *Let* $K(P, Q)$, $K_1(P, Q)$, $K_2(P, Q) \in C(S \times S)$. (a) $K(P, Q) \succeq 0$ *on* $C_0(S)$ *(respectively, on* $L_0(S)$*) if and only if* $\iint_{S \times S} f(P) K(P, Q) \bar{f}(Q) \, dP \, dQ \geqq 0$ *for all* $f \in C_0(S)$ *(respectively, all* $f \in L_0(S)$*)*; (b) *If* $K_1(P, Q) \succeq 0$ *and* $K_2(P, Q) \succeq 0$, *then* $H(P, Q) \equiv K_1(P, Q) K_2(P, Q) \succeq 0$ *and* $\alpha K_1(P, Q) + \beta K_2(P, Q) \succeq 0$ *for* $\alpha, \beta \geqq 0$; *and* (c) *If* $K_1(P, Q)$ *and* $K_2(P, Q)$ *are infinitely divisible, then so is* $H_{\alpha\beta}(P, Q) \equiv K_1^\alpha(P, Q) \times K_2^\beta(P, Q)$ *for all* $\alpha, \beta \geqq 0$.

Proof. Simple limiting arguments, Theorem 1.1, and Proposition 1.9.

COROLLARY 2.2. *If* $K(P, Q)$ *is a continuous kernel on* $S \times S$, *then* $K(P, Q)$ *is infinitely divisible if and only if the argument of* $K(P, Q)$ *can be chosen consistently*

on $S \times S$ and with this choice of argument, $\iint_{S \times S} f(P)K^\alpha(P, Q)\bar{f}(Q)\, dP\, dQ \geqq 0$ for all $\alpha > 0$ and all $f \in C_0(S)$.

Proof. Part (a) of the theorem and the definition of infinite divisibility.

Many theorems, however, can be given a much sharper form in their continuous versions, and a good example is the continuous analogue of Theorem 1.2.

THEOREM 2.3. *Let $f \in C(\mathbf{R}^+)$, $S \subseteq \mathbf{R}$. In order that $f(K(x, y)) \succeq 0$ whenever $K(x, y) > 0$, $K(x, y) \succeq 0$, $K(x, y) \in C(S \times S)$, it is both necessary and sufficient that $f(x) \in C^\infty(\mathbf{R}^+)$ and $f^{(k)}(x) \geqq 0$ for all $x \in \mathbf{R}^+$ and all $k = 0, 1, 2, \ldots$, i.e., f is absolutely monotonic.*

Proof. Let $K(x, y) \equiv \rho e^{x+y}$, $\rho > 0$, so that $K(x, y) \succeq 0$, $K(x, y) > 0$, and $K(x, y)$ is continuous. Let $g(x) \equiv f(\rho e^x)$, so that $g(x)$ is continuous on S and $g(x+y) = f(K(x, y)) \succeq 0$. But it is known [16, p. 275] that $g(x)$ must then be analytic (and even possess a representation $g(z) = \int_{-\infty}^\infty e^{-zt} d\mu(t)$, $d\mu \geqq 0$, convergent for Re $(z) \in S$), so because $\rho > 0$ is arbitrary, we see that $f(x)$ must actually be analytic for all $x \in \mathbf{R}^+$. That all derivatives $f^{(k)}(x) \geqq 0$ for $x \in \mathbf{R}^+$, $k = 0, 1, 2, \ldots$ follows now immediately from Theorem 1.2, a limiting argument, and the following

LEMMA 2.4. *Let $A > 0$ be an $n \times n$ matrix and $S \subseteq \mathbf{R}$. In order that $A \succeq 0$, it is necessary and sufficient that there exist n points $\{x_i\}_{i=1}^n \subset S$ and a continuous kernel $K(x, y) > 0$, $K(x, y) \succeq 0$ such that $K(x_i, x_j) = a_{ij}$.*

We postpone the proof of the lemma and prove now the sufficiency of the condition. If $f \in C^\infty(\mathbf{R})$ and $f^{(k)}(x) \geqq 0$, let $K(x, y) \succeq 0$ be any continuous non-negative kernel on $S \times S$ and let $\phi(x) \in C_0(S)$ have compact support contained in a compact interval $D' \subset S$. Then $K(D', D') \equiv D$ is a compact interval on which we may choose a sequence $\{p_n(x)\}_{n=1}^\infty$ of polynomials with nonnegative coefficients such that the $p_n(x)$ converge uniformly to $f(x)$ on D. This is because we may take for $p_n(x)$ the Bernstein polynomial $B_n(f; x)$ of degree n on D whose coefficients, being formed of positive multiples of finite differences of $f(x)$, are all nonnegative as a consequence of the inequalities $f^{(k)}(x) \geqq 0$. But then $p_n(K(x, y)) \succeq 0$ by Theorem 2.1(b), so

$$\iint_{S \times S} \phi(x) f(K(x, y)) \bar{\phi}(y)\, dx\, dy = \lim_{n \to \infty} \iint_{S \times S} \phi(x) p_n(K(x, y)) \bar{\phi}(y)\, dx\, dy \geqq 0$$

and we see that $f(K(x, y)) \succeq 0$.   Q.E.D.

Proof of the lemma. If $A \succeq 0$, then from the canonical decomposition $A = U \Lambda U^*$ with $U = (u_{ij})_{i,j=1}^n$ unitary and $\Lambda = (\lambda_i \delta_{ij})_{i,j=1}^n$ with $\lambda_i \geqq 0$, we see that $A = (a_{ij})_{i,j=1}^n$ has a representation $a_{ij} = \langle v^i, v^j \rangle = \sum_{k=1}^n v_k^i \bar{v}_k^j$ as a Gram matrix, where $v^i \equiv (\lambda_k^{1/2} \bar{u}_{ik})_{k=1}^n$. If $A > 0$, then $\langle v^i, v^j \rangle > 0$ as well.

Choose $x_0 \in S$ and $h \in \mathbf{R}^+$ such that $\{x_k\}_{k=0}^n \equiv \{x_0 + kh\}_{k=0}^n \subset S$, define the continuous vector valued function

$$v(t) \equiv (k-t)v^{k-1} + (t-k+1)v^k \quad \text{for } k-1 \leq t \leq k, \, 2 \leq k \leq n,$$
$$\equiv v^1 \qquad\qquad\qquad\qquad \text{for } t \leq 1,$$
$$\equiv v^n \qquad\qquad\qquad\qquad \text{for } t \geq n,$$

and let $K(x, y) \equiv \langle v((x-x_0)/h), v((y-x_0)/h) \rangle$ for $x, y \in S$. Then $K(x, y) > 0$, $K(x, y) \succeq 0$, $K(x_i, x_j) = a_{ij}$, and $K(x, y)$ is continuous. The converse follows immediately from a limiting argument.   Q.E.D.

This result shows that a generalization to fractional powers of the Schur product theorem for continuous kernels (Theorem 2.1(b)) is false, for there is the immediate

COROLLARY 2.5. *Let $S \subset \mathbf{R}$, $\alpha > 0$, α not an integer. Then there is defined on $S \times S$ some continuous kernel $K(x, y) \succeq 0$, $K(x, y) > 0$ for which it is not true that $K^\alpha(x, y) \succeq 0$.*

The class of infinitely divisible continuous kernels is then a proper subset of the class of continuous kernels which generate nonnegative quadratic forms. Just as in the matrix case, we must pay special attention to the location of zeros of these kernels, and we find again that they can occur only in very special configurations. If $K(P, Q) \succeq 0$, then a limiting argument from the matrix case shows easily that $K(P_0, P_0) = 0$ for some $P_0 \in S$ implies $K(P_0, S) \equiv K(S, K_0) \equiv 0$. The following theorem says that implications of a similar character are valid even for off-diagonal zeros when $K(P, Q)$ is infinitely divisible. If $C \subset S$, denote the closure of C by $\text{Cl}(C)$, denote by $\varnothing$ the void set, and agree that $K(\varnothing, C) \equiv 0$.

THEOREM 2.6. *Let $K(P, Q) \in C(S \times S)$ be infinitely divisible. Suppose $K(P_0, Q_0) = 0$ for some $P_0, Q_0 \in S$ and let $C_1 \equiv \{P \in S \mid K(P_0, P) = 0\}$, $C_2 = \{P \in S \mid K(P, Q_0) = 0\}$. Then*

(a) *$K(P_0, P)K(P, Q_0) \equiv 0$ for all $P \in S$.*

(b) *The sets C_1 and C_2 are nonvoid, closed, and satisfy $C_1 \cup C_2 = S$.*

(c) *We have $0 \equiv K(P_0, C_1) \equiv K(C_2, Q_0) \equiv K(\text{Cl}(S - C_1), C_1) \equiv K(C_2, \text{Cl}(S - C_2))$.*

(d) *If C_1 is nowhere dense in S (in particular, if $C_1 = \{Q_0\}$), then $K(S, Q_0) \equiv K(S, C_1) \equiv 0$. Similarly for C_2.*

(e) *If $C_1 \neq S$, then there is an open neighborhood $N \subset S$, $P_0 \in N$, with $N \cap C_1 = \varnothing$ and $N \subset C_2$. Similarly for C_2.*

(f) *There exists some $\hat{P} \in S$ for which $K(\hat{P}, \hat{P}) = 0$.*

Proof. If $P_0 = Q_0$, then each statement is either vacuous or trivial, so suppose $P_0 \neq Q_0$.

(a) Let $P \in S$. A limiting argument together with Theorem 1.14 shows that the incidence matrix

$$M \begin{pmatrix} K(P_0, P_0) & K(P_0, P) & K(P_0, Q_0) \\ K(P, P_0) & K(P, P) & K(P, Q_0) \\ K(Q_0, P_0) & K(Q_0, P) & K(Q_0, Q_0) \end{pmatrix} \equiv \begin{pmatrix} \mu_{11} & \mu_{12} & 0 \\ \mu_{12} & \mu_{22} & \mu_{23} \\ 0 & \mu_{23} & \mu_{33} \end{pmatrix}$$

cannot have $\mu_{12}=\mu_{23}=1$, for then of necessity also $\mu_{11}=\mu_{22}=\mu_{33}=1$ and the incidence matrix would be of a form which is forbidden by infinite divisibility. Thus, $\mu_{12}\mu_{23}=0$, so $K(P_0, P)K(P, Q_0)\equiv 0$ for all $P \in S$.

(b) This is immediate from continuity of $K(P, Q)$, the definitions of C_1 and C_2, and part (a).

If $C_1=S$, then there is nothing to prove in parts (c)–(f), so suppose $C_1 \neq S$ in the following. We omit the similar considerations for C_2.

(c) The equation $K(\tilde{P}, P)K(P, \tilde{Q})\equiv 0$ for all $P \in S$ must hold for any pair of points $(\tilde{P}, \tilde{Q})$ such that $K(\tilde{P}, \tilde{Q})=0$ by part (a). Now, $K(P_0, C_1)\equiv 0$ by definition, so $K(P_0, P)K(P, C_1)\equiv 0$ for all $P \in S$ and as $K(P_0, S-C_1)\neq 0$, we conclude that $K(S-C_1, C_1)\equiv 0 \equiv K(\mathrm{Cl}\,(S-C_1), C_1)$ by continuity.

(d) Since $K(S-C_1, C_1)\equiv 0$, $K(\mathrm{Cl}\,(S-C_1), C_1)\equiv 0 \equiv K(S, C_1)$ if C_1 is nowhere dense in S.

(e) Recall $P_0 \in C_2$ and assume $C_1 \neq S$. Then necessarily $P_0 \notin C_1$, so $K(P_0, P_0)\neq 0$. But $K(C_2, \mathrm{Cl}\,(S-C_2))\equiv 0$ implies $K(P_0, \mathrm{Cl}\,(S-C_2))\equiv 0$ so that $P_0 \notin \mathrm{Cl}\,(S-C_2)$. Choose the neighborhood N of P_0 to be disjoint from each of the closed sets C_1 and $\mathrm{Cl}\,(S-C_2)$ for the conclusion.

(f) Since S is connected, the nonvoid closed set $\mathrm{Cl}\,(S-C_1)$ cannot be disjoint from C_1, so there is some point $\hat{P} \in C_1 \cap \mathrm{Cl}\,(S-C_1)\neq \varnothing$, and by part (c), $K(\hat{P}, \hat{P})=0$.   Q.E.D.

We should observe that the infinitely divisible kernel on $S=[-1, 1]\subset R$ defined by

$$K(x, y) \equiv xy \quad \text{if } xy \geq 0,$$
$$\equiv 0 \quad \text{if } xy < 0$$

actually exhibits nontrivially the phenomena described above and shows that the hypothesis in (e) is necessary.

We see that although a zero of an infinitely divisible kernel need not be contained in a full open neighborhood of zeros in $S \times S$, it is always contained in a *one-dimensional* open neighborhood in at least one of the coordinate spaces. Furthermore, if there is any zero at all, then there must be a pair of coordinate *lines* of zeros somewhere. In particular, we have the immediate

COROLLARY 2.7. *Let* $K(P, Q) \in C(S\times S)$ *be infinitely divisible. If* $K(P, P)\neq 0$ *for all* $P \in S$, *then* $K(P, Q)\neq 0$ *for all* $P, Q \in S$.

The following analogue of Corollary 1.6 is the most useful result characterizing infinitely divisible kernels.

THEOREM 2.8. *Let* $K(P, Q) \in C(S\times S)$ *with* $K(P, P)>0$ *for all* $P \in S$. $K(P, Q)$ *is infinitely divisible if and only if* (a) $K(P, Q)$ *has a consistent choice of argument,* (b) $K(P, Q)=\bar{K}(Q, P)\neq 0$ *for all* $P, Q \in S$, *and* (c) $H(P, Q)\equiv \log K(P, Q)\succeq 0$ *on* $L_0(S)$.

Proof. If $K(P, Q)$ is infinitely divisible, then by definition it has a consistent choice of argument and $K(P, Q) \geq 0$, so $K(P, Q) = \bar{K}(Q, P)$. We have just seen that if, in addition, $K(P, P) \neq 0$, then $K(P, Q) \neq 0$, so $H(P, Q)$ is well defined and continuous and only (c) remains to be proved. But $K(P, Q)$ is infinitely divisible on C^n for every $n = 1, 2, \ldots$, so by Corollary 1.6, $H(P, Q) \geq 0$ on L^n for every $n = 1, 2, \ldots$, and so $H(P, Q) \geq 0$ on $L_0(S)$.

Conversely, if (a) and (b) are satisfied, then $H(P, Q)$ is well defined and continuous, so $H(P, Q) \geq 0$ on L^n for every $n = 1, 2, \ldots$. But then by Corollary 1.6, $K^\alpha(P, Q) = \exp\{\alpha H(P, Q)\} \geq 0$ on C^n for every $n = 1, 2, \ldots$ and every $\alpha \geq 0$ so $K(P, Q)$ is infinitely divisible.   Q.E.D.

The converse can also be obtained from the following proposition, which can often be of independent use itself.

PROPOSITION 2.9 (CF. [3, p. 649]). *Let* $H(P, Q) \in C(S \times S)$. *Then* $H(P, Q)$ *is Hermitian and* $H(P, Q) \geq 0$ *on* $L_0(S)$ *if and only if for each compact subset* $C \subseteq S$ *there is a continuous kernel* $K_C(P, Q) \geq 0$ *on* $C_0(C)$ *and a continuous function* $f_C(P)$ *on* C *such that* $H(P, Q) = K_C(P, Q) + f_C(P) + \bar{f}_C(Q)$ *for* $P, Q \in C$.

Proof. If C has measure zero, then the assertion is trivial, so we may assume

$$0 < m \equiv \int_C dP < \infty.$$

Define

$$f_C(P) \equiv \frac{1}{m} \int_C H(P, Q) \, dQ - \frac{1}{2m^2} \iint_{C \times C} H(P, Q) \, dP \, dQ,$$

set $K_C(P, Q) \equiv H(P, Q) - f_C(P) - \bar{f}_C(Q)$, and check that $K_C(P, Q) \geq 0$ on $C_0(C)$. The converse is immediate.   Q.E.D.

There are other simple necessary and sufficient conditions of this same type. For example, if $H(P, Q) \in C(S \times S)$, if P_0 is an arbitrary point of S, and if $H_0(P, Q) \equiv H(P, Q) - H(P, P_0) - H(P_0, Q) + H(P_0, P_0)$, then one can show easily that $H(P, Q) \geq 0$ on L_0 if and only if $H_0(P, Q) \geq 0$.

Theorem 2.6 shows that an infinitely divisible kernel can have zeros only in certain configurations which may be classified [8, pp. 28–30] for the purpose of generalizing Theorem 2.8, but in any particular case the essential condition guaranteeing infinite divisibility of $K(P, Q)$ is, loosely, that $\log K(P, Q) \geq 0$ on L_0. The above proposition provides one criterion for this, but it is helpful to have tractable conditions in certain special cases, and we give some here.

LEMMA 2.10. *Let* $S \subset \mathbf{R}$. *Then* $f \in L_0(S)$ *if and only if* $f(x) = dF(x)/dx$ *for some* $F \in C_0^1(S)$.

Proof. Just set $F(x) \equiv \int_{-\infty}^x f(t) \, dt$ using the definitions of L_0 and C_0.

THEOREM 2.11. *Let* $H(x, y) \in C^2(S \times S)$, $S \subseteq \mathbf{R}$. *Then* $H(x, y) \geq 0$ *on* $L_0(S)$ *if and only if* $K(x, y) = \partial^2 H(x, y)/\partial x \, \partial y \geq 0$ *on* $C_0(S)$. *In particular, if* $H(x, y) \geq 0$, *then* $\partial^2 H(x, y)/\partial x \, \partial y \geq 0$ *and* $\partial^2 H(x, y)/\partial x \, \partial y|_{y=x} \geq 0$.

Proof. This follows from the use of integration by parts, the preceding lemma, and the fact that $C_0^1(S)$ is uniformly dense in $C_0(S)$.

THEOREM 2.12. *Let $H(x, y) \in C(R \times R)$. Then $H(x, y) \succeq 0$ on $L_0(R)$ if and only if for every $h \in R$,*

$$K_h(x, y) \equiv H(x, y) + H(x+h, y+h) - H(x, y+h) - H(x+h, y) \succeq 0$$

on $C_0(R)$.

Proof. If $f \in C_0(R)$, then $[f(x) - f(x-h)] \equiv f_h(x) \in L_0(R)$ for every $h \in R$. Then

$$0 \leq \iint_{R \times R} f_h(x) H(x, y) \bar{f}_h(y) \, dx \, dy = \iint_{R \times R} f(x) K_h(x, y) \bar{f}(y) \, dx \, dy,$$

so $K_h(x, y) \succeq 0$. Conversely, if $f \in L_0(R)$, then for some $F \in C_0^1(R)$,

$$f(x) = \lim_{h \to 0} \frac{1}{h} [F(x) - F(x-h)] = \lim_{h \to 0} \frac{1}{h} F_h(x).$$

But then

$$0 \leq \frac{1}{h^2} \iint_{R \times R} F(x) K_h(x, y) \bar{F}(y) \, dx \, dy$$

$$= \iint_{R \times R} \left\{ \frac{1}{h} F_h(x) \right\} H(x, y) \left\{ \frac{1}{h} \bar{F}_h(y) \right\} dx \, dy \to \iint_{R \times R} f(x) H(x, y) \bar{f}(y) \, dx \, dy$$

as $h \to 0$, so $H(x, y) \succeq 0$ on $L_0(R)$.   Q.E.D.

Similar results are valid in R^n and depend on the

LEMMA 2.13. *Let $S \subseteq R^m$ be an open (always assumed connected) set. Then $f \in L_0(S)$ if and only if*

$$f(x) = \operatorname{div} F(x) = \sum_{k=1}^{m} \frac{\partial F_k}{\partial x^k}(x)$$

for some $F(x) \in C_0^1(S, C^m)$.

Proof. This is a standard result on m-dimensional compact cohomology proved by an elementary but tedious induction on m starting with Lemma 2.10 and a partition of unity argument [15, p. 126, Lemma 4.2].

THEOREM 2.14. *Let $H(x, y) \in C^2(S \times S)$ for some open $S \subseteq R^m$. Then $H(x, y) \succeq 0$ on $L_0(S)$ if and only if*

$$(2.A) \qquad \sum_{i,j=1}^{m} \iint_{S \times S} F_i(x) \frac{\partial^2}{\partial x^i \, \partial y^j} H(x, y) \bar{F}_j(y) \, dx \, dy \geq 0$$

for all $F(x) = (F_i(x))_{i=1}^m \in C_0(S, C^m)$.

Proof. This follows easily by the use of integration by parts (the divergence theorem), the preceding lemma, and the fact that $C_0^1(S, C^m)$ is uniformly dense in $C_0(S, C^m)$.

If we agree that $F(x+he_j) \equiv F(x_1, x_2, \ldots, x_i+h, x_{i+1}, \ldots, x_m)$ whenever $h \in \mathbf{R}$ and $x \in \mathbf{R}^m$, then we have the

THEOREM 2.15. *Let* $H(x, y) \in C(\mathbf{R}^m \times \mathbf{R}^m)$. *Then* $H(x, y) \succeq 0$ *on* $L_0(\mathbf{R}^m)$ *if and only if for every* $h \in \mathbf{R}$ *and every* $F \in C_0(\mathbf{R}^m, \mathbf{C}^m)$,

$$\sum_{i,j=1}^{m} \iint_{\mathbf{R}^m \times \mathbf{R}^m} F_i(x)\{H(x, y) + H(x+he_i, y+he_j)$$
$$- H(x+he_i, y) - H(x, y+he_j)\}\bar{F}_j(x) \, dx \, dy \geqq 0.$$

Proof. Proceed as in Theorem 2.12 using the above two results analogously. Finally, we wish to state Theorem 2.14 in an important special case:

THEOREM 2.16. *Let* $H(z, \zeta)$ *be analytic in* z, *antianalytic in* ζ *for* $z, \zeta \in S \subseteq \mathbf{C}$, S *open. Then* $H(z, \zeta) \succeq 0$ *on* $L_0(S)$ *if and only if*

$$K(z, \zeta) \equiv \partial^2 H(z, \zeta)/\partial z \, \partial \zeta \succeq 0$$

on $C_0(S)$.

Proof. Let $d\sigma(z)$ be the Lebesgue measure in the plane, so that the necessary and sufficient condition (2.A) becomes in this case

$$0 \leqq \sum_{i,j=1}^{2} \iint_{S \times S} F_i(z) \frac{\partial^2}{\partial z_i \, \partial \bar{\zeta}_j} H(z, \zeta) \bar{F}_j(\zeta) \, d\sigma(z) \, d\sigma(\zeta)$$

$$= \iint_{S \times S} (F_1(z) + iF_2(z)) \frac{\partial^2}{\partial z \, \partial \bar{\zeta}} H(z, \zeta)(\bar{F}_1(\zeta) - i\bar{F}_2(\zeta)) \, d\sigma(z) \, d\sigma(\zeta) \geqq 0.$$

This is clearly the case if and only if $\partial^2 H(z, \zeta)/\partial z \, \partial \zeta \succeq 0$ on $C_0(S)$.   Q.E.D.

3. **Applications.** The general theory presented above may be viewed as a generalization of the classical theory of the *infinitely divisible characteristic functions* of probability theory which are precisely the infinitely divisible kernels of the form $K(x, y) \equiv \phi(x-y)$ for some $\phi: \mathbf{R} \to \mathbf{C}$ continuous at zero such that $\phi(0) = 1$. Using the above methods one can characterize these functions easily and derive in an elementary fashion the important Kolmogorov and Levy-Khinchine representation formulae. The same methods can also be used to derive parallel results about the *infinitely divisible completely monotonic functions*, i.e., those functions $\phi \in C^1(\mathbf{R}^+)$ such that $\phi^\alpha(x)$ is a completely monotonic function for all $\alpha > 0$. The function ϕ will have this property if and only if $K_1(x, y) \equiv \phi(x+y)$ is infinitely divisible and $K_2(x, y) \equiv -\phi'(x+y)\phi^{-1}(x+y) \succeq 0$. These difference and sum kernels are discussed in [7], or more fully in [8, pp. 63–93].

The discrete analogues of these continuous kernels are the *infinitely divisible Toeplitz and Hankel matrices*, i.e., matrices $A = (a_{ij})_{i,j=0}^{\infty}$ such that $a_{ij} = a_{i-j}$ or $a_{ij} = b_{i+j}$, respectively, for some sequences of complex numbers $\{a_n\}_{n=-\infty}^{\infty}$, or $\{b_n\}_{n=0}^{\infty}$. Our methods have been used [9] to characterize such sequences and to derive explicit representation formulae for them as trigonometric and algebraic moment sequences. One can solve the natural interpolation problems for these sequences,

and there are applications to the theory of analytic functions with positive real part in the unit disc.

We have already mentioned the recent work of C. Loewner [13] in which he showed that the *Green's function* for Laplace's equation in the unit ball of R^n, $n \geq 2$, is an infinitely divisible kernel, but it is not yet known whether there is a similar generalization of the fact [8, pp. 107–109] that the Green's function of any Sturm-Liouville differential operator $L[u] = (pu')' - qu$, $u(0) = u(1) = 0$, on the unit interval is always infinitely divisible if p and q are positive functions such that $p \in C^1[0, 1]$, $q \in C[0, 1]$.

One can find infinitely divisible kernels among the many *reproducing kernel functions* of families of orthonormal functions, e.g., the Bergman and Szegö kernel functions and the reproducing kernel for the Hermite functions. There is also an analogue of the related theorem of Loève [10, p. 466] that every positive semidefinite kernel $K(x, y)$ is the *covariance kernel* of some second order, normal, mean zero stochastic process $Z(x)$, i.e., $K(x, y) = E(Z(x)Z(y))$. If $K(x, y)$ is a continuous, positive infinitely divisible kernel over [0, 1], then $K(x, y) = E(e^{Z(x)}e^{Z(y)})$ for some second order normal real-valued (but not necessarily mean zero) stochastic process $Z(x)$ and, conversely, every kernel of this form is infinitely divisible [8, pp. 110–112].

Finally, returning to the work of C. Loewner which was the germ of this whole investigation, we recall [11] that a real-valued function $f \in C^1(0, 1)$ is the boundary value of an analytic function mapping the upper half plane $\{z \mid \text{Im} (z) \geq 0\}$ into itself if and only if

$$K_f(x, y) \equiv (f(x) - f(y))/(x - y) \geq 0 \quad \text{on } C_0(0, 1).$$

However, it is now known [4], [6] that f will be the boundary value of a *univalent* (*schlicht*) mapping if and only if f is not constant and $K_f(x, y)$ is infinitely divisible over (0, 1). Some other results of this character can be found in [8, pp. 94–106], where it is shown that if f is a univalent analytic function in the finite plane with the two real slits $(-\infty, 0]$ and $[1, \infty)$ removed, then the kernel

$$H_f(x, y) \equiv |(f(x) - f(y))/(x - y)|$$

is infinitely divisible over (0, 1). We also show that if f is a univalent analytic mapping of the unit disc $UD \equiv \{z \in C \mid |z| < 1\}$ into itself, then the kernel

$$P_f(z, \zeta) \equiv (1 - f(z)\bar{f}(\zeta))/(1 - z\bar{\zeta})$$

is infinitely divisible over UD. These last two results both follow from the classical Grunsky inequalities with our methods.

BIBLIOGRAPHY

1. R. Bellman, *Introduction to matrix analysis*, McGraw-Hill, New York, 1960.

2. R. P. Boas and D. V. Widder, *Functions with positive differences*, Duke Math. J. **7** (1940), 496–503.

3. S. Bochner, *Hilbert distance and positive definite functions*, Ann. of Math. (2) **42** (1941), 647–656.

4. C. H. FitzGerald, *On analytic continuation to a schlicht function*, Proc. Amer. Math. Soc. **18** (1968), 788–792.

5. R. C. Gunning and H. Rossi, *Analytic functions of several complex variables*, Prentice-Hall, Englewood Cliffs, N. J., 1965.

6. R. A. Horn, *On boundary values of a schlicht mapping*, Proc. Amer. Math. Soc. **18** (1968), 782–787.

7. ———, *On infinitely divisible matrices, kernels, and functions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **8** (1967), 219–230.

8. ———, *Infinitely divisible matrices, kernels, and functions*, Ph.D. dissertation, Stanford Univ., Stanford, Calif., 1967.

9. ———, *Infinitely divisible positive definite sequences*, Trans. Amer. Math. Soc. **136** (1969), 287–303.

10. M. Loève, *Probability theory*, 3rd ed., Van Nostrand, Princeton, N. J., 1963.

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

12. C. Loewner, *On Schlicht-monotonic functions of higher order*, J. Math. Anal. Appl. **14** (1966), 320–325.

13. ———, *Determination of the critical exponent of the Green's function*, Symposium on Function Theory, Erevan, USSR, 1965 (1966).

14. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140** (1911), 1–29.

15. S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, Englewood Cliffs, N. J., 1964.

16. D. V. Widder, *The Laplace transform*, Princeton Univ. Press, Princeton, N. J., 1946.

STANFORD UNIVERSITY,
   STANFORD, CALIFORNIA