Z. Wahrscheinlichkeitstheorie verw. Geb. 6, 279-286 (1966)

Spatially Homogeneous Markov Operators*

JAMES R. BROWN

Received July 19, 1965

Introduction

In this paper we continue the study of Markov operators begun in [2, 3]. Among all measure-preserving transformations the rotations on a compact abelian group play a special canonical role. (Cf. [7] and [6], pp. 46—50.) It turns out that among all Markov operators with a finite invariant measure those which act on functions on a compact abelian group and are spatially homogeneous, in a sense to be defined below, play essentially the same role. We show that every such operator T has an integral representation over the set of translation operators. Using this representation we then investigate the spectrum of T and prove a theorem analogous to the representation theorem of P. R. HALMOS and J. VON NEUMANN [7] for measure-preserving transformations with a discrete (pure-point) spectrum.

1. Markov Operators and the Translation Group

Let (X, \mathscr{F}, m) be a finite measure space, and let us denote by L_2 the Banach space of square-integrable, complex-valued functions on X. We shall say that an operator T on L_2 is positive if $f \ge 0 \Rightarrow Tf \ge 0$ ($f \in L_2$). A Markov operator (with invariant measure m) is a positive linear operator T on L_2 satisfying T1 = T*1 = 1. It was shown in [2] that the set M of all such operators is a convex set, which is compact in the weak operator topology. Moreover, there is a one-to-one affine correspondence between M and the set of all doubly stochastic measures on $X \times X$, i.e. positive measures λ on $(X \times X, \mathscr{F} \times \mathscr{F})$ such that $\lambda(A \times X) = \lambda(X \times A) = m(A)$ for $A \in \mathscr{F}$. This correspondence is given by

$$(f, Tg) = \int_{X \times X} f(x) \overline{g(y)} \lambda(dx, dy) \qquad (f, g \in L_2).$$
(1)

Moreover, the group Φ of all invertible measure-preserving transformations φ of (X, \mathscr{F}, m) is canonically embedded in M by setting $T_{\varphi}f(x) = f(\varphi x)$. The induced topology on Φ is the weak topology of [6].

Let X be a compact, abelian group, let \mathscr{F} be the class of Borel subsets of X, and let m be normalized Haar measure on X (m(X) = 1). The group X can be embedded as a subgroup into $\Phi \subseteq M$ by setting

$$T_y f(x) = f(x - y) (x, y \in X, f \in L_2).$$

Our first result is that this is a topological, as well as algebraic, embedding.

^{*} This paper is based in part on the author's dissertation presented for the degree of Doctor of Philosophy at Yale University. Preparation of the paper was partially supported by the Atomic Energy Commission under Contract No. AT (45-1)-1947.

J. R. BROWN:

Theorem 1. The relative topology on X induced by the weak operator topology of M coincides with the given topology of X.

Proof. Since M is compact in the weak operator topology, any Hausdorff topology on M which is weaker than the weak operator topology must coincide with the latter. Thus a subbase for the relative topology on X consists of all sets

$$\begin{split} N(y_0; f, g, \varepsilon) &= \left\{ y \in X : \left| (f, T_y g) - (f, T_{y_0} g) \right| < \varepsilon \right\} \\ &= \left\{ y \in X : \left| \int f(x) \left[\overline{g(x - y)} - \overline{g(x - y_0)} \right] m(dx) \right| < \varepsilon \right\}, \end{split}$$

where $y_0 \in X$, $\varepsilon > 0$, and f and g are *continuous* functions on X. For each such choice of f and g the function $h(x, y) = f(x) \overline{g(x - y)}$ is uniformly continuous on $X \times X$. Hence there exists a neighborhood N_{ε} of $0 \in X$ in the given topology such that

$$\begin{aligned} y - y_0 \in N_{\varepsilon} \Rightarrow \left| f(x) \overline{g(x - y)} - f(x) \overline{g(x - y_0)} \right| < \varepsilon \ (x \in X) \\ \Rightarrow \left| \int f(x) \left[\overline{g(x - y)} - \overline{g(x - y_0)} \right] m(dx) \right| < \varepsilon. \end{aligned}$$

Thus $y_0 + N_{\varepsilon} \subset N(y_0; f, g, \varepsilon)$, and the given topology on X is stronger than the weak operator topology. Since X is compact in its given topology, the two must coincide.

According to the Choquet representation theorem (see, for instance, [9]), for each $T \in M$ there exists a regular probability measure μ on the Borel subsets of M such that

$$(f, Tg) = \int_{M} (f, Sg) \mu(dS) \qquad (f, g \in L_2),$$
 (2)

and such that μ vanishes on any Baire set not containing extreme points of M. Unfortunately, no completely adequate description of the set of extreme points of M is known even when X is the unit interval (circle group), although J. LINDEN-STRAUSS [δ] has given a necessary and sufficient condition for the doubly stochastic measure λ to be extremal. It is easily seen that the set of extreme points of M contains Φ and hence X.

In the remainder of this paper we discuss operators T having a representation of the form (2) with μ concentrated on the set of translation operators. According to Theorem 1, μ can be thought of as an ordinary Borel measure on X, and equation (2) becomes

$$(f, Tg) = \int_{X} \int_{X} f(x) \overline{g(x-y)} m(dx) \mu(dy) \qquad (f, g \in L_2).$$
(3)

We shall see that the condition that T have such a representation has a natural probabilistic interpretation, when T is associated with a Markov transition function, and a natural geometric interpretation.

2. Spatially Homogeneous Operators

Definition. A doubly stochastic measure λ on $X \times X$ (or the associated Markov operator T) is said to be *spatially homogeneous* if

$$\lambda((A+x)\times(B+x)) = \lambda(A\times B) \tag{4}$$

for all $x \in X$, A, $B \in \mathscr{F}$.

We shall denote the set of all spatially homogeneous Markov operators by Σ .

In order to see the probabilistic significance of spatial homogeneity, let us suppose that T is given by a *Markov transition function*:

$$Tf(x) = \int_{\mathcal{X}} f(y) P(x, dy) \qquad (f \in L_2), \qquad (5)$$

where P(x, B) is a measurable function of $x \in X$ for each $B \in \mathscr{F}$ and a probability measure in $B \in \mathscr{F}$ for each $x \in X$. It follows from (1) and (5) that

$$\lambda(A \times B) = \int_{A} P(x, B) m(dx) \qquad (A, B \in \mathscr{F}).$$
(6)

Then λ is spatially homogeneous if and only if

$$P(x + y, B + y) = P(x, B)$$
 $m - a.e.(x)$ (7)

for each $y \in X$, $B \in \mathscr{F}$.

Note. It can be shown that T always has a representation of the form (5) if X is metrizable. In this case it can be shown rather easily that T is a (left) centralizer in the sense of J. G. WENDEL [10], and our Theorem 3 is a special case of his Theorem 1. For the details see [1].

We now turn to the geometric interpretation of spatial homogeneity.

Theorem 2. The Markov operator T is spatially homogeneous iff it commutes with all translations.

Proof. For $A \in \mathscr{F}$ let χ_A denote the characteristic function of A. According to (1), we may rewrite (4) as

$$(\chi_A, T \chi_B) = (T_x \chi_A, T T_x \chi_B) = (\chi_A, T_{-x} T T_x \chi_B)$$

for all $A, B \in \mathcal{F}, x \in X$. However, this is equivalent to

$$T = T_{-x} T T_x$$
 or $T_x T = T T_x$

for all $x \in X$.

Corollary. The convex set Σ is closed, and hence compact, in the weak operator topology.

3. Integral Representation of Spatially Homogeneous Operators

We begin by proving a lemma.

Lemma. The set of extreme points of Σ coincides with X.

Proof. For each doubly stochastic measure $\lambda \in \Sigma$ let us define $\tilde{\lambda}$ on the Borel sets of X by

$$\lambda(B) = \lambda\{(x, y) : x - y \in B\}.$$
(8)

Since the mapping $(x, y) \to x - y$ is a continuous mapping of $X \times X$ onto $X, \tilde{\lambda}$ is a probability measure on X. The correspondence $\lambda \to \tilde{\lambda}$ is thus an affine mapping of Σ into the set of all probability measures on X. Let us show that it is one-to-one.

The mapping $(x, y) \to (x, x - y)$ is a homeomorphism of $X \times X$ onto itself. For fixed $\lambda \in \Sigma$ and $B \in \mathscr{F}$ let us define

$$\nu^{B}(A) = \lambda\{(x, y) : x \in A, x - y \in B\}.$$
(9)

From the spatial homogeneity of λ it follows that

 $v^B(A+z) = \lambda\{(x+z, y+z) : x \in A, x-y \in B\} = v^B(A).$

J. R. BROWN:

By the uniqueness of Haar measure ν^B must be a multiple of *m*. Comparing (8) and (9) we see that

$$\boldsymbol{\nu}^{B}(A) = \tilde{\boldsymbol{\lambda}}(B) \, \boldsymbol{m}(A) \,. \tag{10}$$

Comparing (9) and (10) and noting that the mapping $(x, y) \rightarrow (x, x - y)$ is its own inverse, we see that

$$\lambda(A \times B) = (m \times \tilde{\lambda})\{(x, y) : x \in A, x - y \in B\}$$
⁽¹¹⁾

for all $A, B \in \mathscr{F}$. It follows from (11) that the correspondence $\lambda \to \tilde{\lambda}$ is one-to-one. Moreover, if $\tilde{\lambda}$ is any probability measure on X, then (11) determines a measure

 $\lambda \in \Sigma$ satisfying (8). That is, the correspondence $\lambda \to \tilde{\lambda}$ is onto.

It follows that λ is an extreme point of Σ iff $\tilde{\lambda}$ is an extremal probability measure on X, which is true iff $\tilde{\lambda}$ is concentrated at a point $z \in X$,

$$\lambda(B) = \chi_B(z) \, .$$

Finally, it follows from (8) and (11) that this is true iff λ is concentrated on the graph of the translation $x \to x - z$, that is

$$\lambda(C) = m\{x: (x, x-z) \in C\}.$$

This completes the proof.

Theorem 3. Suppose that $T \in \Sigma$ and let λ be the associated doubly stochastic measure on $X \times X$. Then there is a unique probability measure μ on X such that

$$\lambda(A \times B) = \int_{X} m(A \cap (B+y)) \,\mu(dy) \tag{12}$$

for all $A, B \in \mathcal{F}$. Moreover,

$$Tf(x) = \int_{X} f(x-y) \mu(dy) \qquad m-a. e.$$
 (13)

for all $f \in L_2$.

Proof. Clearly, (12) holds for all $A, B \in \mathscr{F}$ iff (3) holds for all $f, g \in L_2$. Moreover, by the Fubini theorem (3) is equivalent to (13). Thus existence follows from the Lemma and the Choquet theorem. (Since X is closed, we actually need only the Krein-Milman and Riesz representation theorems.)

Since (13) may be read $Tf = f * \mu$, uniqueness of μ follows from properties of the Fourier transform. Thus if \hat{f} and $\hat{\nu}$ denote Fourier and Fourier-Stieltjes transforms of $f \in L_2$ and the regular Borel measure ν , we have

$$f * \nu = 0 (f \in L_2) \Rightarrow \hat{f} \hat{\nu} = 0 (f \in L_2)$$
$$\Rightarrow \hat{\nu} = 0$$
$$\Rightarrow \nu = 0.$$

From the Lemma, the Krein-Milman theorem and the fact that the closed convex hull of a set of operators is the same in the weak operator topology and the strong operator topology ([5], p. 477), we obtain the following result, which is closely related to Theorem 4 of [10].

Corollary. Σ is the closed convex hull of X in the strong operator topology.

4. Spectral Properties of T and the Isomorphism Theorem

Let \hat{X} denote the dual group of X. Then \hat{X} is discrete, and by the Pontryagin duality theorem we can identify its dual with X. For $x \in X$ and $\hat{x} \in \hat{X}$ we shall

282

write $\hat{x}(x) = \langle x, \hat{x} \rangle$. We shall indicate integration with respect to Haar measure by $\int dx$ and $\int d\hat{x}$ on X and \hat{X} , respectively.

Theorem 4. Suppose that $T \in \Sigma$ has the representation (13). Then the spectrum of T is $C = \{\hat{\mu}(\hat{x}) : \hat{x} \in \hat{X}\}$, T has pure point spectrum, and the proper space corresponding to $\lambda \in C$ is the subspace spanned by $\{\hat{x} \in \hat{X} : \hat{\mu}(\hat{x}) = \lambda\}$.

Proof. If $\hat{x} \in \hat{X}$, then $\hat{x} \in L_2$, and we have by (13) that

$$T \, \hat{x}(x) = \int_{X} \langle x - y, \hat{x} \rangle \, \mu(dy)$$

= $\int_{X} \langle x, \hat{x} \rangle \langle \overline{y, \hat{x}} \rangle \, \mu(dy)$
= $\hat{\mu}(\hat{x}) \, \hat{x}(x) \, .$

Thus \hat{x} is a proper function of T corresponding to the proper value $\hat{\mu}(\hat{x})$.

On the other hand, if

$$Tf(x) = \int_X f(x-y) \mu(dy) = \lambda f(x)$$
 a.e.,

then

$$\begin{split} \lambda \widehat{f}(\widehat{x}) &= \int_{X} \overline{\langle x, \widehat{x} \rangle} \int_{X} f(x-y) \, \mu \, (dy) \, dx \\ &= \int_{X} \int_{X} \overline{\langle x, \widehat{x} \rangle} f(x-y) \, dx \, \mu \, (dy) \\ &= \int_{X} \overline{\langle y, \widehat{x} \rangle} \int_{X} \overline{\langle x-y, \widehat{x} \rangle} f(x-y) \, dx \, \mu \, (dy) \\ &= \int_{X} \overline{\langle y, \widehat{x} \rangle} \mu \, (dy) \int_{X} \overline{\langle x, \widehat{x} \rangle} f(x) \, dx \\ &= \widehat{\mu}(\widehat{x}) \, \widehat{f}(\widehat{x}) \, . \end{split}$$

If $f \neq 0$, then \hat{f} does not vanish identically. It follows that $\lambda = \hat{\mu}(\hat{x}) \in C$ for some $\hat{x} \in \hat{X}$, and that

$$f(x) = \int_{\widehat{X}} \langle x, \hat{x} \rangle \widehat{f}(\hat{x}) \, d\hat{x}$$

belongs to the subspace spanned by the set of \hat{x} such that $\hat{f}(\hat{x}) \neq 0$, i.e.

$$\{\hat{x}\in\hat{X}:\hat{\mu}(\hat{x})=\lambda\}.$$

Finally, since \hat{X} spans L_2 , it follows that T has pure point spectrum.

Note. The spectral possibilities for $T \in \Sigma$ are not as simple as in the case of a measure-preserving transformation with pure point spectrum. For instance, the proper values do not in general form a group, but only the range of a positivedefinite function on some discrete group. Nor do the proper values need to have absolute value 1, although, as for any $T \in M$, they are contained in the unit disc. Moreover, the proper values are not necessarily simple even if T is ergodic (in which case 1 is a simple proper value).

Consider, for instance, the symmetric random walk on the group of integers modulo *n*. In this case, $X = \hat{X} = \{0, 1, ..., n-1\}$ with

$$\langle k,j\rangle = e^{2\pi i j k/n},$$

and

$$Tf(k) = \frac{1}{2}f(k-1) + \frac{1}{2}f(k+1) = \sum_{l} f(k-l)\mu(l).$$

Thus

$$\hat{\mu}(j) = \sum_{k} \langle \overline{k, j} \rangle \mu(k)$$

= $\frac{1}{2} \langle \overline{1, j} \rangle + \frac{1}{2} \langle \overline{-1, j} \rangle$
= $\cos(2\pi j/n)$.

If follows that 1 and possibly -1 are simple proper values, while others are double. A similar situation holds for the Brownian motion on the circle group.

Finally, if $\mu = m$ so that

$$Tf(x) = \int_X f(y) dy = \text{const.},$$

then 1 is a simple proper value, while 0 has infinite multiplicity.

It is true, however, that the linearly independent proper functions of $T \in \Sigma$ may be chosen to have constant absolute value 1 and to form a group under pointwise multiplication. We shall show that these properties characterize spatially homogeneous Markov operators up to spatial homomorphism.

Definition. Let T_1 and T_2 be bounded linear operators on $L_2(X_1, \mathscr{F}_1, m_1)$ and $L_2(X_2, \mathscr{F}_2, m_2)$, respectively. We say that T_1 and T_2 are spectrally isomorphic if there exists an invertible isometry U of $L_2(X_1, \mathscr{F}_1, m_1)$ onto $L_2(X_2, \mathscr{F}_2, m_2)$ such that $T_1 = U^{-1}TU$. They are spatially isomorphic if $Uf(x) = f(\psi x)$ $(f \in L_2(X_1, \mathscr{F}_1, m_1))$, where $\psi: X_2 \to X_1$ is an invertible measure-preserving transformation (modulo sets of measure 0).

P. R. HALMOS and J. VON NEUMANN have shown [7] that $T_{\varphi} \in \Phi$ is spatially isomorphic to a translation in a compact group iff it has pure point spectrum. The next theorem extends this result. The proof is a modification of that of HALMOS and VON NEUMANN.

Theorem 5. A Markov operator T on $L_2(X)$ of the finite measure space (X, \mathcal{F}, m) is spatially isomorphic to a spatially homogeneous operator \tilde{T} on a compact abelian group G iff there exists a complete orthonormal system C for $L_2(X)$ consisting of bounded proper functions of T and forming a group under pointwise multiplication.

Note. If T is spatially isomorphic to an ergodic measure-preserving transformation, then it can be shown (see, for instance, [1]) that T is also given by a measure-preserving transformation. For such a transformation our condition is equivalent to requiring that T have pure point spectrum (see [6], p. 34). Thus Theorem 5 is indeed an extension of the above-mentioned theorem.

Proof. For X = G the necessity follows from Theorem 4. More generally, the isometry U_{ψ} which implements the isomorphism carries \tilde{G} onto an orthonormal basis C for $L_2(X)$ and preserves the boundedness and multiplicative properties of that set.

To prove the sufficiency we let the group C have the discrete topology and denote its dual group by G. Then G is compact. Under the identification of Cwith \hat{G} the set C becomes a complete orthonormal system in $L_2(G)$. We shall

 $\mathbf{284}$

indicate the corresponding isometric isomorphism of $L_2(X)$ and $L_2(G)$ by $f \rightarrow \tilde{f} = Wf$. Thus, in particular W maps C onto \hat{G} . Now let $\tilde{T} = WTW^{-1}$. We shall show that \tilde{T} is spatially homogeneous, and that $W = U_{\psi}$ is induced by a measure-preserving transformation ψ .

For each $\tau \in C$ we have

$$\widetilde{T}(W\tau) = W(T\tau) = \lambda_{\tau}(W\tau),$$

where λ_{τ} is the proper value of T corresponding to the proper function τ . Thus $W\tau$ is a proper function of \hat{T} corresponding to the same λ_{τ} .

Let $g \in L_2(X)$. Then $\tilde{g} = Wg$ has the Fourier expansion

$$\tilde{g}(x) = \int_{C} \hat{\tilde{g}}(\tau) \langle \tau, x \rangle d\tau \quad (x \in G)$$

and so

$$T\tilde{g}(x) = \int_{C} \hat{\tilde{g}}(\tau) \lambda_{\tau} \langle \tau, x \rangle d\tau \qquad (x \in G).$$
(14)

If $T_y (y \in G)$ is defined as in §1, and if $h = T_y \tilde{g}$ for fixed $y \in G$, then

$$\begin{split} \hat{h}(\tau) &= \int_{G} h(x) \langle \overline{\tau, x} \rangle \, dx \\ &= \int_{G} \tilde{g}(x - y) \langle \overline{\tau, x} \rangle \, dx \\ &= \int_{G} \tilde{g}(z) \langle \overline{\tau, x + y} \rangle \, dx \\ &= \langle \overline{\tau, y} \rangle \, \hat{\bar{g}}(\tau) \, (\tau \in C) \, . \end{split}$$

$$\end{split}$$
(15)

Combining (14) and (15) gives

$$\begin{split} \widetilde{T}\,T_y g(x) &= \int\limits_C \langle \overline{ au,y}
angle \hat{ar{g}}(au) \, \lambda_ au \langle au,x
angle \, dx \ &= \int\limits_C \hat{ar{g}}(au) \, \lambda_ au \langle au,x-y
angle \, dx \ &= T_y \, \widetilde{T} \, \widetilde{g}(x) \qquad (x \in G) \, . \end{split}$$

Thus \tilde{T} commutes with each $T_y (y \in G)$, and according to Theorem 2 \tilde{T} is spatially homogeneous.

It remains to show that the isometry W of $L_2(X)$ onto $L_2(G)$ is a spatial isomorphism. Since the restriction of W to C is a group isomorphism of C onto \hat{G} , the equality

$$W(fg) = (Wf)(Wg) \tag{16}$$

holds for all $f, g \in C$. By linearity (16) holds for all f, g in the linear space spanned by C. If $B \in \mathscr{F}$, let f be a fixed linear combination of elements of C, and let $g_n \to \chi_B$ in $L_2(X)$. Then f and Wf are bounded functions so that $fg_n \to f\chi_B$ and $(Wf) (Wg_n) \to (Wf) (W\chi_B)$. Thus $W(f\chi_B) = (Wf) (W\chi_B)$. Similarly, letting $f_n \to \chi_B$ gives $W(\chi_B^2) = W(\chi_B) = (W\chi_B)^2$. Thus $W\chi_B \ge 0$. It follows that $Wf \ge 0$ whenever $f \ge 0$, i.e. W is a positive operator. Clearly, W1 = 1 (= W^*1). Since W is an invertible isometry, we have by a trivial modification of Theorem 5 of [2] that $W = U_{\psi}$ for some invertible measure-preserving transformation ψ . **Corollary 5.1.** If T satisfies the hypotheses of Theorem 5, then there exists a probability measure μ on X and a measurable family $\{\varphi_x : x \in X\}$ of measure-preserving transformations of X with pure point spectrum such that

$$Tf(x) = \int_{\mathcal{X}} f(\varphi_y x) \,\mu(dy) \qquad a.e. \tag{17}$$

for each $f \in L_2(X)$.

Proof. This is simply a change of variables in (13) with $T\varphi_y = U_{\psi}T_{\psi y}U_{\psi}^{-1}$ so that φ_y has pure point spectrum.

Corollary 5.2. If T satisfies the hypotheses of Theorem 5, then T is spatially isomorphic to its adjoint T^* .

Proof. If T and \tilde{T} are isomorphic, then so are T^* and $\tilde{T^*}$. Thus we may assume that T is spatially homogeneous. Then

$$Tf(x) = \int_{X} f(x-y) \,\mu(dy)$$

and

$$T^*f(x) = \int_X f(x+y) \,\mu(dy) \,.$$

Let $\psi(x) = -x$. Then

$$U_{\psi}(T^*f)(x) = T^*f(-x) = \int_{X} f(-x+y) \mu(dy) = T(U_w f)(x).$$

Thus $T^* = U_{\psi}^{-1} T U_{\psi}$ as asserted.

Remark. J. R. CHOKSI has recently shown [4] that for non-ergodic transformations with pure point spectrum spectral isomorphism need not imply spatial isomorphism. However, the above proof of Theorem 5 does not depend on ergodicity as in the case of [7]. Thus our theorem applies to non-ergodic transformations.

References

- BROWN, J. R.: Approximation theorems and integral representations for Markov operators. Yale University, Dissertation, 1964.
- [2] Approximation theorems for Markov operators. Pacific J. Math. 16, 13-23 (1966).
- [3] Doubly stochastic measures and Markov operators. Michigan math. J. 12, 367-375 (1965).
- [4] CHOKSI, J. R.: Non-ergodic transformations with discrete spectrum. Ill. J. Math. 9, 307-320 (1965)
- [5] DUNFORD, N., and J. T. SCHWARTZ: Linear operators, Part I. New York: Interscience 1958.
- [6] HALMOS, P. R.: Ergodic theory. New York: Chelsea 1956.
- [7] -, and J. VON NEUMANN: Operator methods in classical mechanics., II. Ann. of Math., II. Ser. 43, 332-350 (1942).
- [8] LINDENSTRAUSS, J.: A remark on extreme doubly stochastic measures. Amer. math. Monthly 72, 379-382 (1965).
- [9] PHELPS, R. R.: Lectures on Choquet's theorem. Princeton: van Nostrand 1966.
- [10] WENDEL, J. G.: Left centralizers and isomorphisms of group algebras. Pacific J. Math.
 2, 251-261 (1952).

Department of Mathematics Oregon State University Corvallis, Oregon 97331 (U.S.A.)