ASYMPTOTIC PERIODICITY OF THE ITERATES
OF MARKOV OPERATORS

BY
A. LASOTA, T. Y. LI¹ AND J. A. YORKE²

ABSTRACT. We say $P: L^1 \rightarrow L^1$ is a Markov operator if
(i) $Pf \geq 0$ for $f \geq 0$ and (ii) $\|Pf\| = \|f\|\text{ if } f \geq 0$. It is shown that any Markov operator $P$ has
a certain spectral decomposition if, for any $f \in L^1$ with $f \geq 0$ and $\|f\| = 1$,
$P^n f \rightarrow \mathcal{F}$ when $n \rightarrow \infty$, where $\mathcal{F}$ is a strongly compact subset of $L^1$. It follows
from this decomposition that $P^n f$ is asymptotically periodic for any $f \in L^1$.

Introduction. In the theory of stationary discrete time Markov processes, the
sequence $\{P^n\}$ of the iterates of a linear operator $P: L^1 \rightarrow L^1$ plays an important
role. This operator is positive, bounded and generalizes the notion of the transition
function [2, 4]. Asymptotic behavior of $\{P^n\}$ depends heavily on the spectral
properties of $P$ [12], but the construction of the spectral decomposition for $P$ is
not an easy problem. It can be solved under some additional hypothesis such as
quasicompactness [14] of $P$ or the uniform stability of $\{P^n\}$ in mean [6]. In general,
there is no technique available for studying the asymptotic behavior of $\{P^n\}$ even
if $P$ is given by a simple explicit formula. A typical example where such a situation
occurs is the statistical theory of deterministic systems. If $S$ is a mapping of the
unit interval into itself and $f$ is a probability density, then the sequence $\{P_S^n\}$,
with

$$(0.1) \quad P_S f(x) = \frac{d}{dx} \int_{S^{-1}[0,x]} f(y) dy,$$

describes the evolution of densities generated by the deterministic system $\{S^n\}$. Despite the simplicity of the formula (0.1) the behavior of $\{P_S^n\}$ is well under-
stood only for a narrow class of transformations. In particular, F. Hofbauer and G.
Keller proved in [3] that for piecewise expanding transformations $S$ the sequences
$\{P_S^n f\} (f \in L^1)$ are asymptotically periodic and M. Misiurewicz obtained a somewhat
analogous result [13] for a class of transformations with negative Schwarzian
derivative.

In the present paper we prove an asymptotic decomposition theorem for a class
of Markov operators acting on an arbitrary space $L^1(X, \Sigma, \mu)$ with a $\sigma$-finite
measure $\mu$. For these operators all the sequences $\{P^n f\}$ with $f \in L^1$ are asymptotically
periodic. Our main assumption is the convergence of $\{P^n f\}$ with normalized non-
negative $f$ to a common compact set $\mathcal{F}$. This pointwise convergence is sometimes

Received by the editors July 27, 1983 and, in revised form, January 5, 1984.
1980 Mathematics Subject Classification. Primary 58F11, 28A65.
Key words and phrases. Markov operator, invariant measures, spectral decomposition, asymptotic
periodicity.
¹This research was supported under National Science Foundation grant MCS 83-01408.
²This research was supported under National Science Foundation grant MCS 81-17967 and its
continuation.

©1984 American Mathematical Society
0002-9947/84 $1.00 + \$0.25 per page

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
easy to verify and we give applications to Markov operators studied by G. Keller and to Markov semigroups arising in the theory of integral equations with an application to the linear Boltzmann equation.

The proof of our main result is quite long, so it will be presented at the end of the paper. Thus the organization of the paper goes as follows. §1 contains the main result—the decomposition theorem. §2 gives some examples of application. §§3–6 are devoted to the proof. Namely, in §§3–4, we derive some simple properties of the limit set of \( \{P^n f\} \) under additional assumptions that \( P1 = 1 \) and \( \mu \) is finite. Then in §5 the proof of the decomposition theorem is given with the same additional conditions. §6 shows how the conditions \( P1 = 1 \) and finiteness of \( \mu \) may be released.

1. Let \( (X, \Sigma, \mu) \) be a measure space with a nonnegative \( \sigma \)-finite measure \( \mu \). Write \( L^1 = L^1(X, \Sigma, \mu) \). A linear operator \( P: L^1 \to L^1 \) will be called a Markov operator if it satisfies the following two conditions:
   
   (a) \( Pf \geq 0 \) for \( f \geq 0 \), \( f \in L^1 \), 
   (b) \( ||Pf|| = ||f|| \) for \( f \geq 0 \), \( f \in L^1 \),

   where \( || \cdot || \) stands for the norm in \( L^1 \). By \( D = D(X, \Sigma, \mu) \) we shall denote the set of all (normalized) densities on \( X \), that is,

   \[ D = \{ f \in L^1 : f \geq 0 \text{ and } ||f|| = 1 \} \].

   We say that an operator \( P \) is strongly constrictive if there exists a strongly compact set \( \mathcal{F} \subset L^1 \) such that

   \[ \lim_{n \to \infty} d(P^n f, \mathcal{F}) = 0 \quad \text{for } f \in D, \]

   where \( d(g, \mathcal{F}) \) denotes the distance between \( g \) and \( \mathcal{F} \), that is, the infimum of \( ||g - f|| \) for \( f \in \mathcal{F} \). For \( g \in L^1 \), the support of \( g \), denoted by \( \text{supp} \, g \), is defined to be the set \{\( x : g(x) \neq 0 \)\}.

   Our main result, Theorem 1.1, describes how \( P^n f \) must approach a finite-dimensional space which is independent of \( f \). Crucial to this result is showing that there are densities \( g_1, \ldots, g_r \) which are a basis for that finite-dimensional space and which permute under the action of \( P \).

   **Theorem 1.1.** Let \( P \) be a strongly constrictive Markov operator. Then there exists a sequence of densities \( g_1, \ldots, g_r \) and a sequence of bounded linear functionals \( \lambda_1, \ldots, \lambda_r \) such that

   \[ \lim_{n \to \infty} \left\| P^n \left( f - \sum_{i=1}^r \lambda_i(f)g_i \right) \right\| = 0 \quad \text{for } f \in L^1. \]

   The densities \( \{g_i\} \) have mutually disjoint supports (\( g_ig_i = 0 \) for \( i \neq j \)) and

   \[ Pg_i = g_{\alpha(i)} \]

   where \( \{\alpha(1), \ldots, \alpha(r)\} \) is a permutation of the integers \( \{1, \ldots, r\} \).

   From Theorem 1.1 it follows immediately that the \( n \)-th iterate \( P^n \) of \( P \) can be written in a very special form.

   **Corollary 1.2.** If \( P \) is a strongly constrictive Markov operator, then

   \[ P^n f = \sum_{i=1}^r \lambda_i(f)g_{\alpha^n(i)} + R_n f \quad \text{for } f \in L^1, \]

   where \( R_n f = f - \sum_{i=1}^r \lambda_i(f)g_{\alpha^n(i)} \).
where \( \alpha^n \) denotes the \( n \)th iterate of the permutation \( \alpha \) and the remainder \( R_n \) converges strongly to zero as \( n \to \infty \). Thus every sequence \( \{P^n f\} \) is asymptotically periodic with a period which does not exceed \( r! \).

It is easy to find an evaluation of the integer \( r \) in (1.1). In fact, since \( P \) is strongly constrictive there exists a function \( g \in L^1 \) which bounds all sequences \( \{P^n f\} \) with \( f \in D \). More precisely, \( g \) has the property that
\[
\lim_{n \to \infty} \| (P^n f - g)^+ \| = 0 \quad \text{for} \quad f \in D,
\]
where \((u)^+ = \max(0, u)\). Then, according to (1.3) and (1.4),
\[
Lf = \lim_{n \to \infty} P^{rn} f = \sum_{i=1}^{r} \lambda_i(f) g_i \leq g \quad \text{for} \quad f \in D.
\]
In particular, setting \( f = g_k \), we have \( Lg_k = g_k \leq g \). Let \( \bar{g} = \sum_{i=1}^{r} g_i \). Then \( \bar{g} \leq g \), since the supports of \( g_i \) are disjoint. Integrating \( \bar{g} \) over \( X \), we have
\[
\int_X g \, d\mu \geq \int_X \bar{g} \, d\mu = \sum_{i=1}^{r} \int_X g_i \, d\mu = r
\]
which means that
\[
(1.5) \quad r \leq \|g\|.
\]
Condition (1.5) enables us to study more precisely the properties of the Markov process \( \{P^n\} \). In some general cases it also guarantees the asymptotical stability of \( \{P^n\} \).

A Markov operator is said to be asymptotically stable if there is a unique \( f_* \in D \) such that
\[
\lim_{n \to \infty} \|P^n f - f_*\| = 0 \quad \text{for all} \quad f \in D.
\]
From Theorem 1.1 and inequality (1.5) follows

**COROLLARY 1.3 (SEE ALSO [10]).** If \( P \) is strongly constrictive and \( \|g\| < 2 \), then \( P \) is asymptotically stable.

From the point of view of applications, the condition \( \|g\| < 2 \) is quite restrictive. Therefore we give the following alternative.

**COROLLARY 1.4.** Let \( P \) be a strongly constrictive Markov operator. Assume that there is a set \( A \subset \Sigma \) with nonzero measure with the property that for every \( f \in D \) there is an integer \( n_0(f) \) such that
\[
(1.6) \quad P^n f(x) > 0 \quad \text{for} \quad x \in A, \quad n \geq n_0(f).
\]
Then \( P \) is asymptotically stable.

**Proof.** It is sufficient to show that \( r = 1 \) in (1.1). But if \( r > 1 \), then there exists an integer \( k \) such that \( A \) is not contained in the support of \( g_k \) and the sequence \( \{P^n f\} \) with \( f = g_k \) does not satisfy (1.6). Thus \( r = 1 \). \( \square \)

**Remark 1.5.** Observe that in order to prove that an operator \( P \) is strongly constrictive it is not necessary to verify the condition \( d(P^n f, F) \to 0 \) for all possible \( f \in D \). Due to the fact that Markov operators are linear contractions, it is sufficient to verify this condition for all \( f \) belonging to an arbitrary dense subset of \( D \).
2. Our first application refers to piecewise expanding transformations. Here we consider the domain as the unit interval. This class of mapping was recently studied by G. Keller [7] who proved a decomposition theorem for the corresponding Markov operators based on the spectral theorem of Ionescu Tulcea and Marinescu [5]; see [7] for a description of how these techniques extended to piecewise analytic maps on \([0, 1] \times [0, 1]\). For simplicity of notation we restrict our attention here to maps on \([0, 1]\). We shall show that such a decomposition immediately follows from Theorem 1.1.

Consider a transformation \(S : [0, 1] \to [0, 1]\) which satisfies the following conditions:

(i) There is a partition \(0 = a_0 < \cdots < a_m = 1\) such that for each integer \(i\) the restriction \(S_i\) of \(S\) to the interval \((a_{i-1}, a_i)\) is a \(C^2\) function;

(ii) \(\inf |S'(x)| > 1\) \((x \neq a_i)\);

(iii) \(\sup (|S''(x)|/(S'(x))^2) < \infty\) \((x \neq a_i)\).

Conditions (i) and (ii) are quite understandable; condition (iii) is equivalent to

\[
\sup_x (|\varphi''_i(x)|/|\varphi'_i(x)|) < \infty \quad \text{for} \ x \in S((a_{i-1}, a_i)),
\]

where \(\varphi_i\) denotes the inverse function to \(S_i\). It is easy to write an explicit formula for \(P_S\), namely

\[
P_S f(x) = \sum_{i=1}^m |\varphi'_i(x)| f(\varphi_i(x)) 1_{C_i}(x),
\]

where \(1_{C_i}\) denotes the characteristic function on the set \(C_i = S_i((a_{i-1}, a_i))\). It is easy to evaluate the total variation, denoted \(V\), of the functions \(P^n_S f\) for large \(n\). Namely, if \(S\) satisfies (i)-(iii), then there exists a constant \(K\) independent of \(f\) such that

\[
\limsup_{n \to \infty} V P^n_S f \leq K
\]

for every \(f \in D[0, 1]\) of bounded variation (see [10]). The set

\[
\mathcal{F} = \{h \in D[0, 1] : V h \leq K\}
\]

is compact in \(L^1\) and, by virtue of (2.1), \(\{P^n_S f\}\) converges to \(\mathcal{F}\). Thus we may summarize our results in the following

**Proposition 2.1.** If \(S : [0, 1] \to [0, 1]\) satisfies conditions (i)-(iii), then the operator \(P_S\) is strongly constrictive. Consequently, for every \(f \in L^1[0, 1]\) the iterates \(P^n_S f\) can be written in the form (1.3) and the sequence \(\{P^n_S f\}\) is asymptotically periodic.

The next examples concern integral operators. We begin with a simple illustrative example. Let

\[
(Pu)(x) = \int_0^1 b(x, y) u(y) \, dy,
\]

where \(b : (0, 1) \times (0, 1) \to [0, \infty)\) is measurable and is stochastic, that is,

\[
\int_0^1 b(x, y) \, dx = 1 \quad \text{for} \ y \in (0, 1).
\]

Assume for some \(b_0 > 0\) that \(b(x, y) \leq b_0\) for all \(x\) and \(y\) in \((0, 1)\).
PROPOSITION 2.2. The operator $P$ in (2.2) is strongly constrictive.

For some choices of $b$, the periodicity of $P$ is nontrivial; that is, the number $r$ of densities mentioned in Theorem 1.1 is at least 2. Suppose, for example, $b(x, y) > 0$ implies that either $x < \frac{1}{2}$ and $y > \frac{1}{2}$ or that $x > \frac{1}{2}$ and $y < \frac{1}{2}$. It follows that the number $r$ of densities with disjoint supports will be at least 2, one having support in $[0, \frac{1}{2}]$ and the other having support in $[\frac{1}{2}, 1]$.

PROOF OF PROPOSITION 2.2. Notice that $(Pf)(x) \leq b_0$ for every $f \in D$. Let $\mathcal{F}_1$ be $\{f : 0 \leq f \leq b_0\}$ and let $\mathcal{F} = P(\mathcal{F}_1)$. Krasnosel’skii [8] proves that since $\mathcal{F}_1$ is weakly compact, $\mathcal{F}$ is strongly compact. Since $P^2f \in \mathcal{F}$, we may trivially conclude $P^nf \to \mathcal{F}$. □

This application does not use the full power of Theorem 1.1 since the convergence to $\mathcal{F}$ is very strong and uniform. Nonetheless variations on this theme suggest themselves: if the problem is defined on $R$ rather than $(0,1)$, and we assume $b(x, y) \leq b_0(x)$ for all $x, y \in R$, where $b_0 \in L^1$, then the result is still true: $P$ is still strongly constrictive. Here, however, it is necessary to use the compactness results in Krasnosel’skii with a little more subtlety to show $P^2(D)$ lies in a strongly compact set. The next example is another variation on the theme, in which the convergence to the compact set $\mathcal{F}$ is not at all uniform and the computations are less trivial.

We now consider a linear version of the Boltzmann equation (see [9])

\begin{equation}
\left(2.3\right) \frac{\partial u(t, x)}{\partial x} + u(t, x) = \int_0^\infty b(x, y)u(t, y) \, dy, \quad t \geq 0, \quad x > 0,
\end{equation}

where the kernel $b(x, y)$ satisfies the following conditions:

(I) The function $b : (0, \infty) \times (0, \infty) \to (0, \infty)$ is measurable and $b(x, y) > 0$ for $x, y > 0$ and stochastic, i.e.,

$\int_0^\infty b(x, y) \, dy = 1 \quad \text{for} \ x > 0.$

(II) There is a function $g \in L^1(0, \infty)$ such that

$b(x, y) \leq g(x)(1 + y) \quad \text{for} \ x > 0, \ y > 0.$

(III) There are constants $\alpha \geq 0, \ 0 \leq \beta < 1$ such that

$\int_0^\infty xb(x, y) \, dx \leq \alpha + \beta y \quad \text{for} \ y > 0.$

All these conditions are automatically satisfied when $b$ is a kernel corresponding to the Tjon-Wu representation of the Boltzmann equation (see [9]); that is,

$b(x, y) = \begin{cases} -e^y \operatorname{Ei}(-y) & \text{for} \ x \leq y, \\ -e^y \operatorname{Ei}(-x) & \text{for} \ x > y, \end{cases}$

where $\operatorname{Ei}$ is the exponential integral

$-\operatorname{Ei}(-x) = \int_x^\infty e^{-y} \frac{dy}{y}.$
It is easy to verify that the operator $Bf(x) = \int_0^\infty b(x, y)f(y)\,dy$ maps $L^1(0, \infty)$ into itself and is Markovian. Thus we may avoid the difficulties related with classical solutions and consider (2.3) as an evolution equation
\[
\frac{du}{dt} + u = Bu, \quad u(0) = u_0
\]
in the space $L^1(0, \infty)$. Since $B$ is a bounded operator, (2.4) has a unique solution for each $u_0 \in L^1$ and it is given by the formula
\[
u(t) = e^{(B - I)t}u_0 = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} B^n u_0.
\]

We are going to prove the following result.

**Proposition 2.3.** If $b$ satisfies (I)–(III), then for every $u_0 \in D$ the solution $u(t)$ of (2.4) has a limit $u_\ast = \lim_{t \to -\infty} u(t)$ in $L^1$ norm, and the limiting function $u_\ast$ is independent of the initial condition $u_0$.

**Proof.** Formula (2.5) implies
\[
\lim_{t \to -\infty} u(t) = \lim_{n \to \infty} B^n u_0
\]
whenever the limit on the right-hand side exists. Thus it is enough to prove that $\{B^n\}$ is asymptotically stable. From (I) it follows that, for every $f \in D(0, \infty)$,
\[
B^n f(x) > 0 \quad \text{for} \quad x > 0, \quad n = 1, 2, \ldots.
\]
Thus the condition (1.6) of Corollary 1.4 is satisfied. It remains to verify that $B$ is strongly constrictive. From (III) an easy calculation gives
\[
\int_0^\infty (1 + x)B^n f(x)\,dx \leq \sigma \|f\| + \beta^n \int_0^\infty (1 + x)f(x)\,dx,
\]
where $\sigma = (1 + \alpha - \beta)/(1 - \beta)$, and from (II) it follows that
\[
|B^n f(x)| \leq \int_0^\infty b(x, y)|B^{n-1} f(y)|\,dy
\leq g(x) \int_0^\infty (1 + y)|B^{n-1} f(y)|\,dy.
\]
Thus finally,
\[
|B^n f(x)| \leq g(x) \left( \sigma \|f\| + \beta^{n-1} \int_0^\infty (1 + y)f(y)\,dy \right).
\]
In particular, when $f \in D(0, \infty)$ and
\[
\int_0^\infty (1 + x)f(x)\,dx < \infty
\]
we have
\[
B^n f(x) \leq g_1(x) + \epsilon_n(x),
\]
where $g_1 = \sigma g \in L^1(0, \infty)$ and $\|\epsilon_n\| \to 0$. Inequality (2.6) implies that for every $f \in D$ that satisfies (2.5) the sequence $\{B^n f\}$ converges to the set $\mathcal{T}_1 = \{f : 0 \leq f \leq g_1\}$ and consequently also to $\mathcal{T}_2 = B(\mathcal{T}_1)$. Every bounded positive integral
operator from $L^1$ to $L^1$ maps sets of the form $\mathcal{F}_1$ into compact sets. Hence $\mathcal{F}_2$ is compact, so $B$ is constritive. □

A result analogous to Proposition 2.3 with the additional assumption that $b(x,y)$ is decreasing in $x$, was proved in [9].

Observe that in the last example we have established only convergence of $P^n f$ to $\mathcal{F}$ for each $f$, not uniform convergence with respect to $f \in D$.

3. In this section and next, we will suppose a strongly constritive Markov operator $P$ on $(X, \Sigma, \mu)$ is given with the additional assumption that $P1 = 1$ and $\mu(X) < \infty$. For $f \in D$, we also assume $P^n f$ converges to a set $\mathcal{F}$ which is a strongly compact subset of $L^1$. If $A \in \Sigma$, we shall denote the characteristic function of $A$ by $1_A$.

**Definition 3.1.** A set $A \in \Sigma$ will be called a nice set if $P^n 1_A$ is the characteristic function for each positive integer $n$. The characteristic function $1_A$ of a nice set $A$ will be called a nice function. The following lemma establishes that the family $A$ of all nice sets forms an algebra $\mathcal{A}$.

**Lemma 3.2.**
(1) If $A$ is a nice set, then $X - A$ is a nice set.
(2) If $A_1, A_2$ are nice sets, then $A_1 \cup A_2$ is a nice set.

**Proof.** (1) Since $1_{X - A} = 1_X - 1_A$, it follows that

$$P^n 1_{X - A} = P^n 1 - P^n 1_A = 1 - P^n 1_A = 1_{X - B_n},$$

where $B_n$ is the set which has $P^n 1_A$ as its characteristic function. Hence $X - A$ is a nice set.

(2) Since $A_1, A_2$ are nice sets, for fixed $n$ we have

$$P^n 1_{A_1} = 1_{B_1} \quad \text{and} \quad P^n 1_{A_2} = 1_{B_2}$$

for some sets $B_1$ and $B_2$ contained in $X$. For $i = 1, 2$, we have

$$1_{A_i} \leq 1_{A_1 \cup A_2} \leq 1_{A_1} + 1_{A_2}. \quad (3.1)$$

Hence,

$$1_{B_i} \leq P^n 1_{A_1 \cup A_2} \leq 1_{B_1} + 1_{B_2}. \quad (3.2)$$

While the inequality on the right-hand side of (3.2) implies that $P^n 1_{A_1 \cup A_2} = 0$ for $x \notin B_1 \cup B_2$, the inequality on the left, along with the inequality $P^n 1_{A_1 \cup A_2} \leq P^n 1 = 1$, gives $P^n 1_{A_1 \cup A_2} = 1$ for $x \in B_1 \cup B_2$. Hence $A_1 \cup A_2$ is a nice set. □

The following corollary with be used later.

**Corollary 3.3.** Let $A_1, A_2$ be two nice sets with disjoint supports and let $P 1_{A_i} = B_i$, $i = 1, 2$. Then $B_1 \cap B_2 = \emptyset$.

**Proof.** Since $1_{A_1 \cup A_2} = 1_{A_1} + 1_{A_2}$, it follows that

$$P 1_{A_1 \cup A_2} = P 1_{A_1} + P 1_{A_2} = 1_{B_1} + 1_{B_2}.$$ 

On the other hand, $P 1_{A_1 \cup A_2} \leq P 1 = 1$; so $1_{B_1} + 1_{B_2} \leq 1$. That is, $B_1 \cap B_2 = \emptyset$. □

The following two lemmas show that the algebra $\mathcal{A}$ of nice sets is finite, that is, there are only finitely many nice sets.
LEMMA 3.4. There exists $\delta > 0$ such that $\mu(A) > \delta$ for nonempty sets $A \in \mathcal{A}$.

PROOF. Choose $A \in \mathcal{A}$ and define $f = 1_A/\mu(A)$. Since $f \in D$, we have $P^n f \to \mathcal{F}$ as $n \to \infty$. Thus, given $\varepsilon \in (0, 1)$, there is a sequence $\{g_n\} \subset \mathcal{F}$ such that

$$
\|P^n f - g_n\| < 1 - \varepsilon \quad \text{for large } n.
$$

Since $\mathcal{F}$ as a compact set is also weakly compact, there is $\delta > 0$, such that for any set $B$ with $\mu(B) < \delta$, we have $\int_B g_n \, d\mu < \varepsilon$. Let $A_n = \text{supp } P^n f$. Then it is easy to see that $\mu(A_n) = \mu(A)$ for all $n$. We claim that $\mu(A_n) > \delta$. Suppose this is not the case. Then $\int_{A_n} g_n \, d\mu < \varepsilon$ implies that

$$
\|P^n f - g_n\| > \int_{A_n} |P^n f - g_n| \, d\mu > \int_{A_n} P^n f \, d\mu - \int_{A_n} g_n \, d\mu \geq 1 - \varepsilon,
$$

which is a contradiction to (3.3). \qed

A nonempty set $B$ is an algebra $\mathcal{B}$ of sets is called an atom if the only subsets of $B$ that are in $\mathcal{B}$ are $B$ and $\emptyset$. Hence, the set of atoms in an algebra must be mutually disjoint. From Lemma 3.4, the number of atoms in $\mathcal{A}$ must be finite since $\mu(X) < \infty$. Let $\{A_1, \ldots, A_r\}$ be the set of atoms in $A$, and write $1_i = 1_{A_i}$ for $1 \leq i \leq r$.

LEMMA 3.5. There exists a permutation $\alpha$ of the integers $1, \ldots, r$ such that $P1_i = 1_{\alpha(i)}$.

PROOF. It follows easily from the definition of nice function that $P1_i$ is a nice function for each $1 \leq i \leq r$. Let $1_{B_i} = P1_i$.

By Corollary 3.3, $1_{B_i} = P1_i$ and $1_{B_j} = P1_j$ has disjoint supports for $i \neq j$ and $B_1, \ldots, B_r$ is a collection of mutually disjoint elements of $A$. Since the algebra $\mathcal{A}$ is finite, $B_i$ must be atoms and consequently the sequence $B_1, \ldots, B_r$ is a permutation of $A_1, \ldots, A_r$. \qed

4. For $f \in L^1$, $\{P^n f\}$ is precompact since $P$ is strongly constrictive. Let $\Omega(f)$ be the set of limit points of $\{P^n f\}$ and $Q = \bigcup_{f \in L^1} \Omega(f)$. We show that $Q$ is a linear subspace of $L^1$. It is clear that all nice sets are in $Q$. The main result of this section is the following

THEOREM 4.1. $Q$ is finite dimensional with basis $\{1_1, \ldots, 1_r\}$.

We begin the proof of this theorem with a set of lemmas. The first two, in fact, are quite special results from the general theory of dynamical systems. We state them in their simple forms for the convenience of the reader.

LEMMA 4.2. If $f \in Q$, then $f \in \Omega(f)$.

PROOF. Let $g \in L^1$ be such that $P^{nk} g \to f$ as $k \to \infty$. That is, given $\varepsilon > 0$, there is a positive integer $N$ such that

$$
\|P^{nk} g - f\| < \varepsilon/2 \quad \text{for } k \geq N.
$$

It follows that, for $k \geq N$,

$$
\|P^{nk-nN} f - f\| < \varepsilon.
$$

Thus, the sequence $\{P^n f\}$ contains a subsequence which converges to $f$. \qed
**Lemma 4.3.** If \( f \in Q \), then \( \| Pf \| = \| f \| \).

**Proof.** First notice \( \| f \| \geq \| Pf \| \geq \cdots \geq \| P^n f \| \geq \cdots \). By Lemma 4.2, there is a subsequence \( \{ n_k \} \) such that \( P^{n_k} f \to f \) as \( k \to \infty \). Hence \( \| P^{n_k} f \| \to \| f \| \), and \( \| P^n f \| \) is constant. □

For a real number \( u \), we shall write \( u^+ = \max(0, u) \) and \( u^- = \max(0, -u) \).

**Lemma 4.4.** If \( \| Pf \| = \| f \| \), then \( \text{supp} Pf^+ \) and \( \text{supp} Pf^- \) are (essentially) disjoint.

**Proof.** Let \( h = \min(Pf^+, Pf^-) \). Suppose \( \text{supp} Pf^+ \) and \( \text{supp} Pf^- \) are not disjoint. Then, \( h \neq 0 \). We then have

\[
\| Pf \| = \| Pf^+ - Pf^- \| = \| (Pf^+ - h) - (Pf^- - h) \| \\
\leq \| Pf^+ - h \| + \| Pf^- - h \| < \| Pf^+ \| + \| Pf^- \| \\
= \| f^+ \| + \| f^- \| = \int_X f^+ \, d\mu + \int_X f^- \, d\mu = \| f \| ,
\]

which is a contradiction. □

**Lemma 4.5.** If \( f_1 \) and \( f_2 \) are nonnegative and have the same support, then \( Pf_1 \) and \( Pf_2 \) have the same support.

**Proof.** We only need to show that \( \text{supp} Pf_1 \subseteq \text{supp} Pf_2 \). Let \( \text{supp} Pf_2 = B_2 \).

Write

\[
P f_1 = P f_1 \cdot 1_{B_2} + P f_1 \cdot 1_{X-B_2}.
\]

For \( c > 2 \), define \( f_c = \min(cf_2, f_1) \). Then \( f_c \) has the following obvious properties:

(1) \( f_c \leq f_1 \),

(2) \( \lim_{c \to \infty} f_c(x) = f_1(x) \),

(3) \( \text{supp} P f_c \subseteq \text{supp} P f_2 = B_2 \).

By (1) and (2) and the Lebesgue dominated convergence theorem, we have

\[
\| P f_1 - P f_c \| = \int_X P f_1(x) - f_c(x) \, d\mu = \int_X (f_1(x) - f_c(x)) \, d\mu \to 0 \quad \text{as} \quad c \to \infty .
\]

By (1) and (3),

\[
\| P f_1 - P f_c \| \geq \int_{X-B_2} (P f_1 - P f_c) \, d\mu = \int_{X-B_2} P f_1 \, d\mu \\
= \| P f_1 \cdot 1_{X-B_2} \| .
\]

Hence, \( P f_1 \cdot 1_{X-B_2} = 0 \). That is, \( \text{supp} P f_1 \subseteq B_2 = \text{supp} P f_2 \). □

**Proposition 4.6.** If \( f \in \Omega(g) \) for some \( g \in L^1 \), then for each real \( v \), \( f^{-1}(-\infty, v) \) is a nice set.

**Proof.** We first assume that \( \mu(f^{-1}(v)) = 0 \). Write \( h_1 = 1_{f^{-1}(-\infty, v)} \) and \( h_2 = 1_{f^{-1}[v, +\infty)} \). Then, letting \( g_0 = f - v \), we see that \( g_0 \in Q \) and

\[
\text{supp} g_0^+ = \{ x : f(x) > v \} = \text{supp} h_2 , \\
\text{supp} g_0^- = \{ x : f(x) < v \} = \text{supp} h_1 .
\]

By Lemmas 4.3, 4.4 and 4.5, we have \( \text{supp} P^n h_1 \cap \text{supp} P^n h_2 \) are disjoint. Hence, \( P1 = 1 \) implies that \( P^n h_1 + P^n h_2 = 1 \). Therefore, \( P^n h_1 \) is a characteristic function for each \( n \geq 0 \).
Suppose \( \mu(f^{-1}(v)) \neq 0 \). Then, since \( \mu(X) < \infty \) and \( f^{-1}(c) \) are disjoint for different \( c > 0 \), the set of \( c \) with \( \mu(f^{-1}(c)) \neq 0 \) is at most countable. Thus its complement is dense. Let \( v_i \) be an increasing sequence with \( v_i \rightarrow v \) and \( \mu(f^{-1}(v_i)) = 0 \). Then, since \( f^{-1}(-\infty, v) = \bigcup_i f^{-1}(-\infty, v_i) \) and the nice sets are an algebra, \( f^{-1}(-\infty, v) \) is also a nice set.

**Proof of Theorem 4.1.** It suffices to prove that every \( f \in Q \) can be written as a linear combination of the \( 1_i \). By Proposition 4.6, for each real \( v \), \( f^{-1}(-\infty, v) \) is a nice set. Since the family of nice sets is finite, there is a finite number of different values, say \( v_1 < \cdots < v_k \), for which \( f^{-1}(-\infty, v_1), \ldots, f^{-1}(-\infty, v_k) \) are different sets. Thus, \( f \) has the form for appropriate choice of real number \( \lambda_i \),

\[
f = \sum_{i=1}^{k-1} \lambda_i (1_{f^{-1}(-\infty, v_{i+1})} - 1_{f^{-1}(-\infty, v_i)}).
\]

Since \( 1_{f^{-1}(-\infty, v_{i+1})} - 1_{f^{-1}(-\infty, v_i)} \) are nice functions, \( f \) is a linear combination of the \( 1_i \).

5. Let \( P : L^1 \rightarrow L^1 \) be a Markov operator and let a sequence of densities \( \{g_i\} \) \( (i = 1, \ldots, r) \) with mutually disjoint supports be given. We assume that \( P \) permutes the densities \( g_i \), i.e.,

\[
P g_i = g_{\alpha(i)},
\]

where \( \alpha \) is a permutation of the integers \( 1, \ldots, r \).

**Definition 5.1.** We say that the \( \{g_i\} \) give an asymptotic decomposition of \( P \) if there is a sequence of linear functionals \( \{\lambda_i\} \) on \( L^1 \) such that

\[
\lim_{n \rightarrow \infty} \left\| P^n \left( f - \sum_{i=1}^{r} \lambda_i(f) g_i \right) \right\| = 0 \quad \text{for} \ f \in L^1.
\]

Observe that since \( g_1, \ldots, g_r \) are linearly independent, the functionals \( \lambda_i \) are necessarily bounded. In verifying that the \( \{g_i\} \) give an asymptotic decomposition the following criterion is quite useful.

**Lemma 5.2.** The densities \( \{g_i\} \) give an asymptotic decomposition if and only if for every \( f \in L^1 \) and \( \varepsilon > 0 \), there are constants \( C_1, \ldots, C_r \) and a positive integer \( n_0 \) such that

\[
\left\| P^{n_0} f - \sum_{i=1}^{r} C_i g_i \right\| \leq \varepsilon.
\]

**Proof.** The "only if" part is obvious. To prove the "if" part let \( n_k \) and \( C_i^k \), \( i = 1, \ldots, r \), be chosen such that, with indices rearranged if necessary,

\[
\left\| P^{n_k} f - \sum_{i=1}^{r} C_i^k g_{\alpha(n_k(i))} \right\| \rightarrow 0.
\]

Since \( \{C_i^k\} \) is a bounded sequence for each \( i \), by choosing a subsequence if necessary, we suppose that, for each \( i \), \( C_i^k \) converges to a constant \( \lambda_i(f) \). Write

\[
\varepsilon_n = \left\| P^n \left( f - \sum_{i=1}^{r} \lambda_i(f) g_i \right) \right\|.
\]
It follows from the subsequence construction that $\varepsilon_{n_k} \to 0$ as $n \to \infty$. Furthermore, $\varepsilon_n$ is a decreasing sequence since $\|P^n h\|$ is nonincreasing for each $h \in L^1$. Hence $\varepsilon_n \to 0$.

It remains to show that $\lambda_i(f)$ are linear. Observe that $\lambda_i(f)$, $i = 1, \ldots, r$, are uniquely defined since the $g_i$ are linearly independent. Moreover, conditions

$$
P^n \left( f_1 - \sum_{i=1}^r \lambda_i(f_1) g_i \right) \to 0 \quad \text{and} \quad P^n \left( f_2 - \sum_{i=1}^r \lambda_i(f_2) g_i \right) \to 0
$$

imply that

$$
P^n \left( f_1 + f_2 - \sum_{i=1}^r (\lambda_i(f_1) + \lambda_i(f_2)) g_i \right) \to 0.
$$

Hence, $\lambda_i(f_1) + \lambda_i(f_2) = \lambda_i(f_1 + f_2)$ for each $i$. \(\square\)

**Proposition 5.3.** Let $P$ be a strongly constrictive Markov operator on measure space $(X, \Sigma, \mu)$. Assume $P1 = 1$ and $\mu(X) < \infty$. Then, there is a sequence of densities $\{g_i\}$, $i = 1, \ldots, r$, with mutually disjoint supports, which gives an asymptotic decomposition of $P$.

**Proof.** Let $\{A_i\}, \{1_i\}$ and $\alpha$ be as in Lemma 3.5. Define $g_i = 1_i/\mu(A_i)$. By virtue of Theorem 4.1, all elements of $Q$ are of the form $\sum_{i=1}^r C_i g_i$. Thus, for $f \in L^1$, there is a subsequence of $\{P^n f\}$ convergent to a function of the form $\sum_{i=1}^r C_i g_i$. Consequently for every $\varepsilon > 0$, condition (5.3) is satisfied by infinitely many integers. \(\square\)

6. In this section, we shall eliminate the conditions $P1 = 1$ and $\mu(X) < \infty$ of Proposition 5.3 and thus give a proof of our main result Theorem 1.1. We assume that the measure space $(X, \Sigma, \mu)$ is $\sigma$-finite and that for any $f \in L^1$, $P^n f$ converges to a strongly compact set $\mathcal{F} \subset L^1$.

Since $\mu$ is $\sigma$-finite, there exists a density $f_0$ with $f_0(x) > 0$ for (almost) all $x$. Since $\{P^n f_0\}$ is precompact, the mean ergodic theorem [1] implies the limit

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f = g
$$

must exist and the limiting function $g$ satisfies $\|g\| = 1$ and $Pg = g$. Write $G = \text{supp} \, g$. Now, for every $A \in \Sigma$, define $\hat{\mu}(A) = \int_A g \, d\mu$, and let $\hat{L} = L^1(X, \Sigma, \hat{\mu})$. Denote the norm in $\hat{L}$ by $\|\cdot\|$. Define $\hat{P} : \hat{L} \to \hat{L}$ by

$$
\hat{P} h = P(hg)/g \quad \text{for } h \in \hat{L}.
$$

where $hg$ is the product $h(x)g(x)$. It is clear that $\hat{P}1 = 1$ and for $h \geq 0$, $\hat{P}h \geq 0$.

Moreover, for $h \geq 0$, we have

$$
|||\hat{P} h||| = \int_X \frac{1}{g} P(hg) \, d\hat{\mu} = \int_X P(hg) \, d\mu = \int_X hg \, d\mu = |||h|||.
$$

(6.1)
Therefore, \( \hat{P} \) is a Markov operator. To prove that \( \hat{P} \) is strongly constritive, define \( \mathcal{F} = \{ f/g : f \in \mathcal{F} \} \). Since

\[
\left\| \frac{f}{g} \right\| = \int_X \frac{f}{g} \, d\mu = \int_G f \, d\mu = \|f\|_G \leq \|f\|,
\]

\( \mathcal{F} \subset \mathcal{L} \). If \( \{ f_n/g \} \subset \mathcal{F} \), then there is a subsequence \( \{ f_{n_k} \} \subset \mathcal{F} \) such that \( f_{n_k} \to f \).

By (6.2) \( \left\| f_{n_k}/g - f/g \right\| \leq \left\| f_{n_k} - f \right\| \). Thus, \( \mathcal{F} \) is strongly compact. Further, for \( \left\| h \right\| = 1 \) and \( h \geq 0 \) we have \( \|hg\| = \|h\| = 1 \), and since \( P \) is strongly constritive there exists a sequence \( \{ f_n \} \subset \mathcal{F} \) such that \( \|P^n(hg) - f_n\| \to 0 \). Consequently,

\[
\left\| \frac{P^n}{g} - \frac{f_n}{g} \right\| = \int_G \left| \frac{1}{g} P^n(hg) - \frac{f_n}{g} \right| g \, d\mu \leq \|P^n(hg) - f_n\| \to 0
\]

which shows that \( \hat{P}^n h \to \hat{f} \). Thus all the conditions of Theorem 5.3 are satisfied by the operator \( \hat{P} \) acting on \( \mathcal{L} \). By virtue of this theorem, we may write

\[
P^n(hg) = \sum_{i=1}^r \lambda_i(h) g_{\alpha(i)}(t) + \varepsilon_n(h)
\]

where \( \left\| \varepsilon_n(h) \right\| \to 0 \) and \( \hat{g}_i \) are densities in \( \mathcal{L} \). From (6.3) we have, by letting \( g_i = g \cdot \hat{g}_i \),

\[
P^n(hg) = \sum_{i=1}^r \lambda_i(h) g_{\alpha(i)}(t) + g \varepsilon_n(h)
\]

with \( \|g \varepsilon_n(h)\| = \|\varepsilon_n(h)\| \to 0 \). From the identity \( \hat{P} g_i = \hat{g}_{\alpha(i)} \) and the definition of \( \hat{P} \), it follows that \( Pg_i = g_{\alpha(i)} \) and from (6.2), \( g_i \) is a density.

To finish the proof of Theorem 1.1, we first give the following lemma concerning the \( f_0 \) in the construction of \( g \).

**Lemma 6.1.** For any \( \varepsilon > 0 \), there exists a positive integer \( m \) such that

\[
\int_{X \setminus G} P^m f_0 \, d\mu \leq \varepsilon.
\]

**Proof.** Suppose that is not the case. Then, there is an \( \varepsilon > 0 \) such that

\[
\int_{X \setminus G} P^m f_0 \, d\mu > \varepsilon \quad \text{for all } m > 0.
\]

It follows that

\[
\int_X \left| \frac{1}{n} \sum_{m=0}^{n-1} P^m f_0 - g \right| \, d\mu \geq \int_{X \setminus G} \left| \frac{1}{n} \sum_{m=0}^{n-1} P^m f_0 - g \right| \, d\mu
\]

\[
= \int_{X \setminus G} \left| \frac{1}{n} \sum_{m=0}^{n-1} P^m f_0 \, d\mu \right| > \varepsilon \quad \text{for all } n > 0,
\]

which contradicts the definition of \( g \). \( \Box \)

Now we return to the proof of Theorem 1.1. Let \( f \in L^1 \) and \( \varepsilon > 0 \) be given. Since \( f_0 > 0 \), there exists a constant \( c > 0 \) and \( q_1 \in L^1 \) with \( \|q_1\| < \varepsilon/4 \) such that
\[ |f| \leq cf_0 + q_1. \] Choose \( m > 0 \) in Lemma 6.1 according to \( \varepsilon/4c \). Then,
\[
\int_{X \setminus G} |P^m f| \, d\mu \leq \int_{X \setminus G} P^m |f| \, d\mu \leq c \int_{X \setminus G} P^m f_0 \, d\mu + \int_{X \setminus G} P^m g_1 \, d\mu
\]
\[
< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}.
\]
On the other hand, there exists a constant \( c_1 > 0 \) and \( q_2 \in L^1 \) with \( \|q_2\| < \varepsilon/4 \) and \( \text{supp} \, q_2 \subset G \) such that
\[ 1_G \cdot |P^m f| \leq c_1 g + q_2. \]
Write \( h = (1_G \cdot P^m f - q_2)/g \). Then \( 1_G \cdot P^m f = hg + q_2 \), with \( \|h\| = \|hg\| \leq c_2 = c_1 + \varepsilon/2 \). Now,
\[ P^m f = 1_{X \setminus G} \cdot P^m f + 1_G \cdot P^m f \]
\[ = 1_{X \setminus G} \cdot P^m f + hg + q_2 = hg + q_3, \]
where
\[ \|q_3\| \leq \|1_{X \setminus G} \cdot P^m f\| + \|q_2\| < \varepsilon/2 + \varepsilon/4 = 3\varepsilon/4. \]
Since \( h \in L_1 \), we have, from (6.4),
\[ P^n (hg) = \sum_{i=1}^r \lambda_i(h)g_{\alpha^n(i)} + g\varepsilon_n(h) \]
with \( \|g\varepsilon_n(h)\| \to 0 \). Choose \( n \) large enough to have \( \|g\varepsilon_n(h)\| < \varepsilon/4 \). Then,
\[ P^{n+m} f = P^n (hg) + P^n q_3 \]
\[ = \sum_{i=1}^r \lambda_i(h)g_{\alpha^n(i)} + g\varepsilon_n(h) + P^n q_3. \]
Hence,
\[ \left\| P^{n+m} f - \sum_{i=1}^r \lambda_i(h)g_{\alpha^n(i)} \right\| < \frac{3}{4} \varepsilon + \frac{1}{4} \varepsilon = \varepsilon, \]
and the condition of Lemma 5.2 is satisfied with \( n_0 = n + m \) and \( C_i = \lambda_{\alpha^{-n}(i)}(h) \) for \( 1 \leq i \leq r \). \( \square \)

**References**

10. A. Lasota and J. A. Yorke, *On the existence of invariant measures for piecewise monotonic trans-
11. _, *Exact dynamical systems and the Frobenius-Perron operator*, Trans. Amer. Math. Soc. 273
261-268.

**Institute of Mathematics, Silesian University, 40-007 Katowice, Poland**

**Department of Mathematics, Michigan State University, East Lansing, Michigan 48824**

**Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742**