PREDICTION THEORY AND FOURIER SERIES IN SEVERAL VARIABLES

BY

HENRY HELSON

Berkeley, California

AND

DAVID LOWDENSLAGER (1)

Urbana, Illinois

1. Introduction

The theory of analytic functions of a complex variable extends only with difficulty and incompletely to functions of several variables. Because the Riemann Mapping Theorem fails in several variables, the description of domains of holomorphy and their analytic transformations has been a major concern. Nevertheless function theory in the bicylinder hardly exists beside the elegant theory of functions in the unit circle. This circumstance is related to the singular fact, never observed so far as we know, that analytic function theory divides into two distinct disciplines in higher dimensions. The theory of analytic functions in several variables has been concerned with functions defined locally and consistently by power series in a domain, whereas much function theory in the circle can be made to depend on group properties of the circle, and generalizes in quite a different way. The study of multiple Fourier series from this point of view is one objective of this paper. The discussion of analyticity in a group-theoretic context was begun by Mackey [13], and recently has been continued with great ingenuity by Arens and Singer [2, 3, 4]. While our work has points of contact with that of Arens and Singer, the methods are different, and we have attained a certain completeness at the expense of generality.

(1) The work of the second-named author was supported by contract Nonr-222 (37) with the Office of Naval Research.

In the work of Kolmogoroff [11] and Wiener [18] on the prediction of second-order stationary stochastic processes, certain theorems about analytic functions in the circle play an essential part. The analytic difficulty is exactly met by a theorem of Szegö; and indeed Szegö's Theorem can be used to prove the various function-theoretic results which would otherwise be used in the proof of the prediction theorem. The second section of this paper is devoted mainly to a generalization of Szegö's Theorem to two or more variables, and this furnishes the solution of a certain prediction problem in several variables. This is not the multiple prediction problem mentioned by Doob [8, p. 594] and treated recently by Wiener [19], of which we shall speak presently.

In the third section we exploit the methods and results of the second section in order to prove a number of theorems in multiple Fourier series generalizing elementary properties of analytic functions of one variable. We obtain an inequality in place of Jensen's Formula, under hypotheses slightly different from those of Arens in a paper not yet published. Then we extend the characterization of functions $w(e^{ix})$ defined on the unit circle having a representation

$$w = |f|^2 \quad \text{almost everywhere,}$$

where f is analytic and of class H^2 inside the circle. A related theorem states that every function analytic in the circle and of class H is the product of two functions in H^2. Finally we extend the theorem of Hardy and Littlewood about functions of class H in the circle:

If
$$f(z) = \sum_0^\infty a_n z^n,$$
then
$$\sum_0^\infty |a_n|/(n+1) < \infty.$$

For simplicity we treat functions of only two variables in this section. In each case the class of functions to which our theorem applies is not the double power series, but rather the functions defined on the torus whose Fourier coefficients a_{mn} vanish for all (m, n) belonging to a *half-plane* (in a sense which must be made precise). The proofs depend on this division of the group of lattice points into disjoint semi-groups, rather than on the local properties of functions defined on the torus. For functions of one variable the theorems are generally proved by removing the zeros of an analytic function in the circle. Of course this technique is not available for functions of several variables, and instead our method depends on the fact that every closed convex set in Hilbert space possesses a unique element of minimal norm.

In section four we discuss Bochner's generalization of a well-known theorem of

F. and M. Riesz [15]. In the form of interest to us, the Riesz Theorem states: *if* $\mu(x)$ *is a complex function of bounded variation on the circle such that*

$$\int e^{-inx}\, d\mu(x) = 0 \quad \text{for } n = 1, 2, \ldots,$$

then μ *is absolutely continuous.* The obvious analogue in several dimensions is trivially false; nevertheless Bochner [6] has found a generalization for set functions μ on the torus. The Riesz Theorem is a convenient tool in proving Szegö's Theorem [1, p. 263]; but some accounts of prediction theory (for example [8]) do not mention it. We have tried to clarify the relation betwen these theorems by giving a new proof of Bochner's Theorem based on the results of preceding sections. It is of methodological interest that our proof does not depend on theorems about analytic functions, as have all the published proofs of the Riesz Theorem.

In section two we generalized Szegö's Theorem to functions of several variables. In section five we consider another kind of generalization: we study functions defined on the unit circle whose values are *matrices*. Wiener [19] was led to the study of matrix-valued functions by a prediction problem different from the one treated in section two. After seeing Wiener's paper we succeeded in extending our method to this case. The fundamental result, as before, is a generalization of Szegö's Theorem. From it flow the solution of a prediction problem, and a number of theorems about matrix-valued analytic functions defined in the circle. Recently Masani and Wiener have completed a paper [14] carrying Wiener's work much further. It is likely that there is a good deal of duplication in our results, although their version of Szegö's Theorem is different from ours. We are happy to accord Masani and Wiener the right of precedence, and to acknowledge our debt to Wiener's paper. We hope nevertheless that the systematic development presented here, as well as our new results, will justify the publication of this section.

In the last section we extend these theorems to their natural degree of generality. We consider functions defined on a compact abelian group whose dual is linearly ordered by a relation consistent with the group structure. The functions may take matrices as values. Then Szegö's Theorem and most of our other results can be extended to this setting, and the proofs are word for word the same as proofs of corresponding theorems in the body of the paper. The torus groups are the best examples of groups to which the analysis applies, but there is no restriction in dimension. In particular, the Bohr compactification of the line (whose dual is the group of real numbers in the discrete topology) is of the type considered.

2. Doubly Stationary Series

Let x_{mn} be an element of a Hilbert space for each integer m and n. We say that $\{x_{mn}\}$ is *doubly stationary* if for all m, n, r, and s we have

$$(x_{m+r,\,n+s},\, x_{mn}) = (x_{rs},\, x_{00}). \tag{1}$$

In this case we define

$$\varrho(r, s) = (x_{rs},\, x_{00}). \tag{2}$$

Then ϱ is a positive definite function on the group of lattice points of the plane. That is, for any complex numbers $\alpha_1, \ldots, \alpha_k$ and integers $r_1, \ldots, r_k$, $s_1, \ldots, s_k$ we have

$$\sum_{i,j=1}^{k} \alpha_i \bar{\alpha}_j \varrho(r_i - r_j,\, s_i - s_j) \geqslant 0. \tag{3}$$

Indeed, using (1) this amounts to

$$\sum \alpha_i \bar{\alpha}_j (x_{r_i s_i},\, x_{r_j s_j}) \geqslant 0,$$

or

$$(\sum \alpha_i x_{r_i s_i},\, \sum \alpha_i x_{r_i s_i}) \geqslant 0.$$

The last inequality is obvious, and so (3) holds.

The theorem of Herglotz, Bochner, and Weil on positive definite functions states that there is a non-negative measure μ defined for Borel sets on the torus

$$0 \leqslant x \leqslant 2\pi, \quad 0 \leqslant y \leqslant 2\pi$$

such that

$$\varrho(r, s) = \int e^{-i(rx+sy)}\, d\mu(x, y)$$

for all integers r and s.

Now let S be any set of lattice points (m, n) in the plane not containing $(0, 0)$, and let $\{a_{mn}\}$ be a set of numbers defined for (m, n) in S, vanishing except for a finite set of indices. Taking (1) and (2) into account we find

$$\left.\begin{aligned} &\|x_{00} + \sum_S a_{mn} x_{mn}\|^2 \\ &\quad = (x_{00}, x_{00}) + \sum \bar{a}_{mn}(x_{00}, x_{mn}) + \sum a_{mn}(x_{mn}, x_{00}) + \sum\sum a_{mn}\bar{a}_{rs}(x_{mn}, x_{rs}) \\ &\quad = \varrho(0,0) + \sum \bar{a}_{mn}\bar{\varrho}(m, n) + \sum a_{mn}\varrho(m, n) + \sum\sum a_{mn}\bar{a}_{rs}\varrho(m-r, n-s) \\ &\quad = \int |1 + \sum_S a_{mn} e^{-i(mx+ny)}|^2\, d\mu(x, y). \end{aligned}\right\} \tag{4}$$

Thus the problem of approximating x_{00} by a linear combination of elements x_{mn} with (m, n) in S is equivalent to minimizing the integral at the end of (4). An explicit evaluation of the infimum is given for the corresponding expression in one variable by the following theorem of Szegö [16]: [1]

If μ is a finite non-negative measure defined on the Borel sets of the circle $|z|=1$ whose absolutely continuous part is $w(e^{ix})\,dx/2\pi$ then we have

$$\exp\left\{\frac{1}{2\pi}\int \log w\,dx\right\} = \inf_P \int |1+P(e^{ix})|^2\,d\mu(x),$$

where P ranges over the trigonometric polynomials of the form

$$P(e^{ix}) = a_1 e^{ix} + a_2 e^{2ix} + \cdots + a_n e^{nix}.$$

The left side is to be interpreted as zero if

$$\int \log w(e^{ix})\,dx = -\infty.$$

The solution of the prediction problem for any set S of lattice points requires an appropriate generalization of Szegö's Theorem. We shall find such a generalization for a very special class of sets S. Before stating our theorem we make some observations which do not require hypotheses on S.

Trigonometric polynomials of the form

$$1 + \sum_S a_{mn} e^{-i(mx+ny)} \tag{5}$$

form a convex subset of the Hilbert space of functions square-summable with respect to μ. The closure of this subset will be called $\mathcal{S}$. If $\mathcal{S}$ contains the null function, then x_{00} lies in the manifold spanned by $\{x_{mn}\}$ for (m, n) in S, and we say that prediction is perfect. Otherwise (and this is the interesting case) any sequence of elements Q_n of $\mathcal{S}$ such that

$$\lim \|Q_n\| = \inf \|G\| \qquad (G \in \mathcal{S})$$

is a Cauchy sequence, and converges to the unique element $1+H$ of $\mathcal{S}$ having minimal norm. We have therefore

$$\inf \int \Big|1 + \sum_S a_{mn} e^{-i(mx+ny)}\Big|^2 d\mu = \int |1+H|^2\,d\mu > 0. \tag{6}$$

---

(1) Szegö stated the theorem for absolutely continuous measures. It was completed by Kolmogoroff and Krein; references are given in [1]. We shall nevertheless refer to the full result as Szegö's Theorem.

For any complex number λ and (m, n) in S the function

$$1 + H(e^{ix}, e^{iy}) + \lambda e^{-i(mx+ny)}$$

belongs to $\mathfrak{S}$, and therefore

$$\int |1 + H(e^{ix}, e^{iy}) + \lambda e^{-i(mx+ny)}|^2 d\mu$$

has a unique minimum at $\lambda = 0$. Hence for every (m, n) in S we have

$$\int [1 + H(e^{ix}, e^{iy})] e^{i(mx+ny)} d\mu = 0. \tag{7}$$

If S is closed under group addition, so that

$$(m, n) \in S \quad \text{and} \quad (m', n') \in S$$

imply

$$(m + m', n + n') \in S,$$

then there is a second orthogonality relation. For each complex λ and each (m, n) in S the function

$$[1 + H(e^{ix}, e^{iy})]\ [1 + \lambda e^{-i(mx+ny)}]$$

belongs to $\mathfrak{S}$, and its norm is minimized at $\lambda = 0$. The conclusion is now

$$\int |1 + H(e^{ix}, e^{iy})|^2 e^{i(mx+ny)} d\mu = 0. \tag{8}$$

By taking the complex conjugate of (8) we see the same is true if $(-m, -n)$ is in S.

It is easy to prove that (7) *characterizes* the minimal element of $\mathfrak{S}$. Indeed, suppose that (7) holds but $1 + G$ is the minimal element. Then

$$\int |1 + H + \lambda(G - H)|^2 d\mu = \int |1 + H|^2 d\mu + |\lambda|^2 \int |G - H|^2 d\mu$$

for every complex λ. This expression is obviously smallest for $\lambda = 0$; but it is at least as small for $\lambda = 1$ if G is the minimal element. Since the minimal function is unique, we conclude that $G = H$.

DEFINITION. S is a *half-plane of lattice points* if

1° $(0, 0) \notin S$
2° $(m, n) \in S$ if and only if $(-m, -n) \notin S$ unless $m = n = 0$
3° $(m, n) \in S$ and $(m', n') \in S$ imply $(m + m', n + n') \in S$.

If S is a half-plane and (8) holds for all (m, n) in S, then by the second condition (8) holds for all m and n except $m = n = 0$. That is, the Fourier-Stieltjes coefficients of the measure $|1+H|^2 d\mu$ all vanish except the central one. Therefore this measure is a multiple of Lebesgue measure. It follows that $1+H$ must vanish almost everywhere with respect to the singular component of $d\mu$, and (7) can be written

$$\int [1+H(e^{ix}, e^{iy})]\, e^{i(mx+ny)}\, d\mu_a = 0 \qquad ((m, n) \in S) \tag{7'}$$

where μ_a is the absolutely continuous part of μ.

We can now state the first generalization of Szegö's Theorem.

THEOREM 1. *Let S be a half-plane of lattice points and let μ be a finite non-negative measure on the torus. Let μ have Lebesgue decomposition*

$$d\mu(x, y) = w(e^{ix}, e^{iy})\, d\sigma + d\mu_s(x, y),$$

where w is non-negative and summable for the measure $d\sigma = dx\,dy/4\pi^2$, and μ_s is singular with respect to $d\sigma$. Then

$$\exp\left\{\int \log w\, d\sigma\right\} = \inf_P \int |1+P|^2\, d\mu, \tag{9}$$

where P ranges over finite sums of the form

$$P(e^{ix}, e^{iy}) = \sum_S a_{mn}\, e^{-i(mx+ny)}. \tag{10}$$

The left side of (9) *is to be interpreted as zero if*

$$\int \log w\, d\sigma = -\infty. \tag{11}$$

Proof. If the infimum in (9) is positive, we have seen that it is equal to

$$\int |1+H|^2\, d\mu,$$

where $1+H$ belongs to $\mathfrak{S}$ and vanishes almost everywhere for μ_s. Hence (7′) holds. Moreover $1+H$ belongs to the convex set $\mathfrak{S}$ formed with the measure $w\,d\sigma$ instead of $d\mu$, and (7′) implies that $1+H$ is the minimal function relative to this measure:

$$\inf_P \int |1+P|^2\, w\, d\sigma = \int |1+H|^2\, w\, d\sigma = \int |1+H|^2\, d\mu.$$

Therefore it will suffice to prove

$$\exp\left\{\int \log w\, d\sigma\right\} = \inf_P \int |1+P|^2\, w\, d\sigma. \tag{12}$$

On the other hand, even if the infimum in (9) is zero it is enough to prove (12). For in that case the infimum in (12) surely vanishes, and having proved (12), obviously (9) holds. So we shall prove (12) for an arbitrary non-negative summable function w. It will be convenient to establish two lemmas.

LEMMA 1. *If w is a non-negative summable function,*

$$\exp\left\{\int \log w \, d\sigma\right\} = \inf_{\psi} \int e^{\psi} w \, d\sigma, \tag{13}$$

where ψ ranges over the real summable functions such that

$$\int \psi \, d\sigma = 0. \tag{14}$$

The geometric and arithmetic means of w are related by the well-known inequality

$$\exp\left\{\int \log w \, d\sigma\right\} \leqslant \int w \, d\sigma.$$

The same remark applies to $e^{\psi} w$, where ψ is any summable function which satisfies (14), and we find therefore

$$\exp\left\{\int \log w \, d\sigma\right\} \leqslant \inf_{\psi} \int e^{\psi} w \, d\sigma. \tag{15}$$

The opposite inequality will be established first assuming $\log w$ is summable. Define

$$\lambda = \int \log w \, d\sigma; \qquad \psi = \lambda - \log w. \tag{16}$$

Then ψ satisfies (14), and we have

$$\int e^{\psi} w \, d\sigma = \int e^{\lambda} d\sigma = \exp\left\{\int \log w \, d\sigma\right\}.$$

Therefore the inequality in (15) must be equality, and the minimal function is given by (16). We shall have to refer to the form of the minimal function again.

If $\log w$ is not summable this argument does not apply, and except in trivial cases no minimal function exists. But $\log (w+\varepsilon)$ is summable for each $\varepsilon > 0$, and by what we have just proved

$$\exp\left\{\int \log (w+\varepsilon) \, d\sigma\right\} = \inf \int e^{\psi} (w+\varepsilon) \, d\sigma \geqslant \inf \int e^{\psi} w \, d\sigma.$$

As ε tends to zero we obtain by the monotone limit theorem

$$\exp\left\{\int \log w \, d\sigma\right\} = 0 \geqslant \inf \int e^{\psi} w \, d\sigma \geqslant 0,$$

from which the statement of the lemma follows.

LEMMA 2. *For any non-negative summable function w we have*

$$\exp\left\{\int \log w\, d\sigma\right\} = \inf_{\psi} \int e^{\psi} w\, d\sigma, \tag{17}$$

where ψ ranges over the real trigonometric polynomials *satisfying* (14).

It suffices to prove the lemma assuming that $\log w$ is summable; for as in Lemma 1, the general case can be treated by a limit process. Divide w by a constant, if necessary, so that

$$\int \log w\, d\sigma = 0. \tag{18}$$

Now let u and v be the positive and negative parts of $\log w$, respectively, so that

$$u,\ v \geqslant 0; \qquad \log w = u - v.$$

Choose a sequence $u_1, u_2, \ldots$ of bounded non-negative functions increasing pointwise to u, and a sequence $v_1, v_2, \ldots$ of bounded non-negative functions increasing to v. By the monotone limit theorem,

$$\lim \int u_n\, d\sigma = \int u\, d\sigma = \int v\, d\sigma = \lim \int v_n\, d\sigma.$$

Consequently for each n there is an m such that

$$\int u_n\, d\sigma \leqslant \int v_m\, d\sigma.$$

In case the inequality is strict, multiply v_m by a constant smaller than one so that equality obtains, and rename the function v_n. We have then

$$0 \leqslant u_n \leqslant u; \qquad 0 \leqslant v_n \leqslant v; \qquad \int u_n\, d\sigma = \int v_n\, d\sigma.$$

Moreover the sequence u_n increases monotonically to u, and it is easy to see that v_n tends pointwise to v. From the construction it follows that

$$0 \leqslant e^{(u-u_n)-(v-v_n)} \leqslant \max\,(1, w).$$

Therefore the Lebesgue dominated convergence theorem applies to give

$$\lim \int e^{v_n - u_n} w\, d\sigma = \lim \int e^{(u-u_n)-(v-v_n)}\, d\sigma = 1.$$

Since the function $\psi = v_n - u_n$ satisfies (14), we have proved that

$$\inf_{\psi} \int e^{\psi} w\, d\sigma \leqslant 1, \tag{19}$$

where now ψ ranges over *bounded* functions satisfying (14). Every bounded function ψ is boundedly the limit of Fejér means of its Fourier series (in one or several dimensions); each approximating function is a trigonometric polynomial which is real if ψ is real, and satisfies (14) if ψ does. Therefore (19) continues to hold if ψ is restricted to real trigonometric polynomials with vanishing integral. In view of (18) and Lemma 1, the inequality of (19) must be equality, and the proof is complete.

We return now to the proof of (12) itself. The most general trigonometric polynomial ψ satisfying (14) can be written, on account of the second property of half-planes, in the form

$$\sum_S a_{mn} e^{-i(mx+ny)} + \sum_S \bar{a}_{mn} e^{i(mx+ny)}. \tag{21}$$

If P denotes the trigonometric polynomial (10), we have

$$\psi = P + \overline{P} = 2 \text{ Re } (P).$$

Therefore the result of Lemma 2 can be restated

$$\exp \left\{\int \log w \, d\sigma\right\} = \inf_P \int |e^P|^2 w \, d\sigma, \tag{22}$$

where P ranges over trigonometric polynomials of the form (10).

On account of the third property of half-planes, it is clear that

$$e^P = 1 + Q,$$

where Q is a continuous function with vanishing integral and having Fourier series of the form (10), although of course Q is not a trigonometric polynomial. Therefore we have

$$\exp \left\{\int \log w \, d\sigma\right\} \geqslant \inf_P \int |1 + P|^2 w \, d\sigma, \tag{23}$$

where P ranges over *all* continuous functions with Fourier series (10). The infimum is not increased if P is restricted to the class of trigonometric polynomials of the form (10), and so we have proved the first half of (12).

The opposite inequality can, paradoxically, be deduced from (23) itself. Replace w in that formula by $|1+Q|^2$, where Q is any polynomial of the form (10):

$$\exp \left\{\int \log |1+Q|^2 d\sigma\right\} \geqslant \inf_P \int |1 + P + Q + PQ|^2 d\sigma \geqslant 1,$$

making use once more of the semi-group property of S. Hence $\log |1+Q|^2$ is summable and

$$\int \log |1+Q|^2 d\sigma \geqslant 0.$$

Therefore we can write

$$|1+Q|^2 = k e^{\psi}; \qquad k \geqslant 1, \qquad \int \psi \, d\sigma = 0. \tag{24}$$

Now if w is an arbitrary non-negative summable function and Q is a polynomial of the form (10) we have by (24)

$$\int |1+Q|^2 w \, d\sigma = k \int e^{\psi} w \, d\sigma \geqslant \inf_{\psi} \int e^{\psi} w \, d\sigma = \exp \left\{ \int \log w \, d\sigma \right\}. \tag{25}$$

But this inequality is exactly the opposite of (23) if we pass to the infimum over Q, and so (12) has been proved. This completes the proof of the theorem.

Theorem 1 is a full generalization of Szegö's Theorem. We have already pointed out its connection with prediction theory; in the next section we shall apply it to multiple Fourier series.

3. Multiple Fourier Series

The first application of Theorem 1 is a partial generalization of Jensen's formula.

THEOREM 2. *Let f be summable on the torus with Fourier series*

$$f(e^{ix}, e^{iy}) \sim b + \sum_{S} b_{mn} e^{-i(mx+ny)}, \tag{26}$$

where S is any half-plane. Then

$$\int \log |f| \, d\sigma \geqslant \log |b|. \tag{27}$$

Proof. By Theorem 1,

$$\exp \left\{ \int \log |f| \, d\sigma \right\} = \inf_{P} \int |1+P|^2 |f| \, d\sigma,$$

where P ranges over the trigonometric polynomials of the form (10). If f is square-summable, we can replace $|f|$ in the last formula by $|f|^2$ and then take the square root of both sides:

$$\exp \left\{ \int \log |f| \, d\sigma \right\} = \inf_{P} \left[\int |(1+P) f|^2 \, d\sigma \right]^{1/2}. \tag{28}$$

If we set in the Fourier series (26) for f we obtain in the product $(1+P)f$ a constant term b, since P has no constant term. By the Parseval equality, the right side of (28) is at least $|b|$, so that (27) holds.

If f is not square-summable, let $\{f_n\}$ be the Fejér means of f. Each f_n is a trigonometric polynomial with constant term b, and the sequence converges to f in L. For any $\varepsilon > 0$ and each n we have

$$\int \log [|f_n| + \varepsilon] \, d\sigma \geqslant \int \log |f_n| \, d\sigma \geqslant \log |b|.$$

Passing to the limit in n with ε fixed,

$$\int \log [|f| + \varepsilon] \, d\sigma \geqslant \log |b|.$$

The result follows by letting ε tend to zero.

For analytic functions of one variable the deficiency of the right side of (27) cannot be evaluated without some stronger hypothesis about the function. (Some consequences of this fact are explored in [5].) It would be interesting to replace (27) by an equation analogous to Jensen's formula if, for example, f is a trigonometric polynomial.

COROLLARY. *If f is summable on the torus, has Fourier series of the form* (26), *and has mean value different from zero, then* $\log |f|$ *is summable.* (¹)

The proof is immediate. In one dimension the corresponding theorem requires no hypothesis on the mean value b, but here some such condition is indispensable. To see that this is so, construct a sequence of functions of one variable, $g_1, g_2, \ldots$, each vanishing on a fixed interval (α, β) in $(0, 2\pi)$, with Fourier series

$$g_m(e^{iy}) = \sum_n a_{mn} e^{-iny}; \qquad \sum_n |a_{mn}| \leqslant 1/m^2.$$

Define
$$f(e^{ix}, e^{iy}) = \sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} a_{mn} e^{-i(mx+ny)}.$$

Then f has absolutely convergent Fourier series and so is summable; moreover its coefficients are restricted to a half-plane. But f vanishes on the set

$$0 \leqslant x \leqslant 2\pi; \qquad \alpha \leqslant y \leqslant \beta,$$

so that $\log |f|$ cannot be summable.

If the coefficients of f are restricted to a sector of opening *smaller than* π, then the conclusion of the corollary holds without any restriction on the mean value of f, provided f is not the null function. The proof is like that of Theorem 2, making use of a construction used again in the proof of Bochner's Theorem in section four. In particular, the conclusion holds if f is an analytic function of two variables, as one can also show easily using Jensen's formula for analytic functions of one variable.

(¹) A similar theorem has been proved by Arens, even without the hypothesis that f has mean value different from zero. He assumes instead that f is defined on a compact group whose dual has an Archimedean order, corresponding here to the case of a half-plane bounded by a line of irrational slope. In the example which follows, the half-plane is bounded by a vertical line, so that the order relation defined by taking S as the set of positive elements is not Archimedean.

THEOREM 3. *Let w be non-negative and summable on the torus, and let S be any half-plane. A necessary and sufficient condition for w to have a representation*

$$w(e^{ix}, e^{iy}) = \left| b + \sum_S b_{mn} e^{-i(mx+ny)} \right|^2; \qquad b \neq 0, \qquad \sum |b_{mn}|^2 < \infty \tag{29}$$

is that

$$\int \log w \, d\sigma > -\infty.\text{(}^1\text{)} \tag{30}$$

Proof. If w has the form (29), then as in the proof of Theorem 2

$$\exp\left\{\int \log w \, d\sigma\right\} = \inf_P \int |(1+P)|^2 w \, d\sigma \geqslant |b|^2 > 0.$$

Thus (30) holds.

Conversely, suppose (30) is true. Then there is a unique function H such that

$$\lim \int |H - P_n|^2 w \, d\sigma = 0 \tag{31}$$

for a sequence of trigonometric polynomials P_n of the form (10), and satisfying

$$\exp\left\{\int \log w \, d\sigma\right\} = e^{\lambda} \equiv |1+H|^2 w.$$

We shall prove that the obvious equality

$$w = \left| \frac{e^{\lambda/2}}{1+H} \right|^2 \tag{32}$$

is a representation for w in the form (29).

By (32), $(1+H)^{-1}$ is square-summable. Its Fourier coefficients are

$$\int e^{-i(mx+ny)} \frac{1}{1+H} d\sigma = \int e^{-i(mx+ny)} \frac{1+\bar{H}}{|1+H|^2} d\sigma = e^{-\lambda} \int e^{-i(mx+ny)} (1+\bar{H}) \, w \, d\sigma.$$

According to (7) this integral vanishes for every (m, n) in S. Therefore the Fourier series of $(1+H)^{-1}$ has the form

$$b + \sum_S b_{mn} e^{-i(mx+ny)}.$$

If $b = 0$, so that

$$\int (1+\bar{H}) \, w \, d\sigma = 0,$$

we should have

$$\int (1+P)(1+\bar{H}) \, w \, d\sigma = 0$$

for every trigonometric polynomial P of the form (10). From this fact and (31) would follow

---

(1) For functions of one variable this was first proved by Szegö [17], using exactly the present method.

$$\int |1+H|^2 w\, d\sigma = 0,$$

which contradicts (30); therefore $b \neq 0$, and the proof is finished.

Let R be a set of lattice points containing the origin and closed under addition. It is an interesting problem, suggested to the authors by Mr. G. Weiss and Professor A. Zygmund, to determine whether every summable function f with Fourier series

$$f(e^{ix}, e^{iy}) \sim \sum_R a_{mn} e^{-i(mx+ny)}$$

can be represented as a product $g \cdot h$, where g and h are square-summable and, like f, have coefficients restricted to R. Even if R is taken to be the set of lattice points in the first quadrant, so that the problem concerns analytic functions of two variables, the answer seems not to be known. Our next theorem treats the case of a half-plane.

THEOREM 4. *Let S be a half-plane and f a summable function with Fourier series*

$$a + \sum_S a_{mn} e^{-i(mx+ny)}; \qquad a \neq 0.$$

There exist square-summable functions g and h with Fourier series of the same form such that $f = g \cdot h$.

Proof. Since the leading coefficient of f is not zero, the corollary of Theorem 2 states that $\log |f|$ is summable. By Theorem 3 and its proof,

$$|f| = |g|^2,$$

where

$$g = \frac{c}{1+H} \sim b + \sum_S b_{mn} e^{-i(mx+ny)}; \qquad b,\ c \neq 0.$$

If we set $h = g^{-1} f$, it is clear that g and h are square-summable, and g at least has Fourier series of the required kind. But we can write

$$h = g^{-1} f = c^{-1}(1+H) f = c^{-1} \lim\, (1+P_n) f,$$

where each P_n is a trigonometric polynomial of the form (10), and the limit is taken in the norm of the space L. It follows that the Fourier coefficients of h are restricted to S (aside from the constant term), and this completes the proof.

We can now derive an analogue for multiple Fourier series of the classical theorem of Hardy and Littlewood which states [10; 20, p. 158]: *if f is summable on the circle with Fourier series*

$$f(e^{ix}) \sim \sum_0^\infty a_n e^{inx},$$

then for a certain absolute constant k

$$\sum_{0}^{\infty} |a_n|/(n+1) \leqslant k \int |f(e^{ix})|\, dx.$$

This theorem follows from Hilbert's inequality and a factorization theorem for analytic functions. The same method of proof works in higher dimensions; first we shall quote an extension of Hilbert's inequality from work of Calderón and Zygmund, and then Theorem 4 will furnish exactly the factorization theorem we need.

In the paper [7] of Calderón and Zygmund, Theorem 14 states: *let K be a function defined on the lattice points, except the origin, and have the form*

$$K(m, n) = \frac{\Omega(e^{i\theta})}{r^2}; \quad m + i n = r e^{i\theta}.$$

Suppose that Ω is continuous on the circle, satisfies a Lipschitz condition of positive order, and also

$$\int_{0}^{2\pi} \Omega(e^{i\theta})\, d\theta = 0.$$

Then there is a constant k depending only on K such that for any square-summable sequences $\{x_{mn}\}$ and $\{y_{rs}\}$ we have

$$\left|\sum{}' K(m+r, n+s)\, x_{mn}\, y_{rs}\right| \leqslant k \left[\sum |x_{mn}|^2 \sum |y_{rs}|^2\right]^{1/2}. \tag{33}$$

(The summation on the left is extended over all indices for which the summand is defined.)

Let S be any half-plane. There is an angle α, uniquely determined up to multiples of 2π, such that every lattice point $m + i n = r e^{i\theta}$ in S satisfies $\alpha \leqslant \theta \leqslant \alpha + \pi$. Define a function Ω to be one for $\alpha \leqslant \theta \leqslant \alpha + \pi$; then extend Ω to the rest of the circle so as to be continuously differentiable and have mean value zero. The corresponding function K is a kernel to which the theorem of Calderón and Zygmund applies. Suppose that $x_{mn} = 0$ unless (m, n) is in S, or $m = n = 0$, and the same for the y_{rs}. Then the only terms which contribute to the sum on the left side of (33) are those for which

$$K(m, n) = 1/(m^2 + n^2).$$

We have therefore the following analogue of Hilbert's inequality:

$$\left|\sum{}' \frac{x_{mn}\, y_{rs}}{(m+r)^2 + (n+s)^2}\right| \leqslant k \left[\sum |x_{mn}|^2 \sum |y_{rs}|^2\right]^{1/2}. \tag{34}$$

THEOREM 5. *Let f be summable on the torus and have Fourier series*

$$f(e^{ix}, e^{iy}) \sim a + \sum_S a_{mn} e^{-i(mx+ny)}.$$

There is an absolute constant k' such that

$$|a| + \sum_S |a_{mn}|/(m^2+n^2+1) \leqslant k' \int |f(e^{ix}, e^{iy})| \, d\sigma. \tag{35}$$

Proof. By continuity it suffices to consider the case $a \neq 0$. Theorem 4 states that f can be written as the product of the square-summable functions

$$g = b + \sum_S b_{mn} e^{-i(mx+ny)}; \qquad b \neq 0$$

and

$$h = c + \sum_S c_{mn} e^{-i(mx+ny)}; \qquad c \neq 0.$$

Then we have

$$a_{mn} = \sum_{r,s} b_{m-r,\, n-s} c_{rs},$$

so that

$$\sum_S \frac{|a_{mn}|}{m^2+n^2} \leqslant \sum{}' \frac{|b_{m-r,\, n-s} c_{rs}|}{m^2+n^2} = \sum{}' \frac{|b_{mn} c_{rs}|}{(m+r)^2+(n+s)^2}.$$

If we set

$$x_{mn} = |b_{mn}|,$$

$$y_{rs} = |c_{rs}|,$$

then (34) applies to give

$$\sum_S |a_{mn}|/(m^2+n^2) \leqslant k \left[\sum |b_{mn}|^2 \sum |c_{rs}|^2 \right]^{1/2}.$$

From the proof of Theorem 4 we know that $|g|^2 = |h|^2 = |f|$; by the Plancherel Theorem therefore

$$\sum_S |a_{mn}|/(m^2+n^2) \leqslant k \int f(e^{ix}, e^{iy}) | \, d\sigma.$$

The statement of the theorem follows trivially from this formula, with $k' = k + 1$.

There is no difficulty in proving an analogous theorem for tori of any finite dimension (since both Theorem 4 and the theorem of Calderón and Zygmund are true in general). We do not know whether there is a generalization to the class of compact groups discussed in the last section.

Theorem 5 applies to a larger class of functions than the double power series. However, as Bochner has remarked, a stronger result holds for the double power series, and can be proved easily from the theorem of Hardy and Littlewood. The theorem is as follows: *if f is summable on the torus with Fourier series*

$$\sum_{m,n=0}^{\infty} a_{mn} e^{i(mx+ny)},$$

then
$$\sum_{m,n=0}^{\infty} |a_{mn}|/(mn+1) \leqslant k \int |f|\, d\sigma.$$

The last theorem of this section generalizes a theorem of Beurling [5].

THEOREM 6. *Let H be the linear subspace of L^2 consisting of the functions whose Fourier series have the form*
$$a + \sum_S a_{mn} e^{-i(mx+ny)}.$$

For any f in H let C_f be the smallest closed linear manifold containing
$$(b+P)f$$
for all constants b and trigonometric polynomials P of the form (10). *We have $C_f = H$ if and only if*
$$\int \log |f|\, d\sigma = \log \left| \int f\, d\sigma \right| > -\infty.$$

It is easy to see that $C_f = H$ if and only if there is a non-zero constant function in the closure of the convex set of functions $(1+P)f$. The proof that this is equivalent to the condition of the theorem is easy to carry out using Theorems 1 and 2 and the Parseval equality.

The problems discussed here for the case where S is a half-plane become much more difficult when S is, for example, the set of lattice points contained in some sector of opening smaller than π. There is no longer any analogy with analytic functions of one variable. It seems to us that these new problems are difficult and interesting.

4. Theorem of Riesz and Bochner

The theorem of F. and M. Riesz [15] (already referred to in the Introduction) states: *if μ is a bounded complex Borel measure on the circle whose Fourier-Stieltjes coefficients vanish for positive indices, then μ is absolutely continuous with respect to Lebesgue measure.* Bochner [6] observed that not every measure on the torus with coefficients restricted to a half-plane is absolutely continuous with respect to the invariant measure on the torus; but Bochner proved that the conclusion holds if the non-vanishing coefficients are all in a sector of opening less than π. The machinery of Bochner's proof is very elaborate. In this section we shall give a new proof of Bochner's Theorem which shows its close connection with prediction theory. On the way we present an example and some lemmas of independent interest.

A certain part of the Riesz Theorem survives in two dimensions, and gives the following preliminary result.

LEMMA 3. *Let μ be a complex measure on the torus without absolutely continuous part. If the Fourier-Stieltjes coefficients*

$$c_{mn} = \int e^{-i(mx+ny)}\, d\mu(x, y)$$

vanish for all (m, n) in a half-plane, then also $c_{00}=0$.

Denote the total variation of μ by ν. Then ν is also singular with respect to $d\sigma$, and so by Theorem 1

$$\inf_P \int |1+P|^2\, d\nu = 0,$$

where P ranges over the trigonometric polynomials of the form (10).

Let $P_1, P_2, \ldots$ be a sequence of such polynomials for which

$$\lim \int |1+P_n|^2\, d\nu = 0.$$

Then clearly

$$\lim \int (1+P_n)\, d\mu = 0.$$

By hypothesis, for each n

$$\int P_n\, d\mu = 0,$$

and so

$$c_{00} = \int d\mu = 0.$$

In the one-dimensional case, having shown in this way that c_0 is zero we can translate the coefficient sequence and prove in turn that $c_{-1}, c_{-2}, \ldots$ all vanish. In two dimensions we cannot conclude anything more from the fact that $c_{00}=0$; indeed there exist *singular* measures μ whose coefficients vanish in a half-plane but not everywhere.

The trivial example, which is mentioned by Bochner, is given by the product of a singular measure $d\gamma(x)$ on the interval with the measure $e^{-iy}\,dy$. The product measure $d\mu$ is clearly singular with respect to two-dimensional Lebesgue measure; its coefficients are given by

$$c_{mn} = \int e^{-imx}\, d\gamma(x) \cdot \int e^{-i(n+1)y}\, dy,$$

and thus vanish for all $n \geqslant 0$.

It is less obvious that there are singular measures whose coefficients vanish on a half-plane bounded by a line $\mathfrak{L}$ of *irrational* slope, and the following construction of such a measure may be of interest.

Project each lattice point onto the real line in the direction parallel to $\mathcal{L}$. Distinct lattice points have distinct projections, since $\mathcal{L}$ has irrational slope. Moreover, the vector sum of two lattice points is projected onto the ordinary sum of their separate projections. So the lattice points are isomorphic as a group with a denumerable dense subgroup of the line, which we endow with the discrete topology and call G. Now $\mathcal{L}$ determines two half-planes; the points of one half-plane are projected into the positive ray of the real line, and the points of the other half-plane into the negative ray.

Now let f be the function on the line equal to one at the origin, and decreasing linearly to zero at 1 and -1. It is well-known that f is positive definite. *A fortiori*, f is positive definite as a function on G. By the general theorem of Herglotz, Bochner, and Weil on positive definite functions, f is the transform of a positive measure on the dual group of G, which is the torus. If we set

$$g(x) = f(x+1) \qquad \text{for } x \in G,$$

then g is the transform of a complex measure μ on the torus, and g vanishes for $x \geqslant 0$. Considered as a function on the group of lattice points, g vanishes on a half-plane bounded by $\mathcal{L}$.

If μ were absolutely continuous, by the general Riemann-Lebesgue Lemma its transform g would tend to zero outside compact sets of G. Since G is discrete, this would mean that $|g(x)| \geqslant \varepsilon$ only for a finite set of x, for any $\varepsilon > 0$. Obviously this is not the case, so μ cannot be absolutely continuous. It will follow from the next theorem that the singular part of μ (in case μ is not itself singular) has coefficients vanishing on the same half-plane as the coefficients of μ, and this is the example we wanted to find.

THEOREM 7. *Let μ be a measure on the torus whose coefficients vanish on a half-plane S. Then the coefficients of its singular and absolutely continuous parts vanish separately on S.*

Proof. Let ν be the total variation of μ. After adding to μ a multiple of Lebesgue measure if necessary, we may assume that

$$\inf_P \int |1+P|^2 \, d\nu > 0,$$

where P ranges over trigonometric polynomials of the form (10). Choose a sequence of polynomials $P_1, P_2, \ldots$ such that

$$\lim \int |1+P_n|^2 \, d\nu = \inf_P \int |1+P|^2 \, d\nu,$$

and denote by $1+H$ the limit of $1+P_n$ in the space of functions square-summable for the measure $d\nu$. Let P and Q be any trigonometric polynomials of the form (10); using the hypothesis on the coefficients of μ we have

$$\int P(1+Q)(1+H)\,d\mu=\lim\int P(1+Q)(1+P_n)\,d\mu=0.$$

Since $1+H$ vanishes almost everywhere for the singular part of μ, we have also

$$\int P(1+Q)(1+H)\,f\,d\sigma=0, \tag{37}$$

where $f\,d\sigma$ is the absolutely continuous part of μ.

The theorem will be proved if we can show that

$$\int P\,f\,d\sigma=0 \tag{38}$$

for arbitrary P of the form (10), for this means that the Fourier coefficients of f vanish on S. We shall need the following relations:

$$(1+H)^{-1} \text{ belongs to } L^2;$$
$$(1+H)^{-1}\sim b+\sum_S b_{mn}\,e^{-i(mx+ny)} \qquad \text{with } b\neq 0;$$
$$(1+H)\cdot f \text{ belongs to } L^2.$$

Choose a sequence of trigonometric polynomials $\{1+Q_n\}$, with each Q_n of the form (10), converging in L^2 to $b^{-1}(1+H)^{-1}$. By (37), for each n

$$\int P(1+Q_n)(1+H)\,f\,d\sigma=0;$$

since $(1+H)\cdot f$ is square-summable we can pass to the limit in n and obtain (38). This completes the proof.

Bochner's Theorem. *Let T be a sector of the plane with opening greater than π radians. Suppose μ is a measure on the torus whose coefficients vanish on T. Then μ is absolutely continuous with respect to Lebesgue measure.*

Proof. It suffices to consider a sector T with center at the origin, so that T contains the union of different half-planes S and S'. Let μ_s be the singular part of μ; by Theorem 7, the coefficients of μ_s vanish on S and also on S'. If any coefficient of μ_s is different from zero, it is easy to see that the coefficient set can be translated so as to bring a non-zero coefficient to the origin, still leaving a half-plane free of non-zero coefficients. (Indeed, find a line $\mathfrak{L}$ through the origin with irrational slope, lying between the lines bounding S and S' in such a way that T contains one of the half-planes bounded by $\mathcal{L}$. If we translate $\mathcal{L}$ we encounter a first lattice point at which μ_s has non-zero coefficient, and the inverse translation is the required one.) But the result of this construction is a coefficient set belonging to a singular measure, vanishing on a half-plane, but not at the origin. This contradicts Lemma 3, and so μ_s is the null measure. This completes the proof.

5. Matrix Valued Analytic Functions

For each point e^{ix} on the circle let $A(e^{ix})$ be an n by n matrix with entries $a_{jk}(e^{ix})$ $(j, k=1, \ldots, n)$. The *normalized trace* of A is the scalar function

$$\operatorname{tr} A(e^{ix}) = \frac{1}{n} \sum_k a_{kk}(e^{ix}).$$

The trace and determinant functions are related by the formula

$$\det e^A = e^{n \operatorname{tr} A},$$

where e^A may be defined by its power series.

The normalized trace has the following properties (and, in fact, is determined by them): for any matrices A, B and scalars a, b

$$\left.\begin{array}{l} \operatorname{tr}(aA+bB) = a \operatorname{tr} A + b \operatorname{tr} B \\ \operatorname{tr}(AB) = \operatorname{tr}(BA), \text{ or equivalently} \\ \operatorname{tr}(U^{-1}AU) = \operatorname{tr} A \text{ if } U \text{ is unitary} \\ \operatorname{tr} A^*A \geqslant 0, \text{ from which follows } \operatorname{tr} A^* = \overline{\operatorname{tr} A} \\ \operatorname{tr} I = 1, \; I \text{ the unit matrix.} \end{array}\right\} \tag{39}$$

If A is a positive definite matrix, there is a unique Hermitian matrix B satisfying

$$e^B = A, \tag{40}$$

and we define

$$B = \log A.$$

By a *trigonometric polynomial*, in the context of matrix functions, we shall mean a finite sum of the form

$$\sum A_k e^{ikx},$$

where each A_k is a constant matrix. The trigonometric polynomial is *analytic* if $A_n = 0$ for $n < 0$.

If each component function a_{jk} of the matrix function A is summable, we shall say that A is summable, or belongs to L. More generally, L^p is to consist of the matrix functions A whose scalar components a_{jk} all belong to the ordinary class L^p. A summable matrix function A has Fourier series

$$A(e^{ix}) \sim \sum A_k e^{ikx},$$

where each A_k is the constant matrix defined by the n^2 scalar equations

$$A_k = \int A(e^{ix}) e^{-ikx} d\sigma(x).$$

(In this section, $d\sigma(x)$ is the measure $dx/2\pi$ on the circle.) More generally, if M is a completely additive matrix-valued function of Borel sets (in other words a matrix whose entries are complex measures), we shall write

$$dM(e^{ix}) \sim \sum A_k e^{ikx}$$

with the A_k defined as

$$A_k = \int e^{-ikx} dM(e^{ix}).$$

It follows from definition that a measurable matrix function A is in L^2 if and only if tr $(A^* A)$ is summable. We shall also need the fact that a measurable positive semi-definite matrix function W is summable if and only if tr W is a summable scalar function.

The ring of constant matrices possesses the natural inner product

$$(A, B) = \operatorname{tr}(B^* A). \tag{41}$$

We can extend this definition to the class of matrix functions in L^2 by setting

$$(A, B) = \int \operatorname{tr}(B^* A)\, d\sigma = \sum_{j,k} \int a_{jk} \bar{b}_{jk}\, d\sigma, \tag{42}$$

where a_{jk} and b_{jk} are component functions of A and B. The Parseval equality holds for square-summable functions A and B with Fourier coefficients A_k and B_k:

$$(A, B) = \sum (A_k, B_k); \tag{43}$$

in this formula the inner product on the left is defined by (42), and those on the right by (41).

The main theorem of this section is an extension of Szegö's Theorem to matrix-valued functions defined on the circle.

THEOREM 8. *Let M be a matrix-valued measure defined on the circle such that $M(E)$ is Hermitian and positive semi-definite for every Borel set E. Let M have Lebesgue decomposition*

$$dM(e^{ix}) = W(e^{ix})\, d\sigma + dM_s(e^{ix}),$$

where W *is a summable matrix function and* M_s *is singular with respect to* $d\sigma$. *Then*

$$\exp\left\{\int \operatorname{tr} \log W \, d\sigma\right\} = \inf_{A_0, P} \int \operatorname{tr} \left[(A_0 + P)^* (A_0 + P) \, d M\right], (^1) \tag{44}$$

where A_0 *ranges over the matrices with determinant one, and* P *over trigonometric polynomials of the form*

$$P(e^{ix}) = \sum_{k>0} A_k e^{ikx}. \tag{45}$$

The left side of (44) *is to be interpreted as zero if*

$$\int \operatorname{tr} \log W \, d\sigma = -\infty. \tag{46}$$

Proof. In outline we can follow the proof of Theorem 1, meeting each new complication as it arises. Let L_M^2 be the set of functions A for which

$$\| A \|_M^2 = \int \operatorname{tr} (A^* A \, d M) < \infty; \tag{47}$$

the norm so defined is positive semi-definite. After identifying functions which differ only on a null-set of dM, L_M^2 is a Hilbert space with inner product

$$(A, B)_M = \int \operatorname{tr} (B^* A \, d M). \tag{48}$$

If the infimum on the right side of (44) is positive, choose and fix A_0 with determinant one, and let H be that element of L_M^2 which is the limit of polynomials P of the form (45) and satisfies

$$\int \operatorname{tr} \left[(A_0 + H)^* (A_0 + H) \, d M\right] = \inf_P \int \operatorname{tr} \left[(A_0 + P)^* (A_0 + P) \, d M\right].$$

The argument leading to (7) and (8) gives analogous orthogonality relations here. If $n > 0$ and G is any non-zero constant matrix, the expression

$$\| A_0 + H + \lambda G e^{inx} \|_M$$

has a unique minimum at $\lambda = 0$. It follows that

$$(A_0 + H, \; G e^{inx})_M = 0 \qquad (n = 1, 2, \ldots). \tag{49}$$

And the expressions

$$\| (A_0 + H)(I + \lambda G e^{inx}) \|_M = \| (A_0 + H) + \lambda (A_0 + H) G e^{inx} \|_M,$$
$$\| (I + \lambda G e^{inx})(A_0 + H) \|_M = \| (A_0 + H) + \lambda G (A_0 + H) e^{inx} \|_M$$

---

(1) The pedantic reader can easily write this symbolic integral literally in terms of the scalar component measures of M.

have minima at $\lambda=0$, uniquely unless

$$(A_0+H)G=0 \quad \text{or} \quad G(A_0+H)=0.$$

In any case we have

$$((A_0+H),\ (A_0+H)Ge^{inx})_M=0 \tag{50}$$

$$(n=1,2,\ldots).$$

and

$$((A_0+H),\ G(A_0+H)e^{inx})_M=0 \tag{51}$$

The definition (48) means that (51) can be written

$$\int e^{-inx}\ \mathrm{tr}\ [(A_0+H)^*\,G^*\,(A_0+H)\,dM]=0. \tag{52}$$

Taking the complex conjugate of (52) and making use of (39),

$$\int e^{inx}\ \mathrm{tr}\ [(A_0+H)^*\,G\,(A_0+H)\,dM]=0.$$

These formulas hold for all G, and for $n=1,2,\ldots$; it is easy to see then that (52) is valid for negative as well as positive integers n. Hence

$$\mathrm{tr}\ [G(A_0+H)\,dM\,(A_0+H)^*]$$

is a constant multiple of scalar Lebesgue measure for each G, so that every component of the matrix measure

$$(A_0+H)\,dM\,(A_0+H)^*$$

is a multiple of Lebesgue measure. This fact can be written

$$(A_0+H)\,dM\,(A_0+H)^*=C\,d\sigma \qquad (C \text{ constant}).$$

Therefore we have

$$(A_0+H)\,dM_s\,(A_0+H)^*=0, \tag{53}$$

$$(A_0+H)\,W\,(A_0+H)^*=C. \tag{54}$$

It follows from (53) that A_0+H vanishes almost everywhere for dM_s, so that (49) takes the alternate form

$$(A_0+H,\ Ge^{inx})_W=0 \qquad (G \text{ arbitrary};\ n=1,2,\ldots), \tag{49'}$$

where the inner product refers to the Hilbert space of matrix functions square-summable for $W\,d\sigma$.

As in the scalar case, we conclude from (49′) that A_0+H has the same minimal property in L_W^2 that it enjoys in L_M^2, and so the infimum on the right side of (44) is not reduced if we replace dM by $W\,d\sigma$. Assuming, then, that this infimum is positive, the theorem will be proved if we show

$$\exp\left\{\int \mathrm{tr}\log W\,d\sigma\right\}=\inf_{A_0,\,P}\int \mathrm{tr}\ [(A_0+P)^*\,(A_0+P)\,W]\,d\sigma. \tag{55}$$

On the other hand, if the infimum in (44) is zero, it still suffices to prove (55), by the same argument as in the scalar case.

LEMMA 4. *Let W be Hermitian, positive semi-definite and summable. Then*

$$\exp\left\{\int \operatorname{tr} \log W \, d\sigma\right\} = \inf_{\Psi} \int \operatorname{tr} (e^{\Psi} W) \, d\sigma, \tag{56}$$

where Ψ ranges over the Hermitian matrix functions with summable trace for which

$$\int \operatorname{tr} \Psi \, d\sigma = 0. \tag{57}$$

The trace of a Hermitian matrix is the average of its proper values; and the determinant is the product of the same numbers. Using the inequality of the arithmetic and geometric means twice we obtain

$$\exp\left\{\int \operatorname{tr} \log W \, d\sigma\right\} = \exp\left\{\frac{1}{n}\int \log \det W \, d\sigma\right\} \leqslant \int (\det W)^{1/n} \, d\sigma \leqslant \int \operatorname{tr} W \, d\sigma. \tag{58}$$

In order to have continued equality it is necessary and sufficient that

$$\operatorname{tr} W \equiv (\det W)^{1/n} \equiv \text{constant},$$

which is to say that W is a constant multiple of the identity matrix.

Let Ψ be a Hermitian matrix function with summable trace satisfying (57); whether or not the positive semi-definite matrix function

$$W' = e^{\frac{1}{2}\Psi} W e^{\frac{1}{2}\Psi}$$

is summable, we have as in (58)

$$\exp\left\{\int \operatorname{tr} \log W' \, d\sigma\right\} \leqslant \int \operatorname{tr} W' \, d\sigma \leqslant \infty.$$

The properties of the trace and determinant functions give

$$n \operatorname{tr} \log W' = \log \det W' = \log \det (e^{\Psi} W) = n \operatorname{tr} \Psi + n \operatorname{tr} \log W,$$
$$\operatorname{tr} W' = \operatorname{tr} (e^{\Psi} W).$$

By (57) we have for every function Ψ

$$\exp\left\{\int \operatorname{tr} \log W \, d\sigma\right\} \leqslant \int \operatorname{tr} (e^{\Psi} W) \, d\sigma. \tag{59}$$

If $\operatorname{tr} \log W$ is summable, define

$$\Psi_0 = \lambda I - \log W; \qquad \lambda = \int \operatorname{tr} \log W \, d\sigma. \tag{60}$$

Then Ψ_0 is Hermitian and satisfies (57), and obviously reduces (59) to equality. This completes the proof if $\operatorname{tr} \log W$ is summable; otherwise a limiting process has to be carried out as in the scalar case.

LEMMA 5. *The statement* (56) *is still true if* Ψ *ranges only over the class of trigonometric polynomials which are Hermitian and satisfy* (57).

Let A and B be commuting Hermitian matrices. They have a common complete set of proper vectors. Define $\max(A, B)$ to be the Hermitian matrix with the same proper vectors and with proper values the larger of the corresponding proper values of A and B. For any Hermitian matrix A, the positive part of A can be defined as $\max(A, 0)$. This construction makes it possible to carry through the proof of Lemma 2 unchanged for the matrix case.

By analogy with the proof for scalar functions, we should like to factor each function e^{Ψ} of Lemma 5 into a product

$$(A_0 + A_1 e^{ix} + \cdots)^* (A_0 + A_1 e^{ix} + \cdots),$$

and then show that it suffices to consider trigonometric polynomials in place of the infinite series. The non-commutativity of matrices introduces a difficulty which must presently be met.

LEMMA 6. *Let* W *be a summable positive semi-definite matrix function for which the infimum of* (55) *is positive. Then* W *has a factorization*

$$W = BB^* \tag{61}$$

where B *is a matrix function in* L^2 *with analytic Fourier series:*

$$B(e^{ix}) \sim \sum_0^{\infty} B_n e^{inx} \quad \text{and } \det B_0 \neq 0. \tag{62}$$

In applying this lemma, we shall need a stronger result than (62) for a narrow class of functions W. It will be convenient to refer later to the proof as well as the statement of the lemma.

By hypothesis, the convex set of trigonometric polynomials of the form

$$I + \sum_{k>0} A_k e^{ikx}$$

is bounded from zero in L^2_W. Let $I + H$ be the unique element of minimal norm in the closure of this set. From (54) we have

$$(I + H)\, W\, (I + H)^* = C, \tag{63}$$

where C is a constant matrix. We assert that C is non-singular. Indeed, otherwise we could find matrices A with determinant one for which

$$\operatorname{tr}(A C A^*) = \int \operatorname{tr}\left[(A + AH)\, W\, (A + AH)^*\right] d\sigma$$

is as close to zero as we please, whereas this quantity is bounded from zero by hypothesis. Clearly then C is Hermitian and positive definite, and so has a non-singular square root. The factorization (63) can be put in the form

$$(A_0 + A_0 H)\, W\, (A_0 + A_0 H)^* = I \qquad (A_0 = C^{-\frac{1}{2}}) \tag{64}$$

or

$$W = B B^*; \qquad B = (A_0 + A_0 H)^{-1}. \tag{65}$$

From (65) it follows that B is square-summable. We shall prove the lemma by showing that its Fourier series is of analytic type:

$$(A_0 + A_0 H)^{-1} \sim B_0 + B_1 e^{ix} + \cdots. \tag{66}$$

It suffices to establish that

$$\int \operatorname{tr}\, [G (A_0 + A_0 H)^{-1}]\, e^{inx}\, d\sigma(x) = 0 \qquad (n = 1, 2, \ldots) \tag{67}$$

for every constant matrix G, since then every component function of $(A_0 + A_0 H)^{-1}$ is analytic. Making use of (64), the left side of (67) is equal to

$$\int \operatorname{tr}\, [G W (A_0 + A_0 H)^*]\, e^{inx}\, d\sigma(x) = \int \operatorname{tr}\, [(A_0 + A_0 H)^* G W]\, e^{inx}\, d\sigma(x) = (A_0^* G e^{inx}, I + H)_W.$$

This inner product vanishes by (49′) for $n = 1, 2, \ldots$, so that (67) holds.

Let Ψ be a trigonometric polynomial satisfying (57). *Then there exists a factorization*

$$e^{\Psi} = A^* A; \qquad A(e^{ix}) \sim \sum_0^{\infty} A_n e^{inx}, \qquad \det A_0 = 1. \tag{68}$$

To prove this fact, we consider the positive definite weight function $W = e^{-\Psi}$. The eigenvalues of W are bounded from zero and from infinity; it follows that the spaces L_W^2 and L^2 have equivalent norms. By a simple calculation we can show that

$$\| A \|^2 = \int \operatorname{tr}\, (A^* A)\, d\sigma \geqslant 1$$

for each analytic trigonometric polynomial A whose leading coefficient A_0 has determinant one. Therefore the infimum of (55) is positive for this function W, and by Lemma 6

$$e^{-\Psi} = B B^*; \qquad B(e^{ix}) \sim \sum_0^{\infty} B_n e^{inx}. \tag{69}$$

If the scalar components of $B(e^{ix})$ are denoted by $b_{ij}(e^{ix})$ for $i, j = 1, 2, \ldots, n$, we have

$$n \operatorname{tr} e^{-\Psi} = \sum_{i,j} | b_{ij} |^2,$$

from which it is clear that the functions b_{ij} are bounded. Now the determinant of B is a sum of products of these functions, and since the b_{ij} are bounded, the Fourier series of det B can be computed formally from the Fourier series of the component functions. Each b_{ij} is analytic, and we find

$$\det B(e^{ix}) \sim \det B_0 + c_1 e^{ix} + c_2 e^{2ix} + \cdots.$$

The one-dimensional version of Theorem 2 (or Jensen's formula) gives

$$\int \log |\det B|^2 \, d\sigma \geqslant \log |\det B_0|^2. \tag{70}$$

The inverse of B is the function $A = A_0 + A_0 H$, obtained as the limit in L^2_W of a sequence of analytic trigonometric polynomials each having constant term A_0. With the present choice of W the sequence converges in L^2 as well, so that A is an analytic element of L^2:

$$A \sim A_0 + A_1 e^{ix} + \cdots; \qquad A_0 = C^{-\frac{1}{2}}.$$

From (69) we have

$$e^{\Psi} = A^* A, \tag{71}$$

so the components of A, like those of B, are bounded functions; exactly as for B then we have

$$\int \log |\det A|^2 \, d\sigma \geqslant \log |\det A_0|^2. \tag{72}$$

Now A and B are bounded functions with analytic Fourier series, so the Fourier series of their product is obtained by formal multiplication of the series for A and B and consequently $A_0 B_0 = I$. It follows that the right side of (70) and of (72) is finite. Adding these inequalities gives zero on both sides. Therefore (70) and (72) are actually equalities.

By assumption Ψ satisfies (57). Making use of (71) we have

$$0 = \int \log \det e^{\Psi} \, d\sigma = \int \log |\det A|^2 \, d\sigma = \log |\det A_0|^2.$$

The determinant of A_0 is at any rate positive, and therefore is equal to one. Thus (71) is a factorization of the kind we wanted.

Now we can prove (55). Let W be a Hermitian positive semi-definite summable matrix function, and let Ψ be a Hermitian trigonometric polynomial satisfying (57). From the result just proved,

$$\int \operatorname{tr} (e^{\Psi} W) \, d\sigma = \int \operatorname{tr} (A^* A W) \, d\sigma,$$

where A is a *bounded* analytic function and $\det A_0 = 1$. Therefore by Lemma 5

$$\exp \left\{ \int \operatorname{tr} \log W \, d\sigma \right\} \geqslant \inf_A \int \operatorname{tr} (A^* A W) \, d\sigma,$$

where A ranges over all bounded functions of analytic type such that $\det A_0 = 1$. Each component function a_{ij} of A is boundedly the limit of Fejér means of its Fourier series, from which it follows that the last inequality remains true when A ranges merely over *trigonometric polynomials* of the same kind. This is one half of (55). The opposite inequality follows as in the scalar case, or can be deduced directly from the one-dimensional version of Theorem 2.

THEOREM 9. *Let W be a Hermitian positive semi-definite summable matrix function defined on the circle. A necessary and sufficient condition for W to have a factorization*

$$W = B B^*, \tag{73}$$

where B is in L^2 with Fourier series of the form

$$B(e^{ix}) \sim \sum_0^\infty B_n e^{inx} \qquad (\det B_0 \neq 0)$$

is that

$$\int \operatorname{tr} \log W \, d\sigma > -\infty. \tag{74}$$

If this condition is satisfied we can choose B so that

$$\int \log |\det B| \, d\sigma = \log |\det B_0|. \tag{75}$$

Proof. Suppose first that W has a factorization of the required kind. For any trigonometric polynomial of the form

$$A(e^{ix}) = A_0 + A_1 e^{ix} + \cdots, \qquad \det A_0 = 1$$

we have clearly

$$\int \operatorname{tr} (A^* A W) \, d\sigma = \int \operatorname{tr} [(A B)^* (A B)] \, d\sigma \geqslant \operatorname{tr} [(A_0 B_0)^* (A_0 B_0)] \geqslant$$
$$\geqslant |\det A_0 B_0|^{2/n} = |\det B_0|^{2/n} > 0. \tag{76}$$

Then (55) shows that $\operatorname{tr} \log W$ is summable.

Conversely, suppose that (74) holds. Then (55) shows that the hypothesis of Lemma 6 is satisfied, so there is a factorization (73). We shall show that the factorization furnished by Lemma 6 has the property (75).

From (76) and (55) once more we see that for any factorization of the form (73) we have

$$\exp \left\{ \int \operatorname{tr} \log W \, d\sigma \right\} \geqslant |\det B_0|^{2/n},$$

which is equivalent to one inequality in (75). Now for any positive definite matrix C, it is an elementary fact that

$$(\det C)^{1/n} = \inf_{A_0} \operatorname{tr} (A_0 C A_0^*),$$

where A_0 ranges over matrices with determinant one. This observation, together with (63) and Theorem 8, implies

$$\exp\left\{\int \operatorname{tr}\log W\,d\sigma\right\} = \inf_A \int \operatorname{tr}(A W A^*)\,d\sigma \leqslant \inf_A \int \operatorname{tr}\left[A(I+H)\,W\,(I+H)^*\right]d\sigma$$

$$= \inf_A \int \operatorname{tr}(A C A^*)\,d\sigma = \inf_{A_0} \operatorname{tr}(A_0 C A_0^*) = (\det C)^{1/n}.$$

(As usual A ranges over the analytic trigonometric polynomials whose constant term A_0 has determinant one, and C is the constant matrix of (63).) Let D be the positive definite square root of C. The analytic function B was defined, in the proof of Lemma 6, to be

$$B = (I+H)^{-1} D, \tag{77}$$

so that

$$D = (I+H)\,B.$$

Let $I+P_n$ be a sequence of trigonometric polynomials converging in L_W^2 to $I+H$, where each P_n is analytic without constant term. It is easy to see that $(I+P_n)\,B$ converges in L^2 to $(I+H)\,B$. The leading coefficient of each function $(I+P_n)\,B$ is B_0, and therefore the constant term in the Fourier series for $(I+H)\,B$ is also B_0. But this function is constant and equal to D, so that $B_0 = D$. The last string of inequalities can be continued then to give

$$\exp\left\{\int \operatorname{tr}\log W\,d\sigma\right\} \leqslant (\det C)^{1/n} = (\det B_0)^{2/n},$$

and this completes the proof of (75).

Theorem 9 is substantially the main result of Wiener's paper [19]. One must remark that the theorem is trivially false without the condition $\det B_0 \neq 0$. For if B is any singular matrix, then $W = B B^*$ is positive semi-definite and has a factorization (73), but (74) does not hold.

The interest of (75) lies in its function-theoretic meaning. If $B = B(e^{ix})$ is square-summable and of analytic type, then we can extend B to a matrix-valued analytic function defined *inside* the circle:

$$B(z) = B_0 + B_1 z + B_2 z^2 + \cdots \qquad (|z| < 1).$$

Obviously $\det B(z)$ is an analytic scalar function. It is not true, as one might expect, that (75) holds whenever $\det B(z)$ is free of zeros in the circle; (75) merely implies that $\det B(z)$ never vanishes. In the scalar case Beurling [5] has defined a factorization for analytic functions into an *inner function*, containing a Blaschke product and perhaps another factor, and an *outer function*, which has no zeros and satisfies the analogue of (75). Besides giving integral representations for these factors, Beurling characterized them by extremal properties. The theorem we have just proved extends Beurling's factorization to matrix-valued functions. For let $A = A(z)$ be an analytic square-summable function, perhaps with zeros. Then AA^* can be written on the boundary in the form BB^*, where B satisfies (75). In particular, $B(z)$ is non-singular inside the circle. Then $B^{-1}A$ is analytic inside the circle, and has exactly the properties of an inner function, while B is an outer function.

In (73) the order of the factors can just as well be reversed. It is clear that a symmetric development of Theorems 7 and 8 would give the other factorization; or we can derive the result directly as follows. If we set

$$\tilde{W}(e^{ix}) = W^*(e^{-ix}),$$

then

$$\tilde{W}(e^{ix}) = B(e^{-ix})\, B^*(e^{-ix}).$$

The function defined by

$$\tilde{B}(z) = B^*(\bar{z})$$

is analytic, and we have

$$\tilde{W} = \tilde{B}^* \tilde{B}.$$

But W and $\tilde{W}$ satisfy (74) at the same time, and so the two kinds of factorization exist together.

THEOREM 10. *Let F be a summable analytic matrix function whose constant term F_0 is non-singular. Then F has a factorization*

$$F = GH,$$

where G and H are square-summable analytic functions.

Let W be the positive semi-definite function $(FF^*)^{\frac{1}{2}}$. Using the proof of Theorem 2 as a model, and inequality (76), we can show that

$$\int \operatorname{tr} \log W\, d\sigma \geqslant \log |\det F_0|^{1/n} > -\infty.$$

By Theorem 9, W has a factorization

$$W = BB^*,$$

where B is the analytic function given by (77). We are going to show that $B^{-1}F$ is square-summable and analytic; then we only have to choose $G = B$ and $H = B^{-1}F$ to obtain the factorization desired.

$B^{-1}F$ is square-summable; for we have

$$(B^{-1}F)(B^{-1}F)^* = B^{-1}FF^*(B^*)^{-1} = B^{-1}W^2(B^*)^{-1} = B^*W(B^*)^{-1} = B^*B.$$

To see that $B^{-1}F$ is analytic we write from (77)

$$B^{-1}F = D^{-1}(I+H)F; \qquad I + H = \lim\, (I + P_n) \quad \text{in } L^2_W.$$

The Fourier coefficients of the components of $B^{-1}F$ are thus given for appropriate choice of constant matrices G by

$$\int \operatorname{tr}\,[G D^{-1}(I+H)F]\,e^{-ikx}\,d\sigma = \int \operatorname{tr}\,[G D^{-1}(I+H)W^2(F^{-1})^*]\,e^{-ikx}\,d\sigma =$$
$$= \int \operatorname{tr}\,[W(F^{-1})^* G D^{-1}(I+H)W]\,e^{-ikx}\,d\sigma. \qquad (78)$$

Now $W(F^{-1})^*$ belongs to L^2_W, since

$$F^{-1}WW(F^{-1})^*W = F^{-1}FF^*(F^{-1})^*W = W$$

is summable. Therefore the last integral in (78) is equal to the limit of similar integrals with $I+H$ replaced by $I+P_n$, so that the Fourier coefficients of the components of $B^{-1}F$ have the form

$$\lim \int \operatorname{tr}\,[G D^{-1}(I+P_n)F]\,e^{-ikx}\,d\sigma.$$

But F is an analytic function, so each such integral vanishes if $k<0$. This says exactly that the coefficients of $B^{-1}F$ vanish for negative indices, and that is what we had to prove.

In the proof we needed the hypothesis that F_0 is non-singular in order to show that $\log W$ is summable, so that W can be factored. The argument is still valid if $F_0 = 0$, provided the first non-vanishing coefficient is non-singular. On the other hand, it may happen that F_0 is different from zero but singular; in the extreme case $\det F$ may vanish everywhere. We do not know whether the theorem remains true in this generality.

We hope to have demonstrated that a certain class of theorems from function theory in the circle can be extended to matrix-valued functions by the method of this paper. It is not our intention to pursue this theory further. We conclude this section by stating the prediction theorem which follows from Theorem 8.

THEOREM 11. *Let $\{x_n^1\}$, $\{x_n^2\}$, ..., $\{x_n^N\}$ be N stationary processes in a Hilbert space, which are moreover mutually stationarily correlated in the sense that the scalar products*

$$(x_m^j,\ x_n^k)$$

depend only on j, k, and $n-m$. Denote by M the matrix-valued positive semi-definite measure with components m_{jk} whose Fourier-Stieltjes coefficients are

$$\int e^{-inx}\,dm_{jk}(x) = (x_n^j,\ x_0^k) \qquad (j,\ k = 1,\ \dots,\ N).$$

Let W be the absolutely continuous part of M. For any matrix $A=(a_{jk})$ of order N consider the elements

$$y^i = \sum_{k=1}^{N} a_{ik} x_0^k \qquad (i = 1, \ldots, N).$$

Denote the distance from y^i to the manifold spanned by all the elements x_n^k with k arbitrary and $n<0$ by γ_i. Then

$$\inf_A \frac{1}{N} \sum_{i=1}^{N} \gamma_i^2 = \exp\left\{\int \operatorname{tr} \log W \, d\sigma\right\}, \tag{79}$$

where A ranges over all constant matrices of determinant one.

The prediction error is once again given in (79) by the exponential expression from Szegö's Theorem, but the problem whose solution is thus given is complicated. We minimize the average square of error in approximating $y^1, \ldots, y^N$ by elements from the combined past of all the processes; and then we choose that set of elements $y^1, \ldots, y^N$ out of the present (obtained from $x_0^1, \ldots, x_0^N$ by means of a transformation with determinant one) for which this average error is as small as possible. It is not obvious that the infimum in (79) is attained, since the matrices of determinant one are not a compact set, but the proof of Theorem 8 shows that a minimal matrix exists.

We proceed to the proof of Theorem 11. First we have to show that a matrix measure M really exists with the properties ascribed to it. It is easy to see that for fixed j and k the sequence

$$(x_n^j, x_0^k)$$

is a linear combination of positive definite sequences, and so there exist complex-valued measures m_{jk} on the circle for which

$$\int e^{-inx} \, d\, m_{jk}(x) = (x_n^j, x_0^k). \tag{80}$$

Let M be the matrix with components m_{jk}. In order to show that M is positive semi-definite it suffices to prove for each set of complex continuous functions $\alpha_1(x), \ldots, \alpha_N(x)$ that

$$\int \sum_{j,k} \alpha_j \bar{\alpha}_k \, d\, m_{jk} \geq 0. \tag{81}$$

Clearly it is enough to prove this inequality when each α_j is a trigonometric polynomial:

$$\alpha_j(x) = \sum_n a_n^j e^{inx}.$$

The left side of (81) is equal to

$$\int \sum a_m^j \bar{a}_n^k e^{-i(n-m)x} \, d\, m_{jk}(x) = \sum a_m^j \bar{a}_n^k (x_{n-m}^j, x_0^k) = \sum a_m^j \bar{a}_n^k (x_{-m}^j, x_{-n}^k) \geq 0.$$

Thus (81) holds, and M is positive semi-definite.

By Theorem 8, we have

$$\exp\left\{\int \operatorname{tr}\log W\,d\sigma\right\} = \inf_A \int \operatorname{tr}\,[A^* A\,dM],$$

where A ranges over analytic trigonometric polynomials whose leading coefficient A_0 has determinant one. Let $A = (a_{jk})$ and

$$a_{jk}(e^{ix}) = \sum_{n \geq 0} a_{jk}^n e^{inx}. \tag{82}$$

For fixed A we have

$$\left.\begin{aligned}\int \operatorname{tr}\,[A^* A\,dM] &= \frac{1}{N}\int \sum_{i,j,k} a_{ij}\,\bar{a}_{ik}\,d m_{jk} = \\ &= \frac{1}{N}\sum_{\substack{i,j,k,\\ m,n}} a_{ij}^m\,\bar{a}_{ik}^n \int e^{-i(n-m)x}\,d m_{jk}(x) = \\ &= \frac{1}{N}\sum_{\substack{i,j,k,\\ m,n,}} a_{ij}^m\,\bar{a}_{ik}^n\,(x_{-m}^j, x_{-n}^k) = \\ &= \frac{1}{N}\sum_i \Big\| \sum_{k,n} a_{ik}^n\,x_{-n}^k \Big\|^2.\end{aligned}\right\} \tag{83}$$

Set

$$y^i = \sum_k a_{ik}^0\,x_0^k \qquad (i = 1, \ldots, N)$$

and recall from (82) that n is summed over non-negative integers in (83). Then the individual terms in the last expression of (83) have the form

$$\|y^i - x\|^2, \tag{84}$$

where x is a linear combination of elements x_n^k for $k = 1, \ldots, N$ and $n < 0$. The numbers a_{ik}^0 are the components of the leading coefficient A_0 of A, and are subject to the requirement that $\det A_0 = 1$. The other coefficients entering into (83) are completely arbitrary, since the other coefficients $A_1, A_2, \ldots$ of A are arbitrary. Therefore the elements x which can appear in (84) are *arbitrary* linear combinations of x_n^k with $n < 0$, and it follows that the infimum over A of (83) is the left side of (79). This completes the proof of the theorem.

6. Extensions of the Theorems

The proof of the Szegö Theorem is valid for functions of more than two variables. Indeed, its natural setting is a compact abelian group K whose dual $\hat{K}$ has a complete linear order compatible with the group structure. (The theory of the Fourier transform on locally compact abelian groups is contained in [12], and this section presupposes some acquaintance with the generalized Fourier transform.) To be precise, $\hat{K}$ is to contain a distinguished subset $\hat{P}$ with the properties

1° $\hat{0} \notin \hat{P}$ (where $\hat{0}$ is the identity in $\hat{K}$)

2° $\hat{x} \in \hat{P}$ if and only if $-\hat{x} \notin \hat{P}$, unless $\hat{x} = \hat{0}$

3° $\hat{x} \in \hat{P}$ and $\hat{y} \in \hat{P}$ imply $\hat{x} + \hat{y} \in \hat{P}$.

The elements of $\hat{P}$ are called *positive.* A total order is defined in $\hat{K}$ by saying

$$\hat{x} > \hat{y} \quad \text{just if} \quad \hat{x} - \hat{y} \in \hat{P}.$$

Let $d\sigma$ be the invariant measure in K, normalized so that K carries total mass one. The value of a character $\hat{x}$ at the point x of K is written $(x, \hat{x})$. With these conventions we have the following generalization of Theorems 1 and 8.

THEOREM 12. *Let M be a positive semi-definite matrix-valued measure defined on Borel subsets of K. Suppose that*

$$dM = W\, d\sigma + dM_s,$$

where W is summable for $d\sigma$ and M_s is singular with respect to $d\sigma$. Then

$$\exp \left\{ \int \operatorname{tr} \log W\, d\sigma \right\} = \inf_{A_0, P} \int \operatorname{tr} \left[(A_0 + P)^* (A_0 + P)\, dM \right],$$

where A_0 ranges over the matrices with determinant one, and P over finite sums

$$P(x) = \sum_{\hat{x} \in \hat{P}} A(\hat{x})(x, \hat{x}).$$

The left side of (78) *is to be interpreted as zero if*

$$\int \operatorname{tr} \log W\, d\sigma = -\infty.$$

We also have analogues of most of the other theorems of the preceding sections. Recall that the Fourier transform of a summable function f on K is the function $\hat{f}$ on $\hat{K}$ defined by

$$\hat{f}(\hat{x}) = \int \overline{(x, \hat{x})}\, f(x)\, d\sigma.$$

We shall say that f *is of analytic type* if $\hat{f}(\hat{x}) = 0$ for all $\hat{x}$ such that $-\hat{x} \in \hat{P}$.

THEOREM 13. *Let f be a summable function of analytic type on K. Then*

$$\int \log |f|\, d\sigma \geq \log |\hat{f}(\hat{0})|.$$

THEOREM 14. *Let W be a summable positive semi-definite matrix function on K. A necessary and sufficient condition that W have a representation in the form*

$$W = BB^*,$$

where B is square-summable, of analytic type, and satisfies $\det \hat{B}(\hat{0}) \neq 0$, *is that*

$$\int \operatorname{tr} \log W \, d\sigma > -\infty.$$

If this condition is satisfied, then we can choose B so that

$$\int \log |\det B(x)| \, d\sigma = \log |\det \hat{B}(\hat{0})|.$$

THEOREM 15. *Let F be a summable matrix function of analytic type with* $\det \hat{F}(\hat{0}) \neq 0$. *Then F has a factorization GH, where G and H are square-summable matrix functions of analytic type.*

THEOREM 16. *Let μ be a complex Borel measure on K whose Fourier-Stieltjes coefficients satisfy*

$$\int \overline{(x, \hat{x})} \, d\mu(x) = 0 \qquad \text{for } \hat{x} \in \hat{P}.$$

Then the coefficients of its singular and absolutely continuous parts separately have the same property.

THEOREM 17. *Let f be a square-summable function of analytic type, and denote by C_f the closed subspace of L^2 spanned by linear combinations of f and functions of the form*

$$\hat{x} \cdot f \qquad (\hat{x} \in \hat{P}).$$

C_f is identical with the set of all square-summable functions of analytic type if and only if

$$\int \log |f| \, d\sigma = \log |\hat{f}(\hat{0})| > -\infty.$$

The special properties of the circle and torus groups have been used very sparingly in the proofs of the preceding sections. In order to prove the theorems just enunciated, we have only to make notational changes on previous pages. It should be observed that Theorem 8 depended on an analogous theorem for scalar functions, so that a scalar version of Theorem 12 has to be given before considering matrix-valued functions. Theorem 16 incorporates that part of the Bochner Theorem which lends itself to elegant statement. The full theorem depends on the comparison of two different order relations in $\hat{K}$, and this is cumbersome to explain in general terms.

We should dwell on one technical point. The properties of the Fejér means of a Fourier series were used in the proof of Lemma 2, Theorem 2, and the corresponding theorems about matrices. On each compact group K there are approximate identities in abundance, and they can even be chosen from the trigonometric polynomials. However, unless K is separable an approximate identity will have to be a *directed system* of functions rather than a simple sequence. For any fixed function f summable on K we can choose from any approximate identity (consisting of trigonometric polynomials) a sequence $e_1, e_2, \ldots$ having the properties we need, even though the sequence depends in general on f. For each n, $e_n * f$ is a trigonometric polynomial with the same mean value as f, and the sequence converges to f in the norm of summable functions (as required for Theorem 2); and if f is bounded, $e_n * f$ has at worst the same bounds as f, and a subsequence converges almost everywhere to f (which is enough to prove Lemma 2). The Fejér kernels have deeper properties than these, but we have not used them.

The compact group with ordered dual seems to be the natural domain on which to state theorems like those of this paper. It is not so obvious what the range of the functions considered should be. Probably our main theorems hold for functions taking values in an algebra of operators on Hilbert space possessing trace and determinant functions of the appropriate kind. The existence and classification of trace functions was a main object of the work of Murray and von Neumann, and more recently Fuglede and Kadison [9] have given a theory of determinants. Unfortunately, only positive definite operators seem to deserve a determinant. Reference to the determinant of other matrices can be avoided in many parts of section five, but there are other places (especially in the proof of (68)) where that is more difficult.

References

[1]. N. I. Achieser, *Vorlesungen über Approximationstheorie*. Berlin. 1953.

[2]. R. Arens, A Banach algebra generalization of conformal mappings of the disc. *Trans. Amer. Math. Soc.* 81 (1956), 501–513.

[3]. R. Arens & I. M. Singer, Function values as boundary integrals. *Proc. Amer. Math. Soc.* 5 (1954), 735–745.

[4]. R. Arens & I. M. Singer, Generalized analytic functions. *Trans. Amer. Math. Soc.* 81 (1956), 379–393.

[5]. A. Beurling, On two problems concerning linear transformations in Hilbert space. *Acta Math.* 81 (1949), 239–255.

[6]. S. Bochner, Boundary values of analytic functions in several variables and of almost periodic functions. *Ann. of Math.* 45 (1944), 708–722.

[7]. A. P. Calderón & A. Zygmund, Singular integrals and periodic functions. *Studia Math.* 14 (1954), 249–271.

[8]. J. L. Doob, *Stochastic Processes*. New York, 1953.

[9]. B. Fuglede & R. V. Kadison, Determinant theory in finite factors. *Ann. of Math.* 55 (1952), 520–530.

[10]. G. H. Hardy & J. E. Littlewood, Some new properties of Fourier constants. *Math. Ann.* 97 (1927), 159–209.

[11]. A. N. KOLMOGOROFF, Stationary sequences in Hilbert's space (Russian). *Bolletin Moskovskogo Gosudarstvennogo Universiteta. Matematika* 2 (1941). (Reference from Mathematical Reviews.)
[12]. L. H. LOOMIS, *Abstract harmonic analysis*. New York, 1953.
[13]. G. MACKEY, The Laplace transform for locally compact Abelian groups. *Proc. Nat. Acad. Sci.* 34 (1948), 156–162.
[14]. N. WIENER & P. MASANI, The prediction theory of multivariate stochastic processes. I. The regularity condition. *Acta Math.*, 98 (1957), 111–150. II. The linear predictor. *Acta Math.*, 99 (1958), 93–137.
[15]. F. & M. RIESZ, Über die Randwerte einer analytischen Funktion. *Quatrième Congrès des math. scandinaves*, 1916, 27–44.
[16]. G. SZEGÖ, Beiträge zur Theorie der toeplitzschen Formen (Erste Mitteilung). *Math. Z.* 6 (1920), 167–202.
[17]. G. SZEGÖ, Über die Randwerte einer analytischen Funktion. *Math. Ann.* 84 (1921), 232–244.
[18]. N. WIENER, *Extrapolation, interpolation, and smoothing of stationary time series*. New York, 1949.
[19]. N. WIENER, On the factorization of matrices. *Comment. Math. Helv.* 29 (1955), 97–111.
[20]. A. ZYGMUND, *Trigonometrical series*. Warsaw, 1935.