PREDICTION THEORY OVER DISCRETE ABELIAN GROUPS

BY

D. M. EAVES

1. **Introduction.** Classical least squares linear prediction theory is concerned with a *stationary stochastic process* (SSP); that is, with a family $X(n)$ $(n=0, 1, -1, \ldots)$ of complex-valued random variables on a probability space (Ω, P), that have zero means, Gaussian joint distributions, and finite covariances $\langle X(n), X(m) \rangle$ depending only on $n-m$. One accomplishment of the theory is an analytical characterization of those SSPs for which some $X(n)$ differs from its conditional expectation relative to those $X(m)$ with $m < n$, the so-called *nondeterministic* SSPs. A second feature of the classical theory is an analytical characterization of the nullness or nonnullness of the *remote past*, that is, of the intersection (over all integers n) of the closed spans in $L_2(\Omega, P)$ of families, $\{X(m) \mid m \leq n\}$. It is the generalization of these two results, replacing the integers with an arbitrary discrete Abelian group, to which this paper is directed.

The classical characterizations referred to are given in terms of the *spectral measure* of the SSP; that is, the unique regular Borel measure $\mu \geq 0$ in the unit circle G, satisfying

$$\langle X(n), X(0) \rangle = \int_G \chi_n(x) \, d\mu(x), \qquad (n = 0, 1, -1, \ldots).$$

(We shall represent G in the form $\{x \mid 0 \leq x < 2\pi\}$, with addition modulo 2π and the appropriate topology; the character χ_n of G is given by $\chi_n(x) = e^{inx}$.) Since the joint distributions are Gaussian, taking conditional expectations relative to a family of random variables is the same as taking orthogonal projections in $L_2(\Omega, P)$ onto the closed span of the family. (See [2, Chapter 3].) Hence the classical analytical criterion for nondeterminacy is given by the following version of Szegö's Theorem:

THEOREM 1.1. *Let $\mu > 0$ be a finite regular Borel measure in the unit circle G; let $d\mu/d\sigma$ be the Radon-Nikodym derivative of its absolutely continuous part relative to Lebesgue measure σ in G; for each k let P_k be the orthogonal projection of $L_2(\mu)$ onto the closed span of all χ_m with $m \leq k$. Then for all n,*

$$\|\chi_n - P_{(n-1)}(\chi_n)\|^2 = \exp \int_G \log \frac{d\mu}{d\sigma} \, d\sigma$$

the norm being that in $L_2(\mu)$.

If μ is the spectral measure of a SSP X of random variables in some $L_2(\Omega, P)$ (as it always is), then this norm is also that of the difference between $X(n)$ and its orthogonal projection in $L_2(\Omega, P)$ onto the closed span of those $X(m)$ with $m \leq n-1$. In other words X is nondeterministic if and only if

$$(1) \qquad \int_G \log \frac{d\mu}{d\sigma} \, d\sigma \; > \; -\infty.$$

The classical criterion for nullness of the remote past of a nondeterministic SSP X is as follows: If μ is the spectral measure of X, then X has null remote past if and only if μ is absolutely continuous relative to σ. Actually we will go somewhat farther than a generalization of this fact, and generalize the following classical facts about X, a nondeterministic SSP:

(2) $X = N + D$ where N is a SSP with null remote past; D is a deterministic SSP; N and D have mutually orthogonal ranges. Letting μ, μ_N, and μ_D respectively be the spectral measures of X, N, and D, $\mu = \mu_N + \mu_D$ is the Lebesgue decomposition of μ (with respect to Lebesque measure σ in the circle).

(3) X has null remote past if and only if

$$X(k) = \sum_{n \geq 0} b_n Z(k-n), \qquad b > 0,$$

where Z is an SSP with Lebesgue measure for its spectral measure, $\sum_n |b_n|^2 < \infty$, each $b_n Z(k-n)$ is the orthogonal projection of $X(k)$ onto the orthogonal complement in the closed span of $X(\{j \mid j \leq k-n\})$, of the closed span of $X(\{j \mid j < k-n\})$;

$$\mu = \left| \sum_{n \geq 0} b_n \Delta^n \right|^2 \sigma,$$

where $\Delta(x) = e^{-ix}$ $(0 \leq x < 2\pi)$. μ is of course absolutely continuous with respect to Lebesgue measure in this case.

In this paper we will postpone a generalization of the nondeterminacy condition (1) until §6 and proceed immediately to generalizations of conditions (2) and (3) for a nondeterministic SSP, as follows: In §2 we first generalize the concept of a SSP and then translate the problem into the neater language of $L_2(\mu)$ where μ is a positive (spectral) measure. We then carry over some basic probabilistic definitions and results from the classical case to the setting where the integers are replaced with a discrete Abelian group $\hat{G}$ and the nonpositive integers are replaced with a proper subsemigroup S of $\hat{G}$ which is, in a sense, not too thin. Definitions generalizing the notions of nondeterminacy and remote past are introduced, relative to the semigroup S. These are used to obtain appropriate generalizations of the SSPs N and D of condition (2).

In §3 and thereafter we fix the semigroup S as the nonpositive half $\{\chi \in \hat{G} \mid \psi(x) \leq 0\}$ of $\hat{G}$ with respect to an arbitrary nontrivial real-valued homomorphism ψ having image $\psi(\hat{G})$ nondense in the reals. In §3 a moving average representation like the one above in condition (3) is developed, for an SSP X

having null remote past. This representation will show that the group structure of interest resulting from our notion of the past is that of a one-parameter subgroup $\{\Delta^n \mid n=0, 1, -1, \ldots\}$ passing through S, rather than the group structure of $\hat{G}$ itself. (We use the multiplicative notation in $\hat{G}$.) This situation is in contrast with another generalization (see [3]) of the classical theory, in which the harmonic analysis of $\hat{G}$ itself plays an essential rôle, but which is devoid of physical concepts resembling the past or remote past. In §§3 and 4 we generalize the remaining part of conditions (2) and (3). (2) and (3) together show that the nondeterministic SSPs with null remote pasts are precisely those with "smooth" (absolutely continuous relative to Haar measure) spectral measures.

In §4 it turns out that the appropriate smoothness condition characterizing our notion of a null remote past is *absolute continuity in a given direction* (in a direction determined by Ψ in a way described in §4). This smoothness condition for complex measures in compact Abelian groups was studied in [2] by de Leeuw and Glicksberg. In §5 we determine the analogue of Szegö's Theorem appropriate to the present context; this result is then used to study the structure of nondeterministic SSPs and to provide the means for constructing examples.

The author is indebted to Professor Ramesh Gangolli for suggesting directional absolute continuity as the appropriate analogue of absolute continuity in the circle.

2. Nondeterminacy and the remote past.

A *stationary stochastic process* (SSP) over a discrete Abelian group $\hat{G}$ is a map X from $\hat{G}$ into a Hilbert space $H \neq \{0\}$ such that H is the closed span of $X(\hat{G})$ and such that the inner product $\langle X(\chi), X(\Delta) \rangle$ depends only on $\Delta^{-1}\chi$. $\langle X(\chi), X(1) \rangle$ is then a positive definite function of $\chi \in \hat{G}$, in the sense that any square matrix with entries $\langle X(\chi_i), X(\chi_j) \rangle$ (given $\chi_1, \chi_2, \ldots$) is nonnegative definite. Equivalently

$$\sum f(\chi)\overline{f(\Delta)}\langle X(\chi), X(\Delta) \rangle \geqq 0$$

whenever $f \in L_1(\hat{G})$ (summing over those countably many pairs χ, Δ where $f \neq 0$). According to Bochner's Theorem [5, pp. 126, 142], since $\langle X(\chi), X(1) \rangle$ is positive definite there is a unique finite Baire measure $\mu \geqq 0$ in the compact group G dual to $\hat{G}$, satisfying

$$\langle X(\chi), X(1) \rangle = \int_G \chi \, d\mu \quad \text{for all } \chi \in \hat{G}.$$

(Here the character χ is regarded as a measurable function on G.) μ is called the *spectral measure* of the given SSP. As in the classical case, there is no loss (or gain) of generality in assuming that $H = L_2(\mu)$ where μ is the spectral measure of the given SSP, and that $X(\chi) = \chi \in L_2(\mu)$ ($\chi \in \hat{G}$). Accordingly we shall speak of a measure $\mu > 0$ in G (rather than an SSP over $\hat{G}$) as having various properties attributable to a SSP.

Throughout this section let S be a proper subsemigroup of $\hat{G}$ containing the identity 1. Let $S^\circ \subseteq S$ consist of those $\chi \in S$ such that for every $\Lambda \in \hat{G}$ eventually $\chi^n \Lambda \in S$ $(n=0, 1, \ldots)$; we assume S° nonempty. A measure $\mu > 0$ in G is said to be *deterministic* with respect to S if for some $\Delta \in S^\circ$, some character $\chi \in \hat{G}$ lies in the closed subspace of $L_2(\mu)$ generated by $\chi \Delta S$. For a subset A of $\hat{G}$ let σA be the subspace of $L_2(\mu)$ generated by A. We omit the easy proofs of the following lemma and theorem.

LEMMA 2.1. $\bigcap_{n>0} \sigma \Delta^n S$ *does not depend on the choice of* $\Delta \in S^\circ$.

This $\bigcap_{n>0} \sigma \Delta^n S$ is called the *remote past* $\mathscr{R}$ *of the measure* μ *relative to* S.

THEOREM 2.2. *The following are equivalent*:
(a) $\sigma S = \mathscr{R}$;
(b) $\sigma S = \sigma \Delta S$ *for some* $\Delta \in S^\circ$;
(c) $\sigma S = \sigma \Delta S$ *for each* $\Delta \in S^\circ$;
(d) μ *is deterministic relative to* S;
(e) $\sigma S = \sigma \hat{G}$.

The proofs of the remaining results stated in this section are similar to those in the classical case, where S is taken to be the semigroup χ_n $(n=0, -1, \ldots)$ of characters of the unit circle. Inasmuch as the proofs can be deduced from Doob's presentation [1, Chapter XII], we omit them. For each $\chi \in \hat{G}$ let P_χ be the orthogonal projection of $L_2(\mu)$ onto $\sigma \chi S$. Fixing an arbitrary $\Delta \in S^\circ$ for the remainder of this section, write $Q_n = P_{\Delta^n} - P_{\Delta^{(n+1)}}$ for each integer n. Let D be the orthogonal projection of $L_2(\mu)$ onto $\mathscr{R}$, let N be the orthogonal projection onto its orthogonal complement $\mathscr{R}^\perp$.

THEOREM 2.3. $N(1) = \sum_{n \geq 0} Q_n(1)$; *the series converges in* $L_2(\mu)$.

THEOREM 2.4. *The maps* $\chi \to N(\chi) \in L_2(\mu)$ *and* $\chi \to D(\chi) \in L_2(\mu)$ *are SSPs over* $\hat{G}$; $N(\chi)$ *has null remote past relative to* S, *while* $D(\chi)$ *is deterministic relative to* S.

We will denote by μ_N and μ_D the spectral measures of $N(\chi)$ and $D(\chi)$.

LEMMA 2.5. *For any* $g \in L_2(\mu)$, $N(g) = 0$ *ae* $-\mu_D$ *and* $D(g) = 0$ *ae* $-\mu_N$; $\mu = \mu_N + \mu_D$; μ_N *and* μ_D *have disjoint measurable supports.*

3. **The moving average.** For the remainder of this paper G is a compact Abelian group; by a *measure* in G we mean a finite complex regular Baire measure; μ will denote a positive measure in G. Except where otherwise noted, μ will be assumed nondeterministic relative to a subsemigroup S of the dual group $\hat{G}$. (We will define the scope of S shortly.) For the time being q will denote any nonnull element of $L_2(\mu)$. Thus the map $\chi \to T_\chi(q) = \chi q$ from $\hat{G}$ into $L_2(\mu)$ is a SSP with spectral measure $|q|^2 \mu$. N and D will be as in §2. Let U be the orthogonal projection of $L_2(\mu)$ onto the closed span $\mathscr{M}$ of all the χq for $\chi \in \hat{G}$, let V be the orthogonal projection onto the orthogonal complement of $\mathscr{M}$. Then $T_\chi U = U T_\chi$ for all χ, so that

the maps $\chi \to U(\chi)$ and $\chi \to V(\chi)$ are SSPs over $\hat{G}$. Let μ_U and μ_V be their spectral measures. Write $\eta = |q|^2\mu$. The following lemma follows from the Stone-Weierstrass theorem, once it is seen that $\mu = \mu_U + \mu_V$ and $q = 0$ $ae - \mu_V$; the subsequent corollary and theorem are then straightforward.

LEMMA 3.1. *The map $f \to fq$ is an isometry from $L_2(\eta)$ onto $L_2(\mu_U)$.*

Henceforth q will denote a particular measurable representative of an equivalence class in $L_2(\mu)$, instead of denoting the equivalence class itself. $1/q$ will denote the function g, where $g(x) = 1/q(x)$ if $q(x) \neq 0$ and $g(x) = 0$ if $q(x) = 0$.

COROLLARY 3.2. $\mu_U = \eta/|q|^2$ *and* $\eta = |q|^2\mu_U$.

THEOREM 3.3. $\mu = \mu_U + \mu_V$ *is the Lebesgue decomposition of μ relative to η.* ($\mu_U \ll \eta$.)

For the remainder of the paper fix $\Psi \not\equiv 0$ an arbitrary real-valued homomorphism of $\hat{G}$ with nondense image, let $S = \{\chi \in \hat{G} \mid \Psi(\chi) \leq 0\}$. Choose Δ so that $\Psi(\Delta)$ is the largest negative value of Ψ, and let Δ be fixed throughout. Further, we now take $Q_0(1)$ for the q discussed above. Thus $\mathcal{M}$ becomes the closed span of all $\chi Q_0(1)$ for $\chi \in \hat{G}$, and $\eta = |Q_0(1)|^2\mu$. η will denote $|Q_0(1)|^2\mu$ throughout the paper.

LEMMA 3.4. *For χ such that $\Psi(\chi) = j\psi(\Delta)$, $\chi Q_k = Q_{k+j}\chi$ for all k and j. (Regarding the χ as unitary transformations of $L_2(\mu)$.)*

Proof. The relations $\chi P_{\Delta^k} = P_{\Delta^{k+j}}\chi$ can be checked; first on the range of P_{Δ^k}, then on its orthogonal complement. These relations follow from the condition $\chi S = S$.

LEMMA 3.5. $\mathcal{M} = \mathcal{R}^\perp$; *hence $\mu_U = \mu_N$ and $\mu_V = \mu_D$.*

Proof. We omit the details. The inequality $\subseteq$ is obvious; the first step in establishing $\supseteq$ is to show (using Theorem 2.3) that any $g \in \mathcal{R}^\perp$ has the form $g = \sum_k Q_k(g)$. The proof now follows from the Lemma 3.4.

Let Γ be the kernel $\Psi^{-1}(\{0\})$ of Ψ. We may assume henceforth that $\hat{G}$ is the direct product of Γ with the integer group Z, in view of the isomorphism

$$\chi \to (-\Psi(\chi)/\Psi(\Delta), \chi\Delta^{-\Psi(\chi)/\Psi(\Delta)})$$

from our heretofore abstract $\hat{G}$ onto $Z \times \Gamma$. (We shall not bother, in what follows, to trace the trivial effect of the choice of Δ upon this representation or upon subsequent results. All of our results could easily be expressed, as most of them are, independently of these representations.) Due to the duality theorem, the representation $G = T \oplus H$, where T is the circle group, is implicit in the above assumption. We shall write $T = [0, 2\pi)$ with the appropriate topology, and denote the group operations in T, H, and G all with $+$. 1 will denote the identity of all multiplicatively written groups; 0 will denote the identity of additive groups. We shall represent Z multiplicatively as $\{\chi_n : n = 0, \pm 1, \ldots\}$, where $\chi_n(t) = e^{int}$ for $t \in T$, and use the internal direct product (sum) notation in $\hat{G}$ (in G). Thus we write $\chi_n\gamma$ for a typical

element of $\hat{G}$, $\Gamma = \{1\} \times \Gamma$, and $s + h = (s, h) \in T \oplus H$. σ will denote normalized Lebesgue measure in T. π will be the canonical projection of G onto H. If τ is a measure on G, define the measure $\tau\pi^{-1}$ in H by $\tau\pi^{-1}(M) = \tau(\pi^{-1}(M))$ for all Baire sets $M \subseteq H$. A computation of Fourier coefficients shows that if $\hat{\tau}$ is supported on Γ then $\tau = \sigma \times \tau\pi^{-1}$.

THEOREM 3.6. *Given* $g \in L_2(\mu)$ *and* $\chi_k\gamma \in \hat{G}$, *there is a unique "coefficient"* b_k, *an element of* $L_2(\eta)$ *in the closed span of* Γ, *such that* $Q_k(g) = b_k Q_k(\Delta^k)$.

Proof. Lemma 3.4 implies that the range of Q_0 is the $L_2(\mu)$-closed span of $\Gamma Q_0(1)$. This and the relation $\Delta^{-k} Q_k = Q_0 \Delta^{-k}$ imply that $s_n Q_0(1) \xrightarrow{n} \Delta^{-k} Q_k(g)$ in $L_2(\mu)$, for some sequence s_n of finite linear sums of elements of Γ. This implies s_n is $L_2(\eta)$-Cauchy; hence $s_n \xrightarrow{n} b_k$ in $L_2(\eta)$ for some $b_k \in L_2(\eta)$. (b_k does not depend on the choice of the sequence s_n.) But this is precisely the condition $s_n Q_0(1) \xrightarrow{n} b_k Q_0(1)$ in $L_2(\mu_U)$, in view of Lemma 3.1. Hence $\Delta^{-k} Q_k(g) = b_k Q_0(1) ae - \mu_U$. By Lemma 2.5 $\Delta^{-k} Q_k(g) = 0 = b_k Q_0(1) ae - \mu_D$; thus $\Delta^{-k} Q_k(g) = b_k Q_0(1) ae - \mu$; hence $Q_k(g) = b_k Q_k(\Delta^k) ae - \mu$.

THEOREM 3.7 (THE MOVING AVERAGE). *For each* $\chi \in \hat{G}$, $N(\chi) = \sum_{n \geq 0} b_n \chi Q_k(\Delta^k) ae - \mu$; *each* $b_n \in L_2(\eta)$ *is in the* $L_2(\eta)$-*closed span of* Γ; $b_0 = 1$ $ae - \eta$;

$$\sum_{n \geq 0} \int_G |b_n|^2 \, d\eta < \infty.$$

Proof. It is enough to consider the case $\chi \equiv 1$. Taking $g \equiv 1$ in Theorem 3.6 and applying Theorem 2.3 yields the result.

Theorems 3.3, 3.7 and Corollary 3.2 immediately yield the following

THEOREM 3.8. $\mu_N = |\sum_{n \geq 0} b_n \Delta^n|^2 \eta$; $\mu = \mu_N + \mu_D$ *is the Lebesgue decomposition of* μ *relative to* η; $\eta = |Q_0(1)|^2 \mu_N$.

Proof. By definition of μ_N, for every $\chi \in \hat{G}$

$$\int_G \chi \, d\mu_N = \int N(\chi) N(1) \, d\mu.$$

Replacing $N(\chi)$ and $N(1)$ with their moving average representations shows that this integral coincides with $\int_G \chi |\sum_{n \geq 0} b_n \Delta^n|^2 \, d\eta$ proving the first assertion.

4. Null remote pasts and smoothness. Thus far we have achieved at least a formal generalization of conditions (2) and (3) (see §1) of the classical theory, to the present setting. The smoothness condition on a positive measure μ, corresponding to our notion of a null remote past, has turned out to be absolute continuity relative to $\eta = |Q_0(1)|^2\mu$. It is a convenient coincidence that in the classical theory this measure is precisely the invariant measure for G, thus giving us a smoothness condition on μ defined intrinsically in terms of G. In our new general setting, however, one can construct μ so as to produce an arbitrary positive measure η

having Fourier transform supported on Γ. There are precisely as many such measures η as there are positive measures on the dual group H, of Γ. Since this is in general a large class of measures (although not in the classical case, where H is a singleton), we need a more direct and natural description of this "smoothness". We must now seek an intrinsic characterization of the condition $\mu \ll \eta$.

A Baire set $E \subseteq G$ is said to be null in the direction Ψ or Ψ-*null* if $T \cap (E - x)$ has zero linear Lebesgue measure for every $x \in G$. A measure τ is called Ψ-*absolutely-continuous* if its total variation at every Ψ-null set is zero. This coincides with the notion of directional absolute continuity examined by De Leeuw and Glicksberg in [2]. Any measure of the form $\sigma \times \alpha$ (α a measure in H) is Ψ-absolutely-continuous, since $\int_T g(s+h) \, ds = 0$ for all $h \in H$, whenever g is the indicator function of a Ψ-null set.

Henceforth terms such as "remote past", "determinacy", etc., will be understood to convey concepts relative to the (fixed) homomorphism Ψ. A measure $\mu > 0$ in G is said to be *strongly nondeterministic* if $f\mu$ is nondeterministic whenever $f \geq 0$ is a bounded Baire function on G constant on each coset $x + T$ ($x \in G$), such that $\int f \, d\mu > 0$. Clearly μ is strongly nondeterministic if and only if, for every bounded complex valued Baire function f nonnull in $L_2(\mu)$ and constant on each $x + T$, the SSP $\chi \to f\chi \in L_2(\mu)$ is nondeterministic.

THEOREM 4.1. *μ is strongly nondeterministic if and only if $\mu\pi^{-1} \ll \eta\pi^{-1}$.*

Proof. To check whether, for a particular f, $f\mu$ is nondeterministic, we need only check whether the $L_2(f\mu)$-distance from 1 to the closed span of ΔS is positive (Theorem 2.2).

Now $1 = Q_0(1) + (1 - Q_0(1))$ yields the two components of 1 that are, respectively, orthogonal in $L_2(f\mu)$ to the span of ΔS, and within the $L_2(f\mu)$-span of ΔS: One can let s_n be a sequence of finite linear sums of elements of Γ, $s_n \xrightarrow{n} \sqrt{f}$ in $L_2(\mu)$. (Take s_n so that $s_n \xrightarrow{n} \sqrt{f}$ in $L_2(H, \mu\pi^{-1})$, treating both the s_n and $\sqrt{f}$ as functions on H.) Hence for any

$$\chi \in \Delta S, \quad \int_G \chi \overline{Q_0(1)} f \, d\mu = \lim_n \int_G \chi s_n \overline{Q_0(1)s_n} \, d\mu = 0.$$

In other words $Q_0(1)$ is orthogonal in $L_2(f\mu)$ to the span of ΔS. Furthermore $1 - Q_0(1) = P_\Delta(1)$ is the $L_2(\mu)$-limit of a sequence s_n of finite linear sums of elements of ΔS; hence

$$\lim_n \int_G |P_\Delta(1) - s_n|^2 f \, d\mu \leq \lim_n \text{SUP}(f) \int_G |P_\Delta(1) - s_n|^2 \, d\mu = 0.$$

Thus $\int_G f |Q_0(1)|^2 \, d\mu = \int_G f \, d\eta$ is the above-named $L_2(f\mu)$-distance.

The problem has now become that of showing $\int_G f \, d\mu > 0$ implies $\int_G f \, d\eta > 0$ for bounded f constant on each $T + x$; but this is immediate from the fact that these two quantities are, respectively, $\int_H f \, d(\mu\pi^{-1})$ and $\int_H f \, d(\eta\pi^{-1})$.

LEMMA 4.2 (THE DIRECTIONAL LEBESGUE DECOMPOSITION). *Given a measure α in G, there exist unique measures α_a and α_s such that $\alpha = \alpha_s + \alpha_s$, α_a is Ψ-absolutely-continuous, and α_s is supported on a Ψ-null set.*

Proof. The construction of α_a can be done exactly as in the traditional Lebesgue decomposition: one removes from the support of a α a Ψ-null Baire set having maximal positive total variation under α.

Let μ be nondeterministic and let $A \subseteq H$ be a Baire set such that $M \rightarrow \mu\pi^{-1}(M \cap A)$ is the absolutely continuous part of $\mu\pi^{-1}$ relative to $\eta\pi^{-1}$. Define $\mu_A(E) = \mu(E \cap \pi^{-1}(A))$ for Baire subsets E of G.

LEMMA 4.3. $(\mu_A)_a = \mu_N$.

Proof. $(\mu_A)_a \geqq \mu_N$, since $\mu \geqq \mu_N$ and $(\mu_A)_a$ was obtained from μ by removing only η-null sets (which are therefore μ_N-null by Theorem 3.8) from the support of μ: by removing $\pi^{-1}(H \sim A)$ to obtain μ_A and then removing a Ψ-null set to obtain $(\mu_A)_a$.

To show $(\mu_A)_a \leqq \mu_N$ it is enough to show $(\mu_A)_a \ll \mu_N$, for then μ_N exceeds $(\mu_A)_a$ by the absolutely continuous part of $\mu - (\mu_A)_a$ relative to μ_N. Therefore we consider any μ_N-null set S, with indicator function I_s. By Theorem 3.8, $0 = \eta(S) = \int_H \int_T I_s(t+h)\, d\sigma(t)\, d(\eta\pi^{-1})(h)$. Thus $\eta\pi^{-1}(N) = 0$, where

$$N = \left\{ h \in H : \int_T I_s(t+h)\, dt > 0 \right\},$$

so that

$$0 = \mu\pi^{-1}(N \cap A) = \mu_A(\pi^{-1}(N)) \geqq (\mu_A)_a(S \cap \pi^{-1}(N)) = 0.$$

It remains only to show $(\mu_A)_a(S \sim \pi^{-1}(N)) = 0$. For this we need only show that $S \sim \pi^{-1}(N)$ is Ψ-null; i.e., that $T \cap ((S \sim \pi^{-1}(N)) - x)$ has zero linear Lebesgue measure in T, for each $x \in G$. This set is in fact empty if $x \in \pi^{-1}(N)$, while if $x = s + \pi(x) \notin \pi^{-1}(N)$, its Lebesgue measure is given by

$$\int_T I_{S \cap \pi^{-1}(H \sim N)}(t+s+\pi(x))\, d\sigma(t) = \int_T I_{S \cap \pi^{-1}(H \sim N)}(t+\pi(x))\, d\sigma(t)$$

$$\leqq \int_T I_s(t+\pi(x))\, d\sigma(t);$$

this is zero since $\pi(x) \in N$.

THEOREM 4.4. *μ has null remote past if and only if it is Ψ-absolutely-continuous and strongly nondeterministic.*

Proof. If μ has null remote past then $\mu \ll \eta = |Q_0(1)|^2 \mu$ in view of Theorem 3.8. The relation $\eta = \sigma \times \eta\pi^{-1}$ implies that for every Ψ-null set E $\eta(E) = 0$, and so $\mu(E) = 0$; this means μ is Ψ-absolutely-continuous. That μ is strongly nondeterministic is the "if" part of Theorem 4.1.

Conversely, suppose μ is Ψ-absolutely-continuous and strongly nondeterministic. Then $\mu_A = \mu$, since $\mu\pi^{-1} \ll \eta\pi^{-1}$ by Theorem 4.1. Thus $\mu_N = (\mu_A)_a = \mu_a = \mu$ by Lemmas 4.2 and 4.3, so that μ_D is identically zero as promised.

5. **Szegő's Theorem and examples.** Given measures α and β in G, $d\alpha/d\beta$ will denote the Radon-Nikodym derivative of the absolutely continuous part of α relative to β. μ will denote an arbitrary positive measure in G. We continue to write $\eta = |Q_0(1)|^2\mu$ as before; if μ is deterministic, define $Q_0(1) \equiv 0$.

THEOREM 5.1. *Let $\tau > 0$ with $\hat{\tau}$ supported on Γ. If $\int_G \log (d\mu/d\tau)\, d\tau > -\infty$ then μ is nondeterministic. Conversely if μ is nondeterministic then*

$$\int_G \log \frac{d\mu}{d\tau}\, d\tau \geqq \int_G \log \frac{d\eta}{d\tau}\, d\tau.$$

Proof. Suppose $\int_G \log (d\mu/d\tau)\, d\tau > -\infty$. Then $\tau\pi^{-1}(A) > 0$, where

$$A = \left\{ h \in H : \int_T \log \left(\frac{d\mu}{d\tau}(t+h) \right) d\sigma(t) > -\infty \right\},$$

and $\tau\pi^{-1}(H \sim A) = 0$. (We have used here the fact that $\tau = \sigma \times \tau\pi^{-1}$.) Now by a classical result of Szegő, $[\int_T |1-q|^2 w\, d\sigma]^{1/2} \geqq \exp \int_T \log (w)\, d\sigma$ for any polynomial q in the variable e^{-it} and any σ-summable function $w \geqq 0$. Furthermore, for each finite linear sum P of elements of $\Delta S = \{\chi \in \hat{G} : \Psi(\chi) < 0\}$, and for each $h \in H$, the function $t \to P(h+t)$ is just such a polynomial. Thus for each such P,

$$\int_G |1-P|^2\, d\mu \geqq \int_A \int_T |1-P(h+t)|^2\, d\mu/d\tau(h+t)\, d\sigma(t)\, d(\tau\pi^{-1})(h)$$

$$\geqq \int_A \left[\exp \int_T \log ((d\mu/d\tau)(h+t))\, d\sigma(t) \right]^2 d(\tau\pi^{-1})(h).$$

This integral is positive, since the bracketed integrand is positive at each $h \in A$. Taking $\text{INF}_P \int |1-P^2|\, d\mu$ over such sums P shows that the $L_2(\mu)$-distance from 1 to the span of ΔS is positive; i.e., μ is nondeterministic.

Conversely, if μ is nondeterministic and $\int_G \log (d\eta/d\tau)\, d\tau > -\infty$, then $\tau \ll \eta$. By Theorems 3.7 and 3.8 $d\mu/d\eta = d\mu_N/d\eta = |\sum_{k \geq 0} b_k \Delta^k|^2 ae - \eta$ and therefore $ae - \tau$; also $b_0 = 1$, $ae - \eta$ and $ae - \tau$. Now

$$\int_H \int_T \left| \sum_{k \geq 0} b_k(h)\Delta^k(h)e^{-ikt} \right|^2 d\sigma(t)\, d(\eta\pi^{-1})(h)$$

$$= \int_G \left| \sum_{k \geq 0} b_k \Delta^k \right|^2 d\eta = \mu_N(G) < \infty$$

implies $\int_T |\sum_{k \geq 0} b_k(h)\Delta^k(h)e^{-ikt}|^2\, d\sigma(t) < \infty$ for almost all h wrt $\eta\pi^{-1}$. By Szegő's classical inequality, for such h the relation

$$\int_T \log \left| \sum_{k \geq 0} b_k(h)\Delta^k(h)e^{-ikt} \right|^2 d\sigma(t) \geqq \log (b_0(h)) = 0$$

is valid. (To see this consider the constant $b_0(h) \equiv 1$ as the density of a measure in T, and note that the partial sums approximating $\sum_{k \geq 1} b_k(h) \Delta^k(h) e^{-ikt}$ are trigonometric polynomials in e^{-it}.) From $\tau = \sigma \times \tau \pi^{-1}$, the inequality $\int_G \log (d\mu/d\eta) \, d\tau \geq 0$ now follows immediately. Finally $\int_G \log (d\mu/d\tau) \, d\tau = \int_S \log (d\mu/d\tau) \, d\tau$, where S is chosen so that $\tau(G \sim S) = 0$ and $G \sim S$ supports the singular part of η relative to τ. Since $d\mu/d\tau$ and $d\mu/d\eta \; d\eta/d\tau$ coincide on all of S except a τ-null subset of S,

$$\int_G \log \frac{d\mu}{d\tau} \, d\tau \geqq \int_S \log \frac{d\eta}{d\tau} \, d\tau = \int_G \log \frac{d\eta}{d\tau} \, d\tau,$$

as promised.

Notice that for a nondeterministic μ, a τ of the right kind with $\int_G \log (d\eta/d\tau) \, d\tau > -\infty$ always exists; namely, $\tau = \eta$ itself. Thus Theorem 5.1 can be reworded as follows:

THEOREM 5.1'. *μ is deterministic if and only if $\int_G \log (d\mu/d\tau) \, d\tau = -\infty$ for every $\tau > 0$ with $\hat{\tau}$ supported in Γ.*

The following is a means of constructing nondeterministic measures:

THEOREM 5.2. *Let $f \geqq 0$ and $v \geqq 0$ (with $\hat{v}$ supported in Γ) be such that $\int_G f \, dv < \infty$ and*

$$v\pi^{-1}\left(\left\{h \in H : \int_T \log f(t+h) \, d\sigma(t) > -\infty\right\}\right) > 0.$$

Then $\mu = fv$ is nondeterministic.

Proof. Pick real M such that $v\pi^{-1}(A) > 0$, where

$$A = \left\{h \in H : \int_T \log f(t+h) \, d\sigma(t) > M\right\}.$$

Let τ be the truncation of v to $\pi^{-1}(A)$; thus $\tau > 0$ and $\hat{\tau}$ is supported on Γ. Furthermore $d\mu/d\tau$ and f agree on $\pi^{-1}(A)$, hence

$$\int_G \log \frac{d\mu}{d\tau} \, d\tau = \int_A \int_T \log f \, d\sigma \, d(v\pi^{-1}) \geqq Mv\pi^{-1}(A) > -\infty.$$

The proof now follows from Theorem 5.1.

Next, we have a recipe for constructing strongly nondeterministic measures:

THEOREM 5.3. *If f and v are as above and $\int \log f \, dv > -\infty$, then $\mu = fv$ is strongly nondeterministic.*

Proof. Let $g \geqq 0$ be constant on cosets $T + x$ and bounded on G, and suppose $\int g \, d\mu > 0$. Taking $\tau = gv$ yields $\tau > 0$ and

$$\int_G \log \frac{d(g\mu)}{d\tau} \, d\tau = \int_G \log f \, dv > -\infty,$$

since g is bounded. By Theorem 5.1, $g\mu$ is nondeterministic; hence μ is strongly nondeterministic.

THEOREM 5.4. *Let μ_i $(i=1,\ldots,n)$ be positive measures in G such that the $\mu_i\pi^{-1}$ have pairwise disjoint supports A_i. For each i let $P_0^i(1)$ be a measurable representative, zero outside of $\pi^{-1}(A_i)$, of the $L_2(\mu_i)$-projection of 1 onto the closed span of ΔS; let $Q_0^i(1)=1_i-P_0^i(1)$ (where 1_i is the indicator function of $\pi^{-1}(A_i)$). Let $\mu=\mu_1+\cdots+\mu_n$. Then μ is nondeterministic if and only if some μ_i is nondeterministic; in either case $Q_0^1(1)+\cdots+Q_0^n(1)=Q_0(1)ae-\mu$.*

Proof. The "if" assertion is immediate from

$$\mathop{\mathrm{INF}}_P \int |1-P|^2\, d\mu \geqq \mathop{\mathrm{INF}}_P \int |1-P|^2\, d\mu_i$$

for each i; letting P range over the finite linear sums of elements of ΔS then yields $\|Q_0(1)\|_\mu \geqq \|Q_0^i(1)\|_{\mu_i}$.

We prove the theorem by showing first, that each $Q_0^i(1)$ is orthogonal in $L_2(\mu)$ to every $\chi \in \Delta S$; and second, that each $P_0^i(1)$ is in the closed $L_2(\mu)$-span of ΔS. The first claim is obvious; we consider the second: Let s_n be a sequence of finite linear sums of elements of ΔS, such that $s_n \xrightarrow{n} P_0^i(1)$ in $L_2(\mu_i)$. Since $1_i s_n \xrightarrow{n} P_0^i(1)$ in $L_2(\mu)$, we need only show that for any $\chi \in \hat{G}$, $1_i\chi$ is in the $L_2(\mu)$-span of $\chi\Gamma$; equivalently, that 1_i is in the $L_2(\mu)$-span of Γ. To this end, consider a sequence r_n of finite linear sums of elements of Γ (i.e., of characters of H) such that r_n converges in $L_2(\mu\pi^{-1})$ to the indicator function of A_i. Then

$$\int_G |1-r_n|\, d\mu = \int_H |1_i-r_n|^2\, d(\mu\pi^{-1}),$$

since the integrand is constant on each coset $T+h$. Since the second integral vanishes as n increases, this proves $Q_0(1)=Q_0^1(1)+\cdots+Q_0^n(1)$ $ae-\mu$. In particular, if μ is nondeterministic then

$$0 < \int |Q_0(1)|^2\, d\mu = \sum_i \int |Q_0^i(1)|^2\, d\mu = \sum_i \int |Q_0^i(1)|^2\, d\mu_i$$

implies some u_i is nondeterministic.

COROLLARY 5.5. *Let μ_i and μ be as above; let $\eta_i=|Q_0^i(1)|^2\mu_i$. Then $\eta=\sum_i \eta_i$.*

COROLLARY 5.6. *Let $\mu=\mu_1+\mu_2$ as above, but with μ_1 deterministic and μ_2 strongly nondeterministic. Then $\mu_A=\mu_2$.*

Proof. By Corollary 5.5, $\eta=\eta_2$. By definition $\mu_A=\mu(\cdot \cap \pi^{-1}(A))$, where A may be taken to be any subset of H such that $\mu\pi^{-1}(\cdot \cap A)$ is the absolutely continuous part of $\mu\pi^{-1}$ relative to $\eta\pi^{-1}$. Under the present conditions A_2 serves admirably in this capacity: If $\eta\pi^{-1}(B)=0$, then $\eta_2\pi^{-1}(B)=0$ implies $0=\mu_2\pi^{-1}(B)=\mu\pi^{-1}(B \cap A_2)$ by Theorem 4.1. Thus $\mu\pi^{-1}(\cdot \cap A_2)\ll\eta\pi^{-1}$. On the other hand

$\mu\pi^{-1} - \mu\pi^{-1}(\cdot \cap A_2)$ is singular relative to η, since it is supported outside of A_2. This finishes the proof, since $\mu_2 = \mu(\cdot \cap \pi^{-1}(A_2))$.

COROLLARY 5.7. *Let $\mu = \mu_1 + \mu_2$ be as above, and with μ_2 Ψ-absolutely continuous. Then $\mu_1 = \mu_D$ and $\mu_2 = \mu_N$; i.e., $L_2(\mu_1)$ (regarded as a subspace of $L_2(\mu)$) is the remote past of μ.*

Proof. By Lemmas 4.3 and 4.2, $\mu_N = (\mu_A)_a = (\mu_2)_a = \mu_2$.

LEMMA 5.8. *Let $\mu = \mu_1 + \cdots + \mu_n$ as in Theorem 5.4, but with each μ_i strongly nondeterministic. Then μ is strongly nondeterministic.*

Proof. Let $g \geqq 0$ be bounded, constant on each coset $T + h$, with $\int g\, d\mu > 0$. For those i such that $\int g\, d\mu_i > 0$ let $\tau_i > 0$ (with $\hat{\tau}_i$ supported on Γ) be such that $\int_G \log (d(g\mu_i)/d\tau_i)\, d\tau_i > \infty$. (The existence of such a τ_i is guaranteed by Theorem 5.1′ and the fact that μ_i is strongly nondeterministic.) For the other indices i let $\tau_i \equiv 0$. Write $\tau = \tau_1 + \cdots + \tau_n$; then $\tau > 0$, $\hat{\tau}$ is supported in Γ, and

$$\int_G \log \frac{d(g\mu)}{d\tau}\, d\tau = \sum_i \int_G \log \frac{d(g\mu)}{d\tau_i}\, d\tau_i$$

$$\geqq \sum_i \int_G \log \frac{d(g\mu_i)}{d\tau_i}\, d\tau_i > -\infty.$$

This proves μ is strongly nondeterministic.

COROLLARY 5.9. *Let H be finite; let $\mu_1, \ldots, \mu_n$ be the truncation of μ to the various cosets $T + h$ ($h \in H$). Then μ is deterministic if and only if all of the μ_i are deterministic; if some μ_i is nondeterministic then it is strongly nondeterministic; in this case $\mu_A = \sum \{\mu_i : \mu_i \text{ is nondeterministic}\}$.*

Proof. The "if and only if" assertion is part of Theorem 5.4.

If some μ_i is nondeterministic then η_i and $\eta_i\pi^{-1}$ are nonnull and only the empty set can be $\eta_i\pi^{-1}$-null. Hence $\mu_i\pi^{-1} \ll \eta_i\pi^{-1}$ implies μ_i is strongly nondeterministic, by Theorem 4.1.

The last assertion is now immediate from Lemma 5.8 and Corollary 5.6, by taking $\sum \{\mu_i : \mu_i \text{ is nondeterministic}\}$ for the μ_2 of Corollary 5.6.

In conclusion, we construct a μ that is absolutely continuous relative to Haar measure $\sigma \times \lambda$ in G (write λ for Haar measure in H); whose derivative (relative to Haar measure) is positive $ae - \sigma \times \lambda$, that is nondeterministic, and whose remote past is a nonnull proper subspace of $L_2(\mu)$. To do this we assume it is possible to write $H = H_1 \cup H_2$ with the H_i disjoint and $\lambda(H_i) > 0$. Let $f_i: T \to (0, \infty)$ be σ-summable ($i = 1, 2$) with $\int_T \log (f_i)\, d\sigma > -\infty$ only for $i = 2$. Then $f_1\sigma$ is a deterministic measure in T while $f_2\sigma$ is nondeterministic. Let $\mu_i = (f_i\sigma) \times \lambda_i$, where λ_i is the truncation to H_i of λ. Then $\mu = \mu_1 + \mu_2$ is absolutely continuous relative to $\sigma \times \lambda$; its derivative at almost every $t + h$ (relative to $\sigma \times \lambda$) has the positive value $f_i(t)$ if $h \in H_i$. Taking $\sigma \times \lambda_2$ for the ν in Theorem 5.3 shows that μ_2 is strongly non-

deterministic. Furthermore μ_1 is deterministic, as we can show by considering any $\tau > 0$ with $\hat{\tau}$ supported on Γ (see Theorem 5.1'): In this case

$$\frac{d\mu_1}{d\tau}(t+h) = \frac{d((f_1\sigma) \times \lambda_1)}{d(\sigma \times \tau\pi^{-1})}(t+h) = f_1(t)\frac{d\lambda_1}{d(\tau\pi^{-1})}(h).$$

(This follows from the fact that $(f_1\sigma) \times (d\lambda_1/d(\tau\pi^{-1}))\tau\pi^{-1}$ is absolutely continuous part of $(f_1\sigma) = \lambda_1$ relative to τ.) If $\tau\pi^{-1}(H_2) > 0$ then the fact that $d\mu_1/d\tau$ is identically zero on $\pi^{-1}(H_2)$ implies $\int_{T \times H_2} \log(d\mu_1/d\tau)\, d\tau = -\infty$, hence $\int_G \log(d\mu_1/d\tau)\, d\tau = -\infty$. If $\tau\pi^{-1}(H_2) = 0$ then $\tau\pi^{-1}(H_1) > 0$; thus

$$\int_G \log\frac{d\mu_1}{d\tau}\, d\tau = \int_{T \times H_1} \log\left(f_1(t)\frac{d\lambda_1}{d(\tau\pi^{-1})}(h)\right) d\sigma(t)\, d(\tau\pi^{-1})(h)$$

$$= \int_{H_1}\int_T \log f_1(t)\, d\sigma(t)\, d(\tau\pi^{-1})(h) + \int_{H_1}\int_T \log\frac{d\lambda_1}{d(\tau\pi^{-1})}\, d\sigma(\tau\pi^{-1})$$

and the first of these integrals is $-\infty$. In other words $\int_G \log(d\mu_1/d\tau)\, d\tau = -\infty$ for any such τ. By Theorem 5.1' μ_1 is deterministic. μ and the μ_i thus satisfy the hypotheses of Corollary 5.7; in particular therefore, the remote past of μ is a non-null proper subspace of $L_2(\mu)$.

REFERENCES

1. K. de Leeuw and I. Glicksberg, *Quasi-invariance and analyticity of measures on compact groups*, Acta Math. **109** (1963), 179–205.

2. J. L. Doob, *Stochastic processes*, Wiley, New York, 1953.

3. H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables*, Acta Math. **99** (1958), 165–201.

4. E. Hewitt and K. A. Ross, *Abstract harmonic analysis*, Springer-Verlag, Berlin, 1963.

5. L. H. Loomis, *An introduction to abstract harmonic analysis*, Van Nostrand, Princeton, N. J., 1953.

SIMON FRASER UNIVERSITY,
    BURNABY, CANADA