TRANSITION PROBABILITY OPERATORS

M. ROSENBLATT
UNIVERSITY OF CALIFORNIA, SAN DIEGO

1. Summary

First, a few remarks are made on invariant probability measures of a transition probability operator. The notion of irreducibility is introduced for a transition operator acting on the continuous functions on a compact space, and implications of this assumption are examined. Finally, the unitary and isometric parts of the operator are isolated and interpreted in terms of the behavior of iterates of the operator.

2. Preliminary remarks

Let us consider a space Ω with a Borel field of subsets $\mathcal{B}$. We shall call $P(x, B)$, $x \in \Omega$, $B \in \mathcal{B}$, a transition probability function if $P(x, \cdot)$ is a probability measure on $\mathcal{B}$ for each $x \in \Omega$ and $P(\cdot, B)$ is a $\mathcal{B}$-measurable function for each $B \in \mathcal{B}$. Higher order transition probability functions can then be introduced recursively starting with $P(\cdot, \cdot)$:

$$(1)\qquad \begin{aligned} P_1(x, B) &= P(x, B), \\ P_{n+1}(x, B) &= \int P_1(x, dy)P_n(y, B), \qquad n = 1, 2, \cdots, \quad x \in \Omega, \quad B \in \mathcal{B}. \end{aligned}$$

It can easily be seen that the functions $P_n(\cdot, \cdot)$ probability measures on $\mathcal{B}$ for each $x \in \Omega$ and measurable in x for each $B \in \mathcal{B}$. Further, a small argument shows that

$$(2)\qquad P_{n+m}(x, B) = \int P_n(x, dy)P_m(y, B), \qquad n, m = 1, 2, \cdots$$

There is an operator T taking bounded measurable functions into bounded measurable functions induced by $P(\cdot, \cdot)$:

$$(3)\qquad (Tf)(x) = \int P(x, dy)f(y).$$

The operator T is positive in that $Tf \geq 0$ if $f \geq 0$; by $f \geq 0$ we mean that $f(x) \geq 0$ for all $x \in \Omega$. Also, T maps the function one onto itself. A convenient

This research was supported in part by the Office of Naval Research Contract Nonr 2216(26) at the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

norm can be introduced for the bounded functions f in terms of the transition function $P(\cdot, \cdot)$. Let

$$(4) \qquad \|f\|_s = \{\inf_c |\sup_x P(x, \{y| |f(y)| \geq c\}) = 0\}.$$

Then $\|Tf\|_s \leq \|f\|_s$, and hence the corresponding norm of the operator T, $\|T\|_s$, is less than or equal to one.

The transition function $P(\cdot, \cdot)$ also induces an operator acting on finite measures and taking them into finite measures. At the risk of a lack in clarity, we shall also denote this operator by T and let

$$(5) \qquad (\mu T)(B) = \int \mu(dx)\, P(x, B) = \nu(B)$$

for any finite measure μ on $\mathcal{B}$. The operator T takes probability measures Q into probability measures since

$$(6) \qquad (QT)(\Omega) = \int Q(dx)\, P(x, \Omega) = \int Q(dx) = 1.$$

We shall call Q an *invariant* probability measure with respect to $P(\cdot, \cdot)$ if

$$(7) \qquad (QT)(B) = \int Q(dx)\, P(x, B) = Q(B)$$

for all $B \in \mathcal{B}$. The set of probability measures is obviously a convex set, and the set of invariant probability measures is a convex subset of the set of probability measures.

Given any specific invariant probability measure Q, we can consider T as an operator on the space $L_2(dQ)$ of square integrable functions with respect to Q. For if $f \in L_2(dQ)$,

$$(8) \qquad \int Q(dx) \left|\int T(f)(x)\right|^2 = \int Q(dx) \left|\int P(x, dy)\, f(y)\right|^2$$
$$\leq \int Q(dx) \int P(x, dy) |f(y)|^2 = \int Q(dx) |f(x)|^2,$$

so that $(Tf)(x)$ is well-defined for almost all x with respect to Q and $Tf \in L_2(dQ)$. Further, $\|Tf\| \leq \|f\|$ in the norm of $L_2(dQ)$, so that T is a contraction operator on $L_2(dQ)$. Since T is a contraction on $L_2(dQ)$, it follows that given any $f \in L_2(dQ)$,

$$(9) \qquad \frac{1}{n} \sum_{k=1}^{n} T^k f \rightarrow \hat{f}$$

converges in $L_2(dQ)$ to $\hat{f} \in L_2(dQ)$ as $n \rightarrow \infty$. We shall say that the invariant measure Q *is ergodic with respect to* T if

$$(10) \qquad \hat{f} = E_Q f = \int f(x) Q(dx)$$

for all $f \in L_2(dQ)$.

THEOREM 1. *Given any two invariant probability measures Q_1, Q_2 ergodic with respect to T, it follows that they are either singular with respect to each other, or else identical.*

If the measures Q_i, $i = 1, 2$, are not singular with respect to each other, there is a set M on which they are absolutely continuous with respect to each other having positive Q_1 and Q_2 measure.

Given any bounded function f,

$$\frac{1}{n} \sum_{k=1}^{n} T^k f \to E_{Q_1} f, \qquad E_{Q_2} f \tag{11}$$

in measure with respect to both Q_1, Q_2 as $n \to \infty$ on M. Then $E_{Q_1} f = E_{Q_2}$ for all bounded f. Let f be the characteristic function of a set B, c_B. It follows that

$$\int Q_1(dx) \frac{1}{n} \sum_{k=1}^{n} (T^k c_B)(x) = Q_1(B) \tag{12}$$

$$= \int Q_1(dx)\, E_{Q_1} c_B = \int Q_2(dx)\, E_{Q_2} c_B = Q_2(B).$$

Thus the two measures Q_1 and Q_2 are identical.

We shall call an invariant probability measure Q *an extreme point of the set of invariant probability measures* if whenever Q is expressible as a convex combination of two invariant probability measures Q_1, Q_2,

$$Q = \alpha Q_1 + (1 - \alpha) Q_2, \qquad 0 < \alpha < 1, \tag{13}$$

it automatically follows that $Q = Q_1 = Q_2$. First, it is easy to see that every ergodic measure Q is an extreme point of the set of invariant probability measures. Since $Q = \alpha Q_1 + (1 - \alpha)Q_2$, $0 < \alpha < 1$, it follows that Q_1 and Q_2 are absolutely continuous with respect to Q. However, theorem 1 then implies that $Q_1 = Q_2 = Q$.

On the other hand, if Q is not ergodic, it cannot be an extreme point of the set of invariant probability measures. Since Q is not ergodic, there is a function $f \in L_2(dQ)$ such that $\hat{f}$ is not constant almost everywhere. Let c be such that $Q(A) = \alpha$, $0 < \alpha < 1$ where $A = \{x | \hat{f}(x) \geq c\}$. The fact that T is a positive contraction on $L_2(dQ)$, taking one into one, implies that $P(x, A) = 1$ for almost all $x(dQ)$ in A. Thus Q_1 and Q_2 are invariant probability measures where

$$\begin{aligned} Q_1(B) &= \frac{1}{\alpha} Q(B \cap A), \\ Q_2(B) &= \frac{1}{1 - \alpha} Q(B \cap \bar{A}), \end{aligned} \tag{14}$$

and

$$Q = \alpha Q_1 + (1 - \alpha) Q_2. \tag{15}$$

Here $\bar{A}$ denotes the complement of A. We therefore have the following theorem.

THEOREM 2. *The extreme points of the set of invariant probability measures with respect to $P(\cdot, \cdot)$ are precisely the ergodic probability measures with respect to $P(\cdot, \cdot)$.*

Blum and Hanson [1] have discussed related questions for point transformations.

3. Transition operators on a compact space

In this section we shall assume that Ω *is a compact Hausdorff space* and $\mathcal{B}$ the Borel field generated by the topology. Further, T *is to take continuous functions into continuous functions*. Therefore, we may take it for granted that $P(x, \cdot)$ is a regular measure on $\mathcal{B}$ for each $x \in \Omega$. Given any fixed point x, let $\sigma_m(x)$ be the spectrum of $P_m(x, \cdot)$, that is, the set of points y such that for every open set O containing y, $P_m(x, O) > 0$. This set is a closed set for every x and $m = 0, 1, \cdots$. By P_0 we shall understand the identity point mapping. The following lemma was obtained in [4].

LEMMA 1. *The spectral set*

$$\sigma_{m+n}(x) = \overline{\bigcup_{y \in \sigma_m(x)} \sigma_n(y)}. \tag{16}$$

The sets $\sigma_n(x)$ correspond to exit from x to points n steps later on. One can also introduce sets $\tau_n(x)$ corresponding to entrance into the immediate neighborhood of x from points n steps before. Let f be any nonnegative continuous function with $f(x) > 0$. Consider

$$S_n(f) = \left\{z \Big| \int P_n(z, dy) f(y) > 0\right\} \tag{17}$$

and set

$$\tau_n(x) = \prod_{f(x) > 0, f \geq 0} S_n(f). \tag{18}$$

Notice that $x \in \tau_n(z)$ if and only if $z \in \sigma_n(x)$. For if $x \in \tau_n(z)$, given any continuous $f \geq 0$ with $f(z) > 0$, it follows that $(T^n f)(x) > 0$, and hence $z \in \sigma_n(x)$. The converse statement follows by going in the reverse direction.

We say that T is *irreducible* acting on the continuous functions if for any continuous $f \geq 0$, $f \not\equiv 0$, and any given point x there is a positive integer $n(f, x)$ such that $\sum_1^n (T^k f)(x) > 0$. The compactness of Ω implies that we can find an $n(f)$ such that $\sum_1^n (T^k f) > 0$ for all x.

LEMMA 2. *The probability $P_n(x, \sigma_n(x))$ is equal to one.*

Consider any open set O containing $\sigma_n(x)$. Let C be the complement of O. For each $y \in C$ there is an open set O_y with $y \in O_y$ such that $P_n(x, O_y) = 0$. C is compact, and hence there is a finite subcovering O_{y_j}, $j = 1, \cdots, n$, of the covering $\{O_y, y \in C\}$ of C. Now $P_n(x, C) \leq \sum P_n(x, O_{y_j}) = 0$ so that $P_n(x, O) = 1$ for every open O with $\sigma_n(x) \subset O$. Thus $P_n(x, \sigma_n(x)) = 1$ by the regularity of $P_n(x, \cdot)$.

THEOREM 3. *The operator T is irreducible acting on continuous functions if and only if*

$$\overline{\bigcup_{n=1}^{\infty} \sigma_n(x)} = \Omega \tag{19}$$

for every x.

For suppose $\overline{\bigcup \sigma_n(x)} = \Omega$ for every x. Consider any $f \geq 0$, $f \not\equiv 0$. Then there is a point z such that $f(z) > 0$. Since $z \in \overline{\bigcup \sigma_n(x)}$ $(= \Omega)$, there is a $z' \in \bigcup \sigma_n(x)$

such that $f(z') > 0$. Then for some n, $(T^n f)(x) > 0$. Thus T is irreducible. Now assume that $\overline{\bigcup \sigma_n(x)} \neq \Omega$. Then there is a $z \notin \overline{\bigcup \sigma_n(x)}$. This implies that there is a continuous f with $f(z) = 1$ and $f = 0$ on $\overline{\bigcup \sigma_n(x)}$. By lemma 2, $(T^n f)(x) = 0$ for each $n > 0$, and hence T is not irreducible.

Given the transition function $P(\cdot, \cdot)$ we can construct the corresponding Markov process starting at x at time zero. Consider the space of sequences $(x_0, x_1, x_2, \cdots)$ with $x_0 = x \in \Omega$, $x_1, x_2, \cdots \in \Omega$. Define the measure $P(\cdots | x_0 = x)$ on product sets of the form $B_1 \times B_2 \times \cdots \times B_k$, $B_i \in \mathcal{B}$, $i = 1, \cdots, k$ by

$$(20) \qquad P(B_1 \times B_2 \times \cdots \times B_k | x_0 = x) = P(x_j \in B_j, j = 1, \cdots, k | x_0 = x)$$
$$= \int_{B_1} P(x, dx_1) \int_{B_2} P(x_1, dx_2) \cdots \int_{B_{k-1}} P(x_{k-1}, B_k),$$

and extend it in the natural way to the Borel field generated by these product sets.

LEMMA 3. *If T is irreducible on continuous functions, then for each point x and every nonvacuous open set O,*

$$(21) \qquad P(x_n \in O \text{ infinitely often}, n = 1, 2, \cdots | x_0 = x) > c(O) > 0.$$

Irreducibility implies that for each continuous $f \geq 0$, $f \not\equiv 0$, there is an n such that $\sum_1^n T^k f > \delta > 0$ for all x. Consider any fixed x and any given nonvacuous open set O. Let z be a point in O. Take f, a continuous function, $0 \leq f \leq 1$, equal to one at z and zero outside of O. Then for some m, $\sum_1^m T^k f > \delta > 0$. But this implies that $\lim_{n \to \infty} (1/n) \sum_1^n T^k f > (\delta/m) > 0$, and this cannot be the case unless, starting with any specific x, there is a positive probability of entering O infinitely often. Further, the probability of entering O infinitely often starting with x is bounded away from zero by some positive number c independent of x but generally depending on O.

THEOREM 4. *If T is irreducible on the continuous functions, for each nonvacuous open set O and each point $x \in \Omega$,*

$$(22) \qquad P(x_n \in O \text{ infinitely often}, n = 1, 2, \cdots | x_0 = x) = 1.$$

Suppose there is an x and a nonvacuous open set O such that

$$(23) \qquad P(x_n \in O \text{ finitely often}, n = 1, 2, \cdots | x_0 = x) > 0.$$

But then there is an integer k such that

$$(24) \qquad P(x_n \in O \ k \text{ times}, n = 1, 2, \cdots | x_0 = x) > 0.$$

For some $n_1, \cdots, n_k$,

$$(25) \qquad P(x_{n_i} \in O, i = 1, \cdots, k; x_n \notin O \text{ otherwise} | x_0 = x) = q > 0.$$

Let

$$(26) \qquad A = \{x_{n_i} \in O, i = 1, \cdots, k; \ x_n \notin O \text{ otherwise}\},$$

$$(27) \qquad A_N = \{x_{n_i} \in O, i = 1, \cdots, k; \ x_n \notin O \text{ otherwise}, n \leq N\},$$

where $n_1 < n_2 < \cdots < n_k < N$. Then, given any $\epsilon > 0$ for some $N(\epsilon)$ sufficiently large,

$$P(A_N|x_0 = x) < P(A|x_0 = x)(1 + \epsilon). \tag{28}$$

However,

$$\begin{aligned} P(A_N \cap \{x_n \in O \text{ infinitely often } n > N\}|x_0 = x) &= \int_{\bar{O}} P(A_N;\ dx_N|x_0 = x)P(x_n \in O \text{ infinitely often } n > N|x_N = x) \\ &\geq cP(A_N|x_0 = x) \end{aligned} \tag{29}$$

where $P(A_N;\ B|x_0 = x) = P(A_N \cap \{x_N \in B\}|x_0 = x)$. Further,

$$P(A_N \cap \{x_n \in O \text{ infinitely often } n > N\}|x_0 = x) > cP(A|x_0 = x), \tag{30}$$

and this is a contradiction if $\epsilon < c$.

4. Iterates of a transition function with respect to an invariant measure

Let Q be an invariant probability measure for the transition function $P(\cdot,\ \cdot)$. We have already remarked that the induced operator T,

$$(Tf)(x) = \int P(x, dy)f(y), \tag{31}$$

acting on $L_2(dQ)$, is a positive contraction taking the function one into itself. Consider the family of functions $\mathfrak{U}_j$ of $L_2(dQ)$ for which T^j is norm-preserving, that is

$$\int |g(x)|^2\, Q(dx) = \int |(T^jg)(x)|^2\, Q(dx) \tag{32}$$

for $g \in \mathfrak{U}_j$. For $f \in L_2(dQ)$,

$$\begin{aligned} |(T^jf)(x)|^2 &= \left|\int P_j(x, dy)f(y)\right|^2 \\ &\leq \int P_j(x, dy)|f(y)|^2 = (T^j|f|^2)(x). \end{aligned} \tag{33}$$

Thus for $g \in \mathfrak{U}_j$,

$$\begin{aligned} \int |g(x)|^2\, Q(dx) &= \int (T^j|g|^2)(x)\, Q(dx) \\ &= \int |(T^jg)(x)|^2\, Q(dx), \end{aligned} \tag{34}$$

and hence,

$$(T^j|g|^2)(x) = |(T^jg)(x)|^2 \tag{35}$$

for almost all x with respect to Q. Given any x for which (35) holds, it follows that $g(y)$ is constant for almost all y with respect to $P_j(x,\ \cdot)$. Thus it is clear that $\mathfrak{U}_j$ precisely consists of those functions with the following property: for almost all x (dQ), $g(\cdot)$ is constant for almost all y with respect to $P_j(x,\ \cdot)$. Obviously, $\mathfrak{U}_j$ is a closed linear space of functions. Further, the essentially bounded functions in $\mathfrak{U}_j$ are an algebra since the product of any two essentially bounded functions g_1, g_2 in $\mathfrak{U}_j$ is an essentially bounded function in $\mathfrak{U}_j$. We thus have the following lemma.

LEMMA 4. *The set $\mathfrak{U}_j$ of functions in $L_2(dQ)$ whose norm is preserved by T^j is a closed linear space. The essentially bounded functions of $\mathfrak{U}_j$ are an algebra containing the function one.*

It immediately follows that there is a Borel field of sets $\mathfrak{B}_j$ induced by $\mathfrak{U}_j$. In fact, $\mathfrak{U}_j$ is precisely the set of square integrable functions (with respect to Q) that are measurable (modulo a set of Q measure zero) with respect to $\mathfrak{B}_j$. Clearly, $\mathfrak{U}_1 \supset \mathfrak{U}_2 \supset \cdots$ and $\mathfrak{B} = \mathfrak{B}_0 \supset \mathfrak{B}_1 \supset \cdots$.

Let

$$\mathfrak{U}_\infty = \prod_{j=1}^{\infty} \mathfrak{U}_j, \qquad \mathfrak{B}_\infty = \prod_{j=1}^{\infty} \mathfrak{B}_j. \tag{36}$$

The set of functions $\mathfrak{U}_\infty$ is the closed linear space of functions whose norm is preserved by all powers T^j, $j = 1, 2, \cdots$, of T and is the set of square integrable functions measurable with respect to $\mathfrak{B}_\infty$. In fact, T induces a measure-preserving set transformation on the sets of $\mathfrak{B}_{j+1}$ taking the sets of $\mathfrak{B}_{j+1}$ into $\mathfrak{B}_j$, $j = 0, 1, \cdots$. For if $B \in \mathfrak{B}_{j+1}$,

$$(T|c_B|^2)(x) = (Tc_B)(x) = |(Tc_B)(x)|^2, \tag{37}$$

so that $(Tc_B)(x)$ is the characteristic function of a set that we shall call τB. Further,

$$\begin{aligned} Q(\tau B) &= \int (Tc_B)(x)\, Q(dx) = \int c_B(x)\, Q(dx) \\ &= Q(B). \end{aligned} \tag{38}$$

Clearly, $\tau B \in \mathfrak{B}_j$ since T takes $\mathfrak{U}_{j+1}$ in $\mathfrak{U}_j$. If $B \in \mathfrak{B}_j$ but not in $\mathfrak{B}_{j+1}$, we have

$$T^k c_B = c_{\tau^k B}, \qquad k = 0, 1, \cdots, j, \tag{39}$$

but on the $(j + 1)$-st application of T, it no longer acts as a set transformation.

Consider now the action of T on the characteristic functions of sets in $\mathfrak{B}_\infty$. Then

$$T^k c_B = c_{\tau^k B} \qquad k = 0, 1, 2, \cdots \tag{40}$$

if $B \in \mathfrak{B}_\infty$ with $\tau^k B \in \mathfrak{B}_\infty$ for $k = 0, 1, 2, \cdots$. Notice that the sets $B \in \mathfrak{B}_\infty$ correspond to a class of events in the backward tail field of the stationary Markov process determined by the transition function $P(\cdot, \cdot)$ and the invariant measure $Q(\cdot)$.

Since $Q(\cdot)$ is an invariant probability measure, we can consider the operator T^* adjoint to T

$$\int (Tf)(x) g(x) Q(dx) = \int f(x) (T^* g)(x) Q(dx) \tag{41}$$

for f, $g \in L_2(dQ)$. The adjoint operator T^* corresponds to the backward transition function $P^*(\cdot, \cdot)$ of the stationary Markov process determined by T and Q. As before, the set $\mathfrak{U}_j^*$ of functions in $L_2(dQ)$ whose norm is preserved by T^{*j} is a closed linear space consisting of the square integrable functions measurable with respect to some Borel field $\mathfrak{B}_j^*$:

$$\mathfrak{U}_1^* \supset \mathfrak{U}_2^* \supset \cdots, \qquad \mathfrak{B} = \mathfrak{B}_0^* \supset \mathfrak{B}_1^* \supset \cdots. \tag{42}$$

We set

$$\mathfrak{U}^*_\infty = \prod_{j=1}^{\infty} \mathfrak{U}^*_j, \qquad \mathfrak{B}^*_\infty = \prod_{j=1}^{\infty} \mathfrak{B}^*_j. \tag{43}$$

The operator T^* induces a measure-preserving set transformation τ^* taking the sets of $\mathfrak{B}^*_{j+1}$ into $\mathfrak{B}^*_j$, $j = 0, 1, \cdots$, and the sets of $\mathfrak{B}^*_\infty$ into $\mathfrak{B}^*_\infty$. The sets $B \in \mathfrak{B}^*_\infty$ correspond to a class of events in the forward tail field of the stationary Markov process determined by T (or T^*) and $Q(\cdot)$. The Borel field $\bar{\mathfrak{B}} = \mathfrak{B}_\infty \cap \mathfrak{B}^*_\infty$ is of special interest because the set of functions $\bar{\mathfrak{U}}$ square integrable and measurable with respect to $\bar{\mathfrak{B}}$ consists precisely of those functions whose norm is preserved by both T^k, $k = 1, 2, \cdots$ and T^{*k}, $k = 1, 2, \cdots$. The map τ acts as a measure-preserving invertible transformation on the sets of $\bar{\mathfrak{B}}$, that is $\tau\tau^*B = \tau^*\tau B = B$ for $B \in \bar{\mathfrak{B}}$ with $Q(\tau B) = Q(\tau^* B) = Q(B)$.

The set of functions $f \in L_2(dQ)$ such that $\int |(T^j f)(x)|^2 \, Q(dx) \to 0$ as $j \to \infty$ are of some interest. Call this set of functions $\mathfrak{D}$. It is a closed linear space by the Minkowski inequality. Let $g \in \mathfrak{D}$ and $h \in \mathfrak{U}_\infty$. By the characterization of functions in $\mathfrak{U}_\infty$, it follows that

$$\begin{aligned} T^j(g\bar{h})(x) &= \int P_j(x, dy) g(y) \bar{h}(y) \\ &= \int P_j(x, dy) g(y) \int P_j(x, dy) \bar{h}(y) \\ &= (T^j g)(x)(T^j \bar{h})(x) \end{aligned} \tag{44}$$

for almost all x with respect to Q. However,

$$\begin{aligned} \left| \int g(x)\bar{h}(x) Q(dx) \right| &= \left| \int (T^j g\bar{h})(x) Q(dx) \right| \\ &= \left| \int (T^j g)(x)(T^j \bar{h})(x) Q(dx) \right| \\ &\leq \left\{ \int |(T^j g)(x)|^2 Q(dx) \int |(T^j \bar{h})(x)|^2 Q(dx) \right\}^{1/2} \end{aligned} \tag{45}$$

which tends to zero as j tends to infinity, so that the $\mathfrak{U}_\infty$ is orthogonal to $\mathfrak{D}$. The space $\mathfrak{D}^*$ is the set of functions $f \in L_2(dQ)$ such that $\int |(T^{*j} f)(x)|^2 Q(dx) \to 0$ as $j \to \infty$, and in the same manner, one can show that $\mathfrak{D}^*$ is orthogonal to $\mathfrak{U}^*_\infty$.

In the case of a countable state space or an almost periodic transition operator T, $\mathfrak{U}_\infty = \mathfrak{U}^*_\infty$ and $\mathfrak{D} = \mathfrak{D}^*$. It is easy to consider a class of transition operators for which this need not be the case. Let Ω be the unit interval $[0, 1]$ and T the transition operator

$$(Tf)(x) = p(x) f(\tfrac{1}{2}x) + (1 - p(x)) f(\tfrac{1}{2} + \tfrac{1}{2}x) \tag{46}$$

acting on the continuous functions on $[0, 1]$. It is assumed that $p(x)$ is a continuous function with $0 \leq p(x) \leq 1$ so that T takes continuous functions into continuous functions (see [2]). There will then be at least one invariant measure $Q(\cdot)$. In particular, consider the simple case in which $p(x) \equiv \frac{1}{2}$ and $Q(\cdot)$ is

Lebesgue measure on [0, 1]. The set $\mathfrak{U}_\infty$ contains precisely the constant functions and $\mathfrak{D}$ all functions orthogonal to them. On the other hand, T^* is given by

$$(T^*f)(x) = f([2x]) \tag{47}$$

where $[y] = y \bmod (0, 1)$. In fact, T^* is given by (47) in the case of any continuous $p(x)$. Thus, $\mathfrak{U}_\infty^* = L_2(dQ)$, and $\mathfrak{D}^*$ is trivial since it contains only the function zero. This is fairly obvious since with $p(x) \equiv \frac{1}{2}$ the stationary Markov process generated by T is purely nondeterministic going forward in time and purely deterministic going backwards in time.

In the case of an almost periodic transition operator, $L_2(dQ)$ is precisely the Hilbert space generated by the two orthogonal spaces $\mathfrak{U}_\infty$ and $\mathfrak{D}$ (see [5]). It is an interesting question as to whether this is still true in general. In particular, if $\mathfrak{U}_\infty$ is trivial, in that it consists only of the constant functions, can one find a function $f \in L_2(dQ)$ such that $\|T^kf\| \downarrow c > 0$ as $k \to \infty$? Notice that if one forgets about transition functions and considers generally contractions on Hilbert space, this is easy to arrange.

In the following simple example, S is a contraction on a Hilbert space of square integrable functions with the property that the constant functions f are the only ones for which $\|f\| = \|T^nf\| \equiv 1$, $n = 1, 2, \cdots$, and yet $\|T^ng\| \downarrow c(g) > 0$ for all $g \not\equiv 0$ as $n \to \infty$. Let $e_0 \equiv 1$, $e_1, e_2, \cdots$ be a complete orthonormal family of functions, say on [0, 1]. Let S be the operator determined by

$$\begin{aligned} Se_0 &= e_0, \\ Se_j &= \lambda_j e_{j+1}, \qquad j = 1, 2, \cdots \end{aligned} \tag{48}$$

with $1 > \lambda_j > 0$ and $\Pi^\infty \lambda_j = L > 0$. The transformation S is clearly a contraction, and if

$$\varphi = \sum_{j=0}^{\infty} c_j e_j, \tag{49}$$

then

$$\|T^n\varphi\|^2 \downarrow |c_0|^2 + \sum_1^\infty |c_j|^2 L^2 / |\prod_1^{j-1} \lambda_i|^2. \tag{50}$$

An even more interesting example is given by a class of irreducible transient chains with an invariant measure. Let the invariant measure be $\pi = (\pi_i)$ so that $\pi T = \pi$ where T is now the matrix of one-step transition probabilities $t_{i,j}$. Notice that f with $f_i \equiv 1$ does not belong to $L_2(\pi)$ since the chain is assumed to be transient. Any nontrivial random walk, that is, with $t_{i,j} = t_{i-j}$ and $t_j < 1$ for all j, has

$$\|T^nf\|^2 = \sum_i |(T^nf)_i|^2 \pi_i \downarrow 0 \tag{51}$$

as $n \to \infty$ if $f \in L_2(\pi)$ with $\pi_i \equiv 1$. In fact, one would expect this to be the case with most irreducible transient chains with an invariant measure. However, the following simple irreducible transient Markov chains do not exhibit this behavior. Let $T = (t_{i,j})$ with

$$t_{i,j} = \begin{cases} 1 - q_i & \text{if } j = i + 1, \\ q_i & \text{if } j = i - 1, \\ 0 & \text{otherwise,} \end{cases} \tag{52}$$

and $0 < q_i < 1$ where $2 \sum q_i < 1$. Then one can show that there is an invariant measure with $\lim_{i \to \infty} \pi_i = \beta > 0$, $\lim_{i \to -\infty} \pi_i = \alpha > 0$ given by

$$\pi_j = \sum_{s=0}^{\infty} \alpha_j^{(s)} \quad \text{with} \quad \alpha_j^{(0)} \equiv 1, \qquad \alpha_j^{(s+1)} = q_j \alpha_j^{(s)} + q_{j+1} \alpha_{j+1}^{(s)}, \qquad s = 0, 1, \cdots. \tag{53}$$

Note that for all $f \in L_2(\pi)$, $\|Tf\| < \|f\|$, and since

$$\sum_i \pi_i |t_{i,j}^{(n)}|^2 \downarrow c_j > 0 \tag{54}$$

as $n \to \infty$, it follows that for most $f \in L_2(\pi)$, $\|T^n f\| \downarrow c > 0$ as $n \to \infty$. It seems natural to call the irreducible transient chains, for which $\|T^n f\| \downarrow 0$ as $n \to \infty$ for all $f \in L_2(\pi)$, purely nondeterministic and to say that the irreducible transient chains with an $f \in L_2(\pi)$, for which $\|T^n f\| \downarrow c > 0$, have a deterministic component.

Let us consider a real-valued function $g \in L_2(dQ)$. We can then consider the distribution function $F_g(\alpha) = Q(\{x | g(x) \leq \alpha\})$ of the function g. The following lemma remarks on a plausible and almost obvious continuity property of the operator T.

LEMMA 5. *If g is real-valued and $\|g\|^2 - \|Tg\|^2 < \epsilon^2$, $\epsilon > 0$, then*

$$F_{Tg}(\alpha - \epsilon) - \epsilon \leq F_g(\alpha) \leq F_{Tg}(\alpha + \epsilon) + \epsilon. \tag{55}$$

Now

$$\begin{aligned} \|g\|^2 - \|(Tg)(x)\|^2 &= \int Q(dx) \left\{ |g(x)|^2 - \left| \int P(x, dy) g(y) \right|^2 \right\} \\ &= \int Q(dx) \int P(x, dy) |g(y) - \int P(x, dz) g(z)|^2. \end{aligned} \tag{56}$$

Thus, if $\|g\|^2 - \|Tg\|^2 < \epsilon^2$, then

$$S_\epsilon = \left\{ x \Big| \int P(x, dy) |g(y) - \int P(x, dz) g(z)|^2 \geq \epsilon \right\} \tag{57}$$

has Q measure less than ϵ. If we look at the measure space of points (x, y) with measure μ generated by

$$R(A \times B) = \int_A Q(dx) P(x, B) = \mu(\{(x, y) | (x, y) \in A \times B\}), \tag{58}$$

then

$$\mu \left\{ (x, y) \Big| \, |f(y) - \int P(x, dz) f(z)| \geq \epsilon \right\} \leq \epsilon. \tag{59}$$

But this immediately yields the conclusion of the lemma. Notice that this implies that the distribution functions $F_{T^n g}$ of $T^n g$ converge as $n \to \infty$.

It is well known that the sequence of operators $T^{*k}\, T^k$ converge as $k \to \infty$ (see [6]). Let $M = \lim_{k \to \infty} T^{*k}\, T^k$ be the limiting self-adjoint operator. The following lemma indicates that the problem of determining when $L_2(dQ)$ is generated by the spaces $\mathfrak{U}_\infty$ and $\mathfrak{D}$ is related to the character of the operator M.

LEMMA 6. *Let the probability measure Q be an invariant measure of the transition probability operator T. The space $L_2(dQ)$ is generated by $\mathfrak{U}_\infty$ and $\mathfrak{D}$ if and only if M is a projection.*

Suppose that $\mathfrak{U}_\infty$ and $\mathfrak{D}$ generate $L_2(dQ)$. It is then clear that M is the projection operator leaving the functions of $\mathfrak{U}_\infty$ invariant and annihilating the functions of $\mathfrak{D}$. Conversely, if M is a projection, the functions left invariant by M are $\mathfrak{U}_\infty$ and the functions annihilated by M are the functions of $\mathfrak{D}$. The space $L_2(dQ)$ is then the direct sum of $\mathfrak{U}_\infty$ and $\mathfrak{D}$.

Notice that if the constant functions are the only functions belonging to $\mathfrak{U}_\infty$, then M is a self-adjoint positive, positive definite operator with Q as an invariant probability measure. Further, one is the only eigenvalue of absolute value one and is simple for M acting on $L_2(dQ)$. Thus, the Markov process with M as transition operator is mixing. A transition probability operator T is strongly mixing if

$$\sup_{\text{ess sup}\,|f|\leq 1} \|(T^n f)(x) - \int f\, dQ\| \to 0 \tag{60}$$

as $n \to \infty$ (see [3]). It is quite clear from lemma 5 that if $\mathfrak{U}_\infty$ is trivial (contains only the constant functions) and there is an f such that $\|T^n f\| \downarrow c > 0$, then M cannot be a strongly mixing operator.

LEMMA 7. *Let Q be an invariant probability measure of T. If T^k is normal for some positive integer k, then $M = M^*$ and $L_2(dQ)$ is generated by $\mathfrak{U}_\infty = \mathfrak{U}_\infty^*$ and $\mathfrak{D} = \mathfrak{D}^*$.*

Since $T^k T^{*k} = T^{*k} T^k$, it follows that $M = \lim_{n\to\infty} T^{nk} T^{*nk} = M^*$. Further, M is a projection for $M = \lim_{n\to\infty} T^{*2nk} T^{2nk} = (\lim_{n\to\infty} T^{*nk} T^{nk})^2 = M^2$. The conclusion follows immediately from lemma 6.

REFERENCES

[1] J. R. BLUM and D. L. HANSON, "On invariant probability measures I," *Pacific J. Math.*, Vol. 10 (1960), pp. 1125–1129.

[2] T. E. HARRIS, "On chains of infinite order," *Pacific J. Math.*, Vol. 5 (1955), pp. 707–724.

[3] A. KOLMOGOROV and YU. A. ROSANOV, "On a strong mixing condition for stationary Gaussian processes," *Teor. Verojatnost. i Primenen.*, Vol. 5 (1960), pp. 222–227. (In Russian; English summary.)

[4] M. ROSENBLATT, "Equicontinuous Markov operators," *Theor. Probability Appl.*, Vol. 9 (1964), pp. 205–222.

[5] ———, "Almost periodic transition operators acting on the continuous functions on a compact space," *J. Math. Mech.*, Vol. 13 (1964), pp. 837–847.

[6] GIAN-CARLO ROTA, "An 'Alternierende Verfahren' for general positive operators," *Bull. Amer. Math. Soc.*, Vol. 68 (1962), pp. 95–102.