ASYMPTOTIC PERIODICITY FOR THE ITERATES OF MARKOV OPERATORS

Réjean Lévesque

A Thesis

in

the Department

of

Mathematics

Presented in Partial Fulfillement of the Requirements
for the degree of Master of Mathematics at
Concordia University
Montréal, Québec, Canada

August, 1987

ISBN 0-315-44870-9

ACKNOWLEDGEMENTS

I take this opportunity to acknowledge Professor Abraham Boyarsky's quality as a teacher and thesis adviser. His patience, guidance and encouragements made my studies at Concordia a rewarding experience.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

In recent years, the study of the probabilistic properties of dynamical systems has received much attention. For example if one deals with a map τ from the unit interval into itself, one studies the asymptotic behavior of the associated Frobenius-Perron operator $P_\tau : L_1(X, \Sigma, \mu) \mapsto L_1(X, \Sigma, \mu)$. It has been proved for large classes of such maps that asymptotically we can expect some regularity in their probabilistic behavior which is reflected in the spectral properties of the associated Frobenius - Perron operator.

In this thesis we shall obtain a spectral decomposition for a Markov operator under the hypothesis that its iterates of any density tend to a common strongly compact set. Such an operator is called constrictive [1]. We present of the proof [1] in chapter 2.

Recently Komornik [2] has obtain the spectral decomposition proved in [1] under the assumption that the iterates of any density converge to a weakly compact subset of $L_1(X, \Sigma, \mu)$.

In this chapter we introduce the underlying motivations for studying dynamical systems from the probabilistic point of view. The last section is devoted to the statement of the main theorem and we also give an example of a class of constrictive Frobenius-Perron operators which are induced by maps of the interval.

Chapter 2 deals with the proof of the main theorem. In chapter 3 we prove some results about random maps. A random map give rise to a convex combination of Markov operators. The asymptotic behavior of such an operator may be very complicated. If we assume that at least one of the operators in the combination is constrictive we obtain an upper bound for the number of elements in the spectral representation of the combination. It is not known in general if convex combinations of constrictive Markov operators have a fixed point. We present two examples of convex combination of constrictive Markov operators which are themselves constrictive. The results obtained in this chapter generalize some earlier results obtained by Pelikan [3].

1.2 Markov Operators.

Let (X, Σ, μ) be a measure space and let $L_1(X, \Sigma, \mu)$ the space of integrable functions on (X, Σ, μ).

1.2.1 Definition

A linear operator P, $P : L_1(X, \Sigma, \mu) \mapsto L_1(X, \Sigma, \mu)$ is said to be **Markov** *if:*

1) $Pf \geq 0$ *for* $f \geq 0$, $f \in L_1(X, \Sigma, \mu)$ *(Positivity)*

2) $\|Pf\| = \|f\|$, $f \geq 0$, $f \in L_1(X, \Sigma, \mu)$ *(norm preserving)*

Using 1) and 2) we will prove some properties of Markov operators which are going to be useful in the sequel. Let $(f)^+ = \max(0, f(x))$, and let $(f)^- = \max(0, -f(x))$.

1.2.2 Proposition

Let P be a Markov operator on $L_1(X, \Sigma, \mu)$, then

M1) $Pf \geq Pg$, *whenever* $f \geq g \in L_1(X, \Sigma, \mu)$

M2) $(Pf)^+ \leq Pf^+$

M3) $(Pf)^- \leq Pf^-$

M4) $|Pf| \leq P|f|$

M5) $\|Pf\| \leq \|f\|$

Proof:

M1) Since $f \geq g$, then $f - g \geq 0$ and $P(f - g) \geq 0$ by positivity of P.

M2) From the definition of f^+ and f^-, it follows that $f = f^+ - f^-$. Hence $(Pf)^+ = (Pf^+ - Pf^-)^+ = \max(0, Pf^+ - Pf^-) \leq \max(0, Pf^+) = Pf^+$

M3) The proof is analogous to M2).

M4) Since $|Pf| = (Pf^+) + (Pf^-)$, and, from M2), M3) $(Pf)^+ \leq Pf^+$, $(Pf)^- \leq Pf^-$ we get that $|Pf| \leq Pf^+ + Pf^- = P(f^+ + f^-) = P|f|$.

M5) We have that $|Pf| \leq P|f|$ and $\int_X P|f|d\mu = \int_X |f|d\mu$, since $|f| \geq 0$, so that

$$\|Pf\| = \int_X |Pf|d\mu \leq \int_X P|f|d\mu = \int_X |f|d\mu = \|f\|.$$

It follows directely from M5) that $\|P\| \leq 1$, and hence that any Markov operator is a contraction.

1.2.3 Definition

Support of $f \equiv \text{supp} f = \text{cl}\{x : f(x) \neq 0\}$

1.2.4 Proposition

$\|Pf\| = \|f\|$ if and only if Pf^+ and Pf^- have disjoint supports.

Proof:

We have $|Pf^+ - Pf^-| \leq |Pf^+| + |Pf^-|$. Integrating over X .

$$\int_X |Pf^+(x) - Pf^-(x)|d\mu \leq \int_X |Pf^+(x)|d\mu + \int_X |Pf^-(x)|d\mu$$

Equality occurs if and only if there is no set $A \in \Sigma$, $\mu(A) > 0$, such that $Pf^+(x) > 0$ and $Pf^-(x) > 0$ for $x \in A$; that is,$Pf^+(x)$ and $Pf^-(x)$ have disjoint support. Since $f = f^+ - f^-$, the left hand integral is simply $\|Pf\|$. The right hand side is $\|Pf^+\| + \|Pf^-\| = \|f^+\| + \|f^-\| = \|f\|$ and this completes the proof of the proposition.

We introduce now the idea of a fixed point for a Markov operator, which is fundamental in studying dynamical systems from the probabilistic point of view.

1.2.4 Definition

If P is a Markov operator and, for some $f \in L_1(X, \Sigma, \mu)$, we have $Pf = f$ then f is called **a fixed point** *for P.*

1.2.5 Proposition

If $Pf = f$, then $Pf^+ = f^+$ and $Pf^- = f^-$.

Proof:

From $Pf = f$ we have $f^+ = (Pf)^+ \leq Pf^+$ and $f^- = (Pf)^- \leq Pf^-$ so

$$\begin{aligned}
0 &\leq \int_X [Pf^+ - f^+]d\mu + \int_X [Pf^- - f^-]d\mu \\
&= \int_X [Pf^+ + Pf^-]d\mu - \int_X [f^+ + f^-]d\mu \\
&= \int_X P|f|d\mu - \int_X |f|d\mu \\
&= \|P|f|\| - \||f|\|.
\end{aligned}$$

Since P is a contraction, the last equality is less than or equal to 0. This is only possible if $Pf^+ = f^+$ and $Pf^- = f^-$.

1.2.6 Definition

Let (X, Σ, μ) be a measure space and let

$$D(X, \Sigma, \mu) = \{f \in L_1(X, \Sigma, \mu) \mid f \geq 0, \|f\| = 1\}.$$

Any $f \in D(X, \Sigma, \mu)$ is called a **density**.

Consider now the set function

$$\mu_f(A) = \int_A f d\mu, \qquad f \in D(X, \Sigma, \mu), \quad A \in \Sigma.$$

It is easily checked that μ_f is a measure, $\mu_f(X) = 1$, and that $\mu_f(A) = 0$ whenever $\mu(A) = 0$. The measure μ_f is then said to be **absolutely continuous with respect to** μ and f is the density associated with μ_f.

1.2.7 Definition

For any $f \in D(X, \Sigma, \mu)$ such that $Pf = f$, the measure μ_f is called an **absolutely continuous measure invariant under** P

1.3 An Example of Markov Operator: Frobenius-Perron Operators

We begin with two definitions.

1.3.1 Definition

Let (X, Σ, μ) be a measure space. A transformation $\tau : X \mapsto X$ is **measurable** *if $\tau^{-1}(A) \in \Sigma$ for all $A \in \Sigma$.*

1.3.2 Definition

A measurable transformation $\tau : X \mapsto X$ on a measure space (X, Σ, μ) is **non-singular** *if $\mu(\tau^{-1}(A)) = 0$ for all $A \in \Sigma$ such that $\mu(A) = 0$.*

Non-singular tranformations represent an important class of transformations on a measure space. More precisely, the transformations induce Markov operators on $L_1(X, \Sigma, \mu)$ in the following way:

1) Let $f \in L_1(X, \Sigma, \mu)$ and $f \geq 0$. Consider

$$\int_{\tau^{-1}(A)} f(x) d\mu, \qquad A \in \Sigma.$$

Since $\tau^{-1}(\bigcup_i A_i) = \bigcup_i \tau^{-1}(A_i)$, it follows that this integral defines a finite measure. Thus,by the Radon-Nikodym theorem, there is a unique element in $L_1(X, \Sigma, \mu)$, which we denote by $P_\tau f$, such that

$$\int_A P_\tau f d\mu = \int_{\tau^{-1}(A)} f(x) d\mu, \qquad A \in \Sigma.$$

2) If $f \in L_1(X, \Sigma, \mu)$ is not non-negative, write $f = f^+ - f^-$ and define $P_\tau f = P_\tau f^+ - P_\tau f^-$. From this, we have:

$$\int_A P_\tau f(x) d\mu = \int_A f^+(x) d\mu - \int_A f^-(x) d\mu$$

that is

$$\int_A P_\tau f d\mu = \int_{\tau^{-1}(A)} f(x)d\mu, \qquad A \in \Sigma.$$

1.3.3 Definition

Let (X, Σ, μ) be a measure space. If $\tau : X \mapsto X$ is a non-singular transformation, the unique operator P_τ induced by τ is called the **Frobenius-Perron** *operator associated with τ.*

It is straightforward to show from the definition of P_τ that a Frobenius-Perron operator has the following properties:

FP1) $P_\tau(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1(P_\tau f_1) + \lambda_2(P_\tau f_2)$ for all $f_1, f_2 \in L_1(X, \Sigma, \mu)$, $\lambda_1, \lambda_2 \in \mathbf{R}$. (linearity)

FP2) $P_\tau f \geq 0$ if $f \geq 0$. (positivity)

FP3) $\int_X P_\tau f d\mu = \int_X f d\mu$. (preserves integrals)

FP4) If

$$\tau^n = \overbrace{\tau \circ \tau \circ \ldots \circ \tau}^{n \text{ times}}$$

and P_{τ^n} is the Frobenius-Perron corresponding to τ^n, then $P_{\tau^n} = P_\tau^n$ where P_τ is the Frobenius-Perron operator corresponding to τ.

One can see directly from FP1)-FP3) that a Frobenius-Perron operator is a Markov operator and that it discribes the evolution of any $f \in L_1(X, \Sigma, \mu)$ under the transformation τ.

In some special cases we can obtain an explicit form for P_τ.

If $X = [a,b]$, an interval on the real line with the usual Lebesgue measure and $A = [a,x]$, then

$$\int_a^x P_\tau f(s)d\mu = \int_{\tau^{-1}[a,x]} f(s)d\mu$$

and by differentiating

$$P_\tau f(x) = \frac{d}{dx}\int_{\tau^{-1}[a,x]} f(s)d\mu.$$

If the transformation τ is differentiable and invertible, then an explicit form of $P_\tau f$ is available. In this case τ must monotone. Suppose it is increasing and that τ^{-1} has continuous derivative. Then $\tau^{-1}[a,x] = [\tau^{-1}(a), \tau^{-1}(x)]$ and then

$$P_\tau f(x) = \frac{d}{dx}\int_{\tau^{-1}(a)}^{\tau^{-1}(x)} f(s)d\mu = f(\tau^{-1}(x))\frac{d}{dx}(\tau^{-1}(x))$$

by Leibnitz's Rule.

If τ is decreasing, then the sign in the last expression is reversed. Thus in the general one dimensional case, for τ differentiable and invertible with $\frac{d}{dx}(\tau^{-1}(x))$ continuous,

$$P_\tau f(x) = f(\tau^{-1}(x))\left|\frac{d}{dx}(\tau^{-1}(x))\right|.$$

Frobenius-Perron operator: An example.

We now derive the Frobenius-Perron operator for piecewise monotone $\mathbf{C}^1$ function in $[0,1]$.

1.3.4 Theorem

Let $\phi_i \in \mathbf{C}^1[b_{i-1}, b_i]$ and monotone where $0 = b_0 < b_1 < \ldots < b_q = 1$. Assume also for the sake of convenience that each ϕ_i can be extended as a monotone $\mathbf{C}^1$ function on $[0,1]$. Then for

$$\phi = \sum_{i=1}^{q} \phi_i \chi_{B_i}$$

where $B_i = [b_{i-1}, b_i]$ we have

$$P_\phi f(x) = \sum_{i=1}^{q} f(\psi_i(x))\sigma_i(x)\chi_{J_i}(x)$$

where $\psi_i = \phi_i^{-1}$, $\sigma_i = |\psi_i|$, $J_i = \phi_i(B_i)$.

Proof:

Set $A_i(x) = \phi_i^{-1}([0,x]) \cap B_i = \psi_i([0,x]) \cap B_i$. Then

$$\int_{A_i(x)} f(s)ds = \pm \int_{\psi_i(0)}^{\psi_i(x)} f(s)\chi_{B_i}(s)ds.$$

we want $\int_{A_i(x)} f \geq 0$ when $f \geq 0$. Since ϕ_i is monotone, ψ is monotone and ϕ_i and ψ_i are either both increasing or both decreasing. Therefore

$$\frac{\psi_i'(x)}{|\psi_i'(x)|} = \frac{\psi_i'(y)}{|\psi_i'(y)|}$$

for all $x, y \in [0,1]$. We use this to set the sign. Thus

$$\int_{A_i(x)} f(s)ds = \frac{\psi_i'(x)}{|\psi_i'(x)|} \int_{\psi_i(0)}^{\psi_i(x)} f(s)\chi_{B_i}(s)ds$$

$$\frac{d}{dx}\int_{A_i(x)} f(s)ds = \frac{\psi_i'(x)}{|\psi_i'(x)|}\frac{d}{dx}\int_{\psi_i(0)}^{\psi_i(x)} f(s)\chi_{B_i}(s)ds$$
$$= \frac{\psi_i'(x)}{|\psi_i'(x)|} f(\psi_i(x))\chi_{B_i}(\psi_i(x))\psi_i'(x)$$

by Leibnitz's Rule

$$= \frac{|\psi_i'(x)|^2}{|\psi_i'(x)|} f(\psi_i(x))\chi_{B_i}(\psi_i(x))$$
$$= f(\psi_i(x))\sigma_i(x)\chi_{B_i}(\psi_i(x)).$$

Note that

$$\chi_{B_i}(\psi_i(x)) = 1 \iff \psi_i(x) \in B_i$$
$$\iff x \in \phi_i(B_i) = J_i$$
$$\iff \chi_{J_i} = 1.$$

Therefore $\chi_{J_i}(x) = \chi_{B_i}(\psi_i(x))$ and we get

$$\frac{d}{dx}\int_{A_i(x)} f(s)ds = f(\psi_i(x))\sigma_i(x)\chi_{J_i}(x).$$

For $\phi = \sum_{i=0}^{q} \phi_i \chi_{B_i}$, $\phi^{-1}([0,x]) = \bigcup_{i=0}^{q} A_i(x)$, where the A_i's are disjoint since the B_i's are disjoint. Thus

$$P_\phi f(x) = \frac{d}{dx}\int_{\phi^{-1}([0,x])} f(s)ds$$
$$= \frac{d}{dx}\sum_{i=0}^{q}\int_{A_i(x)} f(s)ds.$$

Therefore

$$P_\phi f(x) = \sum_{i=1}^{q} f(\psi_i(x))\sigma_i(x)\chi_{J_i}(x)$$

1.4 Invariant Measure, Ergodicity, Mixing and Exactness

Let $\tau : X \mapsto X$ be a non singular measurable transformation on a measure space X. Let $x \in X$ be a starting point and let the corresponding orbit be defined by the sequence of iterations $\{x_n = \tau^n(x)\}_{n=1}^{\infty}$.

The main idea is not to study individual orbits corresponding to a single point but to rather study the evolution of distributions or densities of infinitely many initial conditions, by means of the iterates of the Frobenius-Perron operator induced by τ. If P_τ is the Frobenius-Perron associated with τ the effect of P_τ goes "forward" in time, that is, if f is the probability density on X at time n then $P_\tau f$ is the density on X at time $n+1$. Given these observations an important question is: Does there exists an $f^* \in D$ such that the action of τ on X does not change on f^* ? That is, does there exists an $f^* \in D$ such that $P_\tau f^* = f^*$? The existence of such an f^* would tell us something about the probabilistic behaviour of τ.

1.4.1 Definition

A measure μ is said to be invariant under τ if $\mu(\tau^{-1}(A)) = \mu(A)$ for all $A \in \Sigma$. If $P^n f \to f^*$ in $L_1(X, \Sigma, \mu)$ sense, then $Pf^* = f^*$ and

$$\int_A f^* d\mu = \int_{\tau^{-1}(A)} f^* d\mu \qquad \text{for all } A \in \Sigma$$

Thus $\nu(A) = \int_A f^*(x) d\mu$ is an absolutely continuous measure invariant under τ.

There are simple mappings $\tau : X \mapsto X$ having an absolutely invariant measure. For example if X is some closed interval of the real line, μ the Borel measure in X and τ such that $\tau(X) = X$, then there are infinitely many such measures.

However, the existence of an absolutely continuous invariant measure in itself does not imply any stochastic or turbulent like behaviour of the dynamical system defined by τ. Let τ have a unique absolutely continuous invariant measure. Then erratic behaviour can be investigated by means of the different modes of convergence of $\{P^n f\}$ to the density of the invariant measure. We now describe these modes of convergence.

1.4.2 Definition

A transformation τ is said to be **ergodic** *if there exists no non-trivial set of X which is invariant under τ, more precisely, τ is ergodic if for all $A \in \Sigma$ for which $\tau^{-1}(A) = A$, $\mu(A) = 0$ or $\mu(X \setminus A) = 0$.*

As an example let us consider the rotation F on the unit circle S^1, where $F(x) = x + \theta$ and $\theta \in [0, 2\pi]$ is constant. Obviously the measure induced by the arc length is invariant under F. But depending on whether θ is rational or irrational, F is not ergodic or ergodic respectively.

This example shows that ergodic transformations are not necessarily very irregular.

From the operator point of view, it can be shown that ergodicity of the transformation τ corresponds to the convergence of the average of the successive iterations of P,

$$\frac{1}{n}\left(\sum_{k=1}^{n} P^k f\right) \to f^* \qquad \text{for all } f \in \mathcal{D}.$$

where f^* is the density of the unique absolutely continuous measure invariant under τ. This is a consequence of Birkhoff's Ergodic Theorem. This type of convergence is also called Cesaro.

The next category of asymptotic behaviour is called mixing.

1.4.3 Definition

A transformation τ is called **mixing** *if*

$$\lim_{n\to\infty} \mu(A \cap \tau^{-n}(B)) = \mu(A)\mu(B), \qquad \text{for all } A, B \in \Sigma.$$

Roughly speaking, this condition means that if one starts with a set A of initial conditions, then after many iterations the fraction of solutions points lying in some (arbitrarily given measurable) set B equals the product of the measure of the sets A and B. Mixing is loosely called irregular or chaotic behaviour.

It can be shown that mixing for a transformation implies that the iterates of its associated operator P in $L_1(X, \Sigma, \mu)$ converges weakly to the density of the unique absolutely continuous invariant measure. This type of convergence necesserary implies the Cesaro convergence of P and therefore mixing transformation are ergodic.

The last category of asymptotic behaviour is called exactness.

1.4.4 Definition

A transformation is said to be **exact** *if*

$$\lim_{n\to\infty} \mu(\tau^n(A)) \to \mu(X) \qquad \text{for } A \in \Sigma, \mu(A) > 0.$$

The iterates of the operator P of an exact transformation converge strongly for all $f \in D$ to the unique f^* representing the unique absolutely continuous measure. It therefore follows that exactness implies mixing which in turns implies ergodicity.

In the next section we describe another possible type of asymptotic behaviour which does not necessarly fit in the preceding hierarchy of assymptotic behaviour, namely asymptotic periodicity.

1.5 Constrictive Markov Operator, Asymptotic Periodicity

Let (X, Σ, μ) be a measure space with non-negative σ-finite measure μ. As before, let D denote the set of all normalized densities on X.

1.5.1 Definition

We say an operator P is **strongly constrictive** *if there exists a strongly compact set $F \subset L_1(X, \Sigma, \mu)$ such*

$$\lim_{n\to\infty} d(P^n g, F) = 0 \qquad \text{for } g \in D,$$

where $d(g, F)$ denotes the distance between g and F, that is, $\inf \|g - f\|$ for $f \in F$.

The main result proved in this thesis is the following.

1.5.2 Theorem[1]

Let P a strongly constrictive Markov operator. Then there exists a finite sequence of densities $g_1, g_2, \ldots, g_r$ and a finite sequence of bounded linear functionals $\lambda_1, \lambda_2, \ldots, \lambda_r$ such that

$$\lim_{n\to\infty} \|P^n\Big(f - \sum_{i=0}^{n} \lambda_i(f) g_i\Big)\| = 0 \qquad \text{for } f \in L_1(X, \Sigma, \mu).$$

The densities $\{g_i\}$ have mutually disjoint supports and $P g_i = g_{\alpha(i)}$ where

$\{\alpha(1), \alpha(2), \ldots, \alpha(r)\}$ *is a permutation of the integers* $\{1, 2, \ldots, r\}$.

Since

$$P^n\Big(f - \sum_{i=0}^{n} \lambda_i(f) g_i\Big) = P^n f - P^n\Big(\sum_{i=0}^{n} \lambda_i(f) g_i\Big)$$
$$= P^n f - \Big(\sum_{i=0}^{n} \lambda_i(f) g_{\alpha^n(i)}\Big)$$

and $P^n\Big(f - \sum_{i=0}^{n} \lambda_i(f) g_i\Big)$ converges in norm to 0 as $n \to \infty$, we could rewrite:

$$P^n f = \sum_{i=0}^{n} \lambda_i(f) g_{\alpha^n(i)} + R_n f,$$

where α^n denotes the n^{th} iterate of the permutation α and $R_n f$ converges in norm to 0 as $n \to \infty$. Thus, every sequence$\{P^n f\}$ is asymptotically periodic, with period which does not exceed $r!$ since the permutations of r elements form a cyclic group the period of which does not exceed $r!$.

We now give an example of a constrictive operator. Consider a measurable transformation $S : [0, 1] \mapsto [0, 1]$, which satisfies the following conditions:

i) There is a partition $0 = a_0, a_1, \ldots, a_m = 1$ such that for each integer i the restriction S_i of S to the interval (a_{i-1}, a_i) is a $\mathbf{C}^2$ function.

ii) $\inf |S'(x)| > 1$, $(x \neq a_i)$ where $'$ denote the derivative.

iii) $\sup \dfrac{|S''(x)|}{|S'(x)|} < \infty$.

If T_i denote the inverse function S_i^{-1}, it is easy to write an explicit formula for P_S:

$$P_S f(x) = \sum_{i=1}^{m} T_i(x)[f(T_i(x))\chi_{C_i}(x)]$$

where χ_{C_i} denotes the characteristic function on the set $C_i = T_i(a_{i-1}, a_i)$. We can evaluate the total variation on $[0,1]$, denoted $\bigvee_0^1$ of the function $P_S^n f$ for large n [5]. Namely, if S satisfies i)-iii), then there exists a constant K independent of f such

$$\limsup_{n\to\infty} \bigvee_0^1 P_S^n f \leq K,$$

for every $f \in D[0,1]$ of bounded variation, The set

$$F = \{h \in D \mid \bigvee_0^1 h \leq K\}$$

is compact in $L_1(X, \Sigma, \mu)$, by Helley's Theorem and since the densities of boundeds variation are dense in D, $\{P_S^n f\}$ converges to F for any $f \in D$. We summarize these results in the following proposition.

1.5.3 Proposition

If $S : [0,1] \mapsto [0,1]$ satisfies conditions i)-iii), then the operator P_S is strongly constrictive. Consequently, for every $f \in L_1[0,1]$, the iterates $P_S^n f$ can be written in the form of the theorem 1.5.2 and the sequence $\{P_S^n f\}$ is assymptoticaly periodic.

CHAPTER 2

PROOF OF THE MAIN THEOREM

2.1 Proof of the Main Theorem: Case $\mu(X) < \infty$ and $P\chi_X = \chi_X$

We now proceed to the proof of the main theorem. We assume we have a strongly constrictive Markov operator P on (X, Σ, μ), with the additional assumption that $P\chi_X = \chi_X$ and $\mu(X) < \infty$. Later in this chapter we will delete the assumption $P\chi_X = \chi_X$ and $\mu(X) < \infty$ to prove the theorem in the general case.

2.1.1 Definition

A set $A \in \Sigma$ will be called a **nice** *set if $P\chi_A$ is a characteristic function for each positive integer n. In this case χ_A is called a* **nice** *function. In the following lemma, we show that the family of nice sets forms an algebra $\mathcal{A}$.*

2.1.2 Lemma

i) If A is a nice set, then $X \setminus A$ is a nice set.

ii) If A_1, A_2 are nice sets, then $A_1 \cup A_2$ is a nice set.

Proof:

i) From $\chi_{X\setminus A} = \chi_X - \chi_A$, it follows that

$$P^n\chi_X - P^n\chi_A = \chi_X - \chi_{B_n}$$

where B_n is the set which has $P^n\chi_A$ as its characteristic function. Hence $X \setminus A$ is a nice set.

ii) Since A_1, A_2 are nice sets, for fixed n, we have $P^n\chi_{A_1} = \chi_{B_1}, P^n\chi_{A_2} = \chi_{B_2}$ for somme sets B_1, B_2 contained in X. For $i = 1,2$ we have

$$\chi_{A_i} \leq \chi_{A_1 \cup A_2} \leq \chi_{A_1} + \chi_{A_2}.$$

Hence

$$\chi_{B_i} \leq P^n\chi_{A_1 \cup A_2} \leq \chi_{B_1} + \chi_{B_2}.$$

If $x \notin B_1 \cup B_2$, then $P^n\chi_{A_1 \cup A_2} = 0$.Since $P^n\chi_{A_1 \cup A_2} \leq \chi_X$, $x \in B_1 \cup B_2$ implies $P^n\chi_{A_1 \cup A_2} = 1$. So $A_1 \cup A_2$ is a nice set.

2.1.3 Corollary

Let A_1, A_2 be two nice sets which are disjoint and let $P\chi_{A_i} = \chi_{B_i}$, $i = 1,2$ Then $B_1 \cap B_2 = \emptyset$.

Proof:

Since $\chi_{A_1 \cup A_2} = \chi_{A_1} + \chi_{A_2}$, it follows that

$$P\chi_{A_1 \cup A_2} = P\chi_{A_1} + P\chi_{A_2} = \chi_{B_1} + \chi_{B_2}$$

On the other hand $P\chi_{A_1 \cup A_2} \leq P\chi_X = \chi_X$ so that $\chi_{B_1} + \chi_{B_2} \leq \chi_X$. This proves that $B_1 \cap B_2 = \emptyset$.

Now we use constrictiveness of P to show that in this case, the algebra of nice set is finite, that is, there are only finitely many nice sets.

2.1.4 Lemma

There exists a $\delta > 0$ such that $\mu(A) > \delta$ for all nice set $A \in \Sigma$.

Proof:

Choose $A \in \Sigma$ and define $f = \chi_A/\mu(A)$. Since P is constrictive, there exists $F \subset L_1(X, \Sigma, \mu)$ such that $P^n f \to F$ as $n \to \infty$ where $f \in D$. Thus, given $\varepsilon \in (0,1)$, there exists a sequence $\{g_n\} \subset F$ such that $\|P^n f - g_n\| < 1 - \varepsilon$. Since F is strongly compact, it is also weakly compact. Therefore, given $\varepsilon > 0$, there is a δ such that for all $B \in \Sigma$ such that $\mu(B) < \delta$ we have

$$\int_B g_n d\mu < \varepsilon.$$

Let $A_n = \operatorname{supp} P^n f$, then , since $P^n f = P^n \chi_A/\mu(A) = \chi_{A_n}$, we get that $\mu(A) = \mu(A_n)$ for all n. We claim $\mu(A) > \delta$. Assume the contrary. Then

$$\int_{A_n} g_n d\mu < \varepsilon$$

implies that

$$\begin{aligned} \|P^n f - g_n\| &\geq \int_{A_n} |P^n f - g_n| d\mu \\ &\geq \int_{A_n} P^n f d\mu - \int_{A_n} g_n d\mu \\ &\geq 1 - \varepsilon. \end{aligned}$$

which contradicts $\|P^n f - g_n\| < 1 - \varepsilon$.

A nonempty set B in an algebra $\mathcal{B}$ of sets is called an atom if the only subsets of positive measure 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 2.1.4, the number of atoms in $\mathcal{A}$ must finite since $\mu(X) < \infty$. Let $\{A_1, A_2, \ldots, A_r\}$ be the set of atoms of $\mathcal{A}$ and write $\chi_i = \chi_{A_i}$ for $i = 1, 2, \ldots, r$.

2.1.5 Lemma

There exists a permutation α of $\{1, 2, \ldots, r\}$ such that $P\chi_i = \chi_{\alpha(i)}$.

Proof:

It follows easily from the definition of nice functions that $P\chi_i$ is a nice function for each $i = 1, 2, \ldots, r$. Let $\chi_{B_i} = P\chi_i$. By Corollary 2.1.3 χ_{B_i} and χ_{B_j} have disjoint supports for $i \neq j$ and $B_1, B_2, \ldots, B_r$ is a collection of mutually disjoint elements of $\mathcal{A}$. Now each B_i is an atom. Suppose this is not the case, then for at least one i, there is a set $B_i' \subset B_i$, such that $\mu(B_i') > 0$. Since $\mu(A_i) = \mu(B_i)$ and $\mu(X) = \mu(\bigcup_{i=1}^r A_i) = \mu(\bigcup_{i=1}^r B_i)$, B_i must intersects an A_i on a set of positive measure. This contradicts the fact that the A_i's are atoms. Then $B_1, B_2, \ldots, B_r$ is a permutation of the $A_1, A_2, \ldots, A_r$.

For $f \in L_1(X, \Sigma, \mu)$, $\{P^n f\}$ is precompact since P is strongly constrictive. Let $\Omega(f)$ be the set of limit points of $\{P^n f\}$ and $Q \equiv \bigcup_{f \in L_1(X,\Sigma,\mu)} \Omega(f)$. We claim that all nice functions are in Q.First, observe that all possible unions of the atoms in $\mathcal{A}$ is also in $\mathcal{A}$ since it is an algebra. If a set A is in $\mathcal{A}$ then it can be written as a union of atoms of $\mathcal{A}$ since $\mu(A) \geq \delta > 0$ from Lemma 2.1.4. Therefore, the set of atoms $\{A_1, A_2, \ldots, A_r\}$ generates the algebra $\mathcal{A}$, and therefore $\mathcal{A}$ is a finite algebra. Now for any $A \in \mathcal{A}$, $A = \bigcup_{k=1}^j A_k$ for some $j \in \{1, 2, \ldots, r\}$, and $\chi_A = \chi_{A_1} + \chi_{A_2} + \cdots + \chi_{A_j}$, since the A_i's are disjoint. Therefore the sequence $P^n \chi_A = \sum_{k=1}^j P^n \chi_{A_k}$ contains a subsequence convergent to χ_A because P permutes $\{\chi_1, \chi_2, \ldots, \chi_r\}$. Then $\chi_A \in \Omega(\chi_A) \subset Q$ for any nice set A, which prove the claim.

We now state:

2.1.6 Theorem

Q is a finite dimensional linear space with basis $\{\chi_1, \chi_2, \ldots, \chi_r\}$.

Before proving it we need to prove a few lemmas.

2.1.7 Lemma

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

Proof:

If $f \in Q$, then there exists a $g \in L_1(X, \Sigma, \mu)$ and a sequence $\{P^{n_i} g\}$ of $\{P^n g\}$ such that $\{P^{n_i} g\} \to f$ as $i \to \infty$. That is given $\varepsilon > 0$, there is a positive integer N such that $\|P^{n_k} g - f\| < \frac{\varepsilon}{2}$ for $k \geq N$. Then $\|P^{n_k} g - P^{n_{k+1} - n_k} f\| < \frac{\varepsilon}{2}$. But $k + 1 > N$ so that $\|P^{n_{k+1}} g - f\| < \frac{\varepsilon}{2}$, that is

$$\|P^{n_{k+1} - n_k} f - f\| < \|P^{n_{k+1}} g - P^{n_{k+1} - n_k} f\| + \|P^{n_{k+1}} g - f\|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

then there is a subsequence of $\{P^n f\}$ which converges to f and so $f \in \Omega(f)$.

2.1.8 Lemma

If $f \in Q$, then $\|Pf\| = \|f\|$.

Proof:

First notice that $\|f\| \geq \|Pf\| \geq \ldots \|P^n f\| \geq \ldots$. From the ~~preceding~~ lemma, there is a subsequence $\{P^{n_k} f\} \to f$, that is $\|P^{n_k}\| \to \|f\|$. Hence $\|P^{n_k} f\|$ is a constant, which implies $\|Pf\| = \|f\|$.

2.1.9 Lemma

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 prove that $\operatorname{supp} Pf_1 \subset \operatorname{supp} Pf_2$. The reverse inclusion is proved in exactly the same way. Let $\operatorname{supp} Pf_2 = B_2$. Then

$$Pf_1 = \chi_{B_2} Pf_1 + \chi_{X \setminus B_2} Pf_1 .$$

For $c > 0$, define $f_c = \min(cf_2, f_1)$. Then f_c has the proprieties:

1) $f_c \leq f_1$
2) $\lim_{c\to\infty} f_c(x) = f_1(x)$
3) $\operatorname{supp} Pf_c \subset \operatorname{supp} Pf_2 = B_2$

From 1) and 2) and the Lebesgue Dominated Convergence Theorem, we have

$$\|Pf_1 - Pf_c\| = \int_X P(f_1 - f_c)d\mu = \int_X (f_1 - f_c)d\mu \to 0$$

as $c \to \infty$. From 1) and 3),

$$\|Pf_1 - Pf_c\| = \int_{X\setminus B_2} (Pf_1 - Pf_c)d\mu = \int_{X\setminus B_2} Pf_1 d\mu = \|Pf_1\chi_{X\setminus B_2}\|.$$

So that $Pf_1\chi_{X\setminus B_2} = 0$. Thus $\operatorname{supp} Pf_1 \subset B_2 = \operatorname{supp} Pf_2$.

2.1.10 Proposition

If $f \in Q$, then for each $v \geq 0$, $f^{-1}(-\infty, v)$ is a nice set.

Proof:

We first assume that $\mu(f^{-1}(v)) = 0$. Write $h_1 = \chi_{f^{-1}(-\infty,v)}$ and $h_2 = \chi_{f^{-1}[v,\infty)}$. Then $h_1 + h_2 = \chi_X$. Letting $g = f - v$, we see that $g \in Q$, and

$$\operatorname{supp} g^+ = \{x : f(x) > v\} = \operatorname{supp} h_2$$

$$\operatorname{supp} g^- = \{x : f(x) < v\} = \operatorname{supp} h_1$$

Now $\operatorname{supp} g^+$ and $\operatorname{supp} g^-$ are disjoint and by combining Lemmas 2.1.8, 2.1.9 and Proposition 1.2.4 we get that

$$\operatorname{supp} P^n h_1 \cap \operatorname{supp} P^n h_2 = \emptyset$$

Now $P^n(h_1+h_2) = P^n\chi_X = \chi_X = P^n h_1 + P^n h_2$, therefore, $P^n h_1$ is a characteristic function for each $n \geq 0$. Suppose now, $\mu(f^{-1}(v)) > 0$. Then, since $\mu(x) < \infty$ and

$f^{-1}(c)$ are disjoint for different $c > 0$, the set of c's with $\mu(f^{-1}(c)) > 0$ is at most countable. thus its complement is dense. Let v_i be an increasing sequence with $v_i \to v$ and $\mu(f^{-1}(v_i)) = 0$. Then since $f^{-1}(-\infty, v) = \bigcup_i f^{-1}(-\infty, v_i)$ and the nice set form an algebra, $f^{-1}(-\infty, v)$ is also a nice set.

Proof of Theorem 2.1.6

It suffices to prove that every $f \in Q$ can be written as a linear combination of the χ_i's. By Proposition 2.1.10, 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 $v_1, v_2, \ldots, v_k$, for which $f^{-1}(-\infty, v_1), f^{-1}(-\infty, v_2), \ldots, f^{-1}(-\infty, v_k)$ are different sets. Thus f has the form for the appropriate choice of real number β_i,

$$f = \sum_{i=1}^{k-1} \beta_i (\chi_{f^{-1}(-\infty, v_{i+1})} - \chi_{f^{-1}(-\infty, v_i)}).$$

Since $\chi_{f^{-1}(-\infty, v_{i+1})} - \chi_{f^{-1}(-\infty, v_i)}$ are nice functions, f is a linear combination of the χ_i's.

Let $P : L_1(X, \Sigma, \mu) \mapsto L_1(X, \Sigma, \mu)$ be a Markov operator and let a sequence of densities $\{g_i\}$ $(i = 1, 2, \ldots, r)$ with mutually disjoint supports be given. We assume that P permutes the densities g_i, i.e. $Pg_i = g_{\alpha(i)}$, where α is a permutation of the integers $1, 2, \ldots, r$.

2.1.11 Definition

We say that the $\{g_i\}$ induce an **asymptotic decomposition** *of P if there is a sequence of linear functionals $\{\lambda_i\}_{i=1}^{r}$ on $L_1(X, \Sigma, \mu)$ such that*

$$\lim_{n \to \infty} \|P^n(f - \sum_{i=i}^{r} \lambda_i(f) g_i)\| \to 0 \qquad \text{for } f \in L_1(X, \Sigma, \mu).$$

Since the g_i's are linearly independant, the functionals λ_i are nescessarly bounded.

2.1.12 Lemma

The densities $\{g_i\}$ *induce an asymptotic decomposition if and only if for every* $f \in L_1(X, \Sigma, \mu)$ *and* $\varepsilon > 0$, *there are constant* $c_1, c_2, \ldots, c_r$ *and a positive integer* n_0 *such that*

$$\|P^n f - \sum_{i=1}^{r} c_i g_i\| \le \varepsilon, \qquad n \ge n_0.$$

Proof:

The "only if" part is obvious. To prove the "if" part, let n_k and c_i^k, $i = 1, 2, \ldots, r$, be chosen such that, with indices rearranged if nescessary,

$$\|P^{n_k} f - \sum_{i=1}^{r} c_i^k g_{\alpha^{n_k}(i)}\| \to 0. \tag{*}$$

Now $\{c_i^k\}$ is a bounded sequence for each i. By choosing a subsequence if nescessary, we suppose that, for each i, c_i^k converges to a constant $\lambda_i(f)$. Write

$$E_n = \|P^n(f - \sum_{i=1}^{r} \lambda_i(f) g_i)\|.$$

It follows from (*) that $E_{n_k} \to 0$ as $n \to \infty$. Furthermore, E_n is a decreasing sequence since P being a Markov operator, $\|P^n h\|$ is nonincreasing for each $h \in L_1(X, \Sigma, \mu)$. Hence $E_n \to 0$. It remains to show that the λ_i's are linear. Observe that $\lambda_i(f)$, $i = 1, 2, \ldots, r$ are uniquely defined since the g_i are linearly independent. Moreover, from

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

$P^n(f_2 - \sum_{i=1}^r \lambda_i(f_2) g_i) \to 0$ imply that

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

It follows that $\lambda_i(f_1) + \lambda_i(f_2) = \lambda_i(f_1 + f_2)$ for each i.

2.1.13 Proposition

Let P be a strongly constrictive Markov operator on a measure space (X, Σ, μ). Assume $P\chi_X = \chi_X$ and $\mu(x) < \infty$. Then, there is a sequence of densities $\{g_i\}$, $i = 1, 2, \ldots, r$ with mutually disjoint supports, which gives an asymptotic decomposition of P.

Proof:

Let $\{A_i\}$ and χ_i be as in Lemma 2.1.5. Define $g_i = \chi_i/\mu(A_i)$. By virtue of Theorem 2.1.6, all elements of Q are of the form $\sum_{i=1}^{r} c_i g_i$. Thus, for $f \in L_1(X, \Sigma, \mu)$, there is a subsequence of $\{P^n f\}$ which converges to a function of the form $\sum_{i=1}^{r} c_i g_i$. Then, by Lemma 2.1.12 we have an asymptotic decomposition of f.

2.2 Proof of the Main Theorem: the General Case

We are now going to prove the main theorem in the general case, that is, by deleting the conditions $\mu(X) < \infty$ and $P\chi_X = \chi_X$. We assume only that X is a σ-finite measure space and that P is constrictive. Since μ is σ-finite, there exists a density f_0 with $f_0(x)$ a.e. Since $\{P^n f_0\}$ is precompact, P being constrictive, the mean ergodic Theorem implies that the limit $\frac{1}{n}\sum_{i=1}^{n-1} P^i f_0 = g$ must exist and the limiting function g satisfies $\|g\| = 1$ and $Pg = g$.

Write $G = \operatorname{supp} g$. Now for every $A \in \Sigma$, define $\overline{\mu}(A) = \int_A g d\mu$, and let $\overline{L_1} = L_1(X, \Sigma, \overline{\mu})$. Denote the norm in $\overline{L_1}$ by $|||\ |||$ and define $\overline{P} : \overline{L_1} \mapsto \overline{L_1}$ by $\overline{P}h = P(hg)/g$ for $h \in \overline{L_1}$, where $hg = h(x)g(x)$. It is clear that $\overline{P}\chi_X = \chi_X$ and if $h \geq 0$, $\overline{P}h \geq 0$. Moreover, for $h \geq 0$, we have

$$|||\overline{P}h||| = \int_X P(hg)/g d\overline{\mu} = \int_X P(hg) d\mu$$
$$= \int_X hg d\mu = \int_X h d\overline{\mu} = |||h|||.$$

Therefore $\overline{P}$ is a Markov operator.

Let F be the strongly compact set to which all the iterates of P converge. Define $\overline{F} = \{f/g : f \in F\}$. Since

$$|||f/g||| = \int_X f/g d\overline{\mu} = \int_G f d\mu = |||f|_G||| \le \|f\|$$

$\overline{F} \subset \overline{L_1}$. If $\{f_n/g\} \subset \overline{F}$, then there is a subsequence $\{f_n\} \subset F$ such that $f_n \to f$. Then from the preceding inequality, $|||f_n/g - f/g||| \le \|f_n - f\|$. Since every infinite sequence of $\overline{F}$ contains a strongly convergent subsequence, we conclude that $\overline{F}$ is strongly compact. Further, for $h \ge 0$ and $|||h||| = 1$ we have $\|hg\| = |||h||| = 1$, and since P is strongly constrictive, there exists a sequence $\{f_n\} \subset F$ such that $\|P^n(hg) - f_n\| \to 0$. Consequently,

$$\begin{aligned}|||(P^n(hg) - f_n)/g||| &= \int_G |P^n h - f_n/g| g d\mu \\ &\le \|P^n(hg) - f_n\| \to 0,\end{aligned}$$

which implies that $\overline{P}^n h \to \overline{F}$.

We have shown that $\overline{P}$ has the following properties:

1) $\overline{P}\chi_X = \chi_X$
2) $\overline{\mu}(X) < \infty$
3) $\{\overline{P}^n h\} \to \overline{F}$ for all densities $h \in \overline{L_1}$, where $\overline{F}$ is a strongly compact compact subset of $\overline{L_1}$. We may then apply Proposition 2.1.13 to conclude that

$$\overline{P}^n h = \sum_{i=1}^{r} \lambda_i(h) \overline{g}_{\alpha^n(i)} + \varepsilon_n(h), \qquad \text{for } h \in \overline{L_1}$$

where $|||\varepsilon_n(h)||| \to 0$ as $n \to \infty$, α is a permutation of $\{1, 2, \ldots, r\}$ and $\overline{g}_i$ are densities in $\overline{L_1}$ with mutually disjoint support. By letting $g_i = g\overline{g}_i$ we have

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

with $\|g\varepsilon_n(h)\| = |||\varepsilon_n(h)||| \to 0$. From the identity $\overline{P}\overline{g}_i = \overline{g}_{\alpha(i)}$ and from

$$\|g_i\| = \int_X g\overline{g}_i d\mu = \int_X \overline{g}_i d\overline{\mu} = 1$$

we conclude that g_i is a density.

To get an asymptotic decomposition of $f \in L_1(X, \Sigma, \mu)$ under P, we still have to prove a number of lemmas.

2.2.1 Lemma

Let f_0 be the density introduced at the beginning of this section. This density is such that $\frac{1}{n}\sum_{i=1}^{n-1} P^i f_0 = g$. For any $\varepsilon > 0$, there exists an integer m_0 such that

$$\int_{X \setminus G} P^{m_0} f_0 d\mu \leq \varepsilon.$$

Proof:

Suppose this is not the case. There is an $\varepsilon > 0$ such that $\int_{X \setminus G} P^m f_0 d\mu > \varepsilon$ for all $m > 0$. It follows that

$$\begin{aligned} \int_X |\frac{1}{n}\sum_{i=1}^{n-1} P^i f_0 - g| d\mu &\geq \int_{X \setminus G} |\frac{1}{n}\sum_{i=1}^{n-1} P^i f_0 - g| d\mu \\ &= \int_{X \setminus G} \frac{1}{n}\sum_{i=1}^{n-1} P^i f_0 d\mu > 0 \end{aligned}$$

for all $n > 0$, which contradicts the definition of g.

We can now complete the proof of the main result.

Let $f \in L_1(X, \Sigma, \mu)$ and $\varepsilon > 0$ be given. Since $f_0 > 0$, there exists a constant $c > 0$ and $q_1 \in L_1(X, \Sigma, \mu)$ with $\|q_1\| < \frac{\varepsilon}{4}$ such that $|f| \leq cf_0 + q_1$. Choose m_0 as in Lemma 2.2.1 such that

$$\int_{X \setminus G} P^{m_0} f_0 d\mu < \frac{\varepsilon}{4c}.$$

Then

$$\int_{X\setminus G} |P^m f| d\mu \leq \int_{X\setminus G} P^m |f| d\mu$$
$$\leq c \int_{X\setminus G} P^m f_0 d\mu + \int_{X\setminus G} P^m q_1 d\mu$$
$$< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}.$$

On the other hand, there exists a constant $c_1 > 0$ and $q_2 \in L_1(X, \Sigma, \mu)$ with $\|q_2\| < \frac{\varepsilon}{4}$ and $\operatorname{supp} q_2 \subset G$ such that $\chi_G |P^m f| \leq c_1 g + q_2$. Write $h = (\chi_G |P^m f| - q_2)/g$. Then $\chi_G |P^m f| = hg + q_2$, with $|||h||| = \|hg\| < \infty$. Now

$$P^m f = \chi_{X\setminus G} P^m f + \chi_G P^m f$$
$$= \chi_{X\setminus G} P^m f + hg + q_2$$
$$= hg + q_3,$$

where $< \|q_3\| \leq \|\chi_{X\setminus G} P^m f\| + \|q_2\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \frac{3\varepsilon}{4}$. Since $h \in \overline{L_1}$, we have

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

with $\|g\varepsilon_n(h)\| \to 0$. Choose n large enough so that $\|g\varepsilon_n(h)\| < \frac{\varepsilon}{4}$. Then

$$P^{n+m} f = P^n(hg + q_3) = \sum_{i=1}^{r} \lambda_i(h) g_{\alpha^n(i)} + g\varepsilon_n(h) + P^n q_3.$$

Hence $\|P^{n+m} f_0 - \sum_{i=1}^{r} \lambda_i(h) g_{\alpha^n(i)}\| < \frac{3\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon$ and the conditions of Lemma 2.1.12 are satisfied with $n_0 = n + m_0$ and $c_1 = \lambda_{\alpha^{-n}(i)}(h)$ for $1 \leq i \leq r$. This proves the main result in its full generality.

CHAPTER 3

SPECTRAL DECOMPOSITION FOR CONVEX COMBINATION OF MARKOV OPERATORS

3.1 Random Maps

In this chapter we present some results about convex combination of Markov operators. The main question once again is to investigate the asymptotic behaviour of the iterates of these convex combinations. We are espacially interested in the case when one of the Markov operator is constrictive.

The motivation to study this problem arises from the theory of random mappings. More precisely, let τ_1 and τ_2 be two non-singular maps from a closed interval into itself. Let τ_1 have a probability α of being applied at any time, and τ_2 a probability $(1-\alpha)$ of being applied at any time. Consider now the stationary stochastic process defined by

$$x_{n+1} = \beta\tau_1(x_n) + (1-\beta)\tau_2(x_n),$$

where β is a random variable which assumes the value 1 with probability $0 \leq \alpha \leq 1$ and the value 0 with probability $(1-\alpha)$. Consider the event $\{x_{n+1} \leq x\}$. Then

$$\begin{aligned}\Pr\{x_{n+1} \leq x\} &= \Pr\{\beta\tau_1(x_n) + (1-\beta)\tau_2(x_n) \leq x\} \\ &= \Pr\{\beta\tau_1(x_n) + (1-\beta)\tau_2(x_n) \leq x \mid \beta = 1\} \\ &\quad + \Pr\{\beta\tau_1(x_n) + (1-\beta)\tau_2(x_n) \leq x \mid \beta = 0\}\end{aligned}$$

that is

$$\Pr\{x_{n+1} \leq x\} = \alpha \Pr\{\tau_1(x_n) \leq x\} + (1-\alpha)\Pr\{\tau_1(x_n) \leq x\}.$$

If we let x_n have probability density function $f(x) \in L_1(X, \Sigma, \mu)$, then by the Radon-Nykodim Theorem we get

$$P_\alpha f(x) = \alpha P_1 f(x) + (1-\alpha) P_2 f(x)$$

where P_1, P_2 denote the Frobenius-Perron operators of τ_1 and τ_2, respectively.

This is a convex combination of Frobenius-Perron operators, and it is easily checked that P_α is a Markov operator.

An absolutely continuous measure μ is said to be invariant for the stationary stochastic process defined above if its density function $f \in L_1(X, \Sigma, \mu)$ is a fixed point of the operator P_α.

In the following sections, we show that if we assume the convex combination of constrictive to be constrictive we get an upper bound for the number of densities in the spectral representation of P_α.

We will begin by giving an example of a convex combination of constrictive operators such that P_α. is constrictive.

3.2 Example of Convex Combinations of Markov Operators Which Are Constrictive

i) Let $I = [0,1]$ and $L_1 \equiv L_1([0,1], m)$ be the space of Lebesgue integrable functions where m is Lebesgue measure on I. Let C denote the set of non-singular piecewise $\mathbf{C}^2$ maps on I. Let $\tau \in C$. The Frobenius-Perron operator

$P_\tau : L_1 \mapsto L_1$ associated with τ is defined by

$$P_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}[0,1]} f(s)ds.$$

Let $\{\tau_i\}_{i=1}^n \subset C$ and let $0 \leq \lambda_i \leq 1$ be the probability that the map τ_i is applied at any given iteration, where $\sum_{i=1}^n \lambda_i = 1$. This gives rise to a random map $T = \{\tau_i, \lambda_i\}_{i=1}^n$, and a combination $P_T = \sum_{i=1}^n \lambda_i P_{\tau_i}$.

3.2.1 Theorem[3]

Let T be a random map. If, for all $x \in I$, $\sum_{i=1}^n \lambda_i/|\tau_i'(x)| \leq \gamma < 1$, then for all $f \in L_1$

1) $\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} P_T^j f = f^*$ exists in L_1.

2) $P_T f^* = f^*$

3) $\bigvee_0^1 f^* \leq B\|f\|$ for some $B > 0$ which is independent of f, where $\bigvee_0^1 f$ denotes the variation of f over $[0,1]$.

The key of this proof is the inequality [4]:

$$\bigvee_0^1 P_T f \leq \alpha \bigvee_0^1 f + k\|f\|_1$$

for somme $0 < \alpha < 1$ and $k > 0$ independent of f. By an induction argument, we obtain

$$\bigvee_0^1 P_T^i f \leq \alpha^i \bigvee_0^1 f + \frac{k\|f\|_1}{1-\alpha}.$$

Therefore, for every $f \in D$ of bounded variration,

$$\limsup_{j\to\infty} \bigvee_0^1 P_T^j f \leq k_1$$

where $k_1 = k/(1-\alpha)$. Now the set F defined by

$$F = \{g \in D \mid \bigvee_0^1 g \leq k_1\}$$

is weakly precompact by Helley's Theorem. Since the set

$$\{f \in D \mid f \text{ of bounded variation}\}$$

is dense in D, the last inequality tells us that P_T is constrictive.

ii) Let $K : X \times X \mapsto \mathbf{R}$ be a measurable function that satisfies $K(x,y) \geq 0$ and $\int_X K(x,y)dx = 1$. K is called a **stochastic kernel.** Let us define the integral operator $P_1 : L_1 \mapsto L_1$ by

$$P_1 f(x) = \int_X K(x,y)f(y)dy, \qquad \text{for all } f \in L_1.$$

Clearly P_1 is a Markov operator.

3.2.2 Theorem

Let P_1 be the integral operator whose stochastic kernel K satisfies $K(x,y) \leq g(x)$, $g \in L_1$. Let P_2 be any Markov operator. Then $P_\lambda = \lambda P_1 + (1-\lambda)P_2$, $0 < \lambda < 1$, is constrictive.

Proof:

Since $K(x,y) \leq g(x)$,

$$P_1 f(x) = \int_x K(x,y)f(y)dy \leq g(x) \int_X f(y)dy = g(x),$$

where $f(x) \in D$. Therefore

$$P_\lambda f(x) \leq \lambda g(x) + (1-\lambda)P_2 f.$$

We shall prove by induction that

$$P_\lambda^n f(x) \leq \lambda \sum_{i=0}^{n-1} (1-\lambda)^i P_2^i g + (1-\lambda)^n P_2^n f.$$

It is true for $n = 1$. Assume it is true for $n = k$. Then

$$\begin{aligned} P_\lambda^{k+1} f &= P_\lambda(P_\lambda^k f) = \lambda P_1(P_\lambda^k f) + (1-\lambda)P_2(P_\lambda^k f) \\ &\leq \lambda g(x) + (1-\lambda)P_2(\lambda \sum_{i=0}^{k-1}(1-\lambda)^i P_2^i g(x) + (1-\lambda)^k P_2^k f(x)) \\ &= \lambda g(x) + \lambda \sum_{i=0}^{k-1}(1-\lambda)^{i+1} P_2^{i+1} g(x) + (1-\lambda)^{k+1} P_2^{k+1} f(x) \\ &= \lambda \sum_{i=0}^{k}(1-\lambda)^i P_2^i g(x) + (1-\lambda)^{k+1} P_2^{k+1} f(x). \end{aligned}$$

Since P_2 is a Markov operator $\|P_2 g(x)\| \leq \|g(x)\|$, for all i. Hence, for all n,

$$\begin{aligned} \|\lambda \sum_{i=0}^{n-1} P_2^i g(x)\|_1 &\leq \lambda \sum_{i=0}^{n-1}(1-\lambda)^i \|g\|_1 \\ &\leq \|g\|_1(1-(1-\lambda)^n) \\ &\leq \|g\|_1. \end{aligned}$$

thus

$$P_\lambda^n f(x) \leq \tilde{g}(x) + (1-\lambda)^n P_2^n f(x) \quad \text{for } f \in D$$

where $\tilde{g}(x) = \lambda \sum_{i=0}^{\infty}(1-\lambda)^i P_2^i g(x)$, and $\limsup_{n\to\infty} P^n f(x) \leq \tilde{g}(x)$. Since $\tilde{g} \in L_1$, $P_\lambda^n f$ converges to the weakly precompact set $\{f \in L_1 \mid f(x) \leq \tilde{g}(x)\}$. Hence P_λ is constrictive.

3.3 Spectral Representation of Combination of Constrictive Markov Operators

LetP_1 and P_2 be constrictive Markov operators on $L_1(X, A, m)$. Then for any $f \in L_1$, $P_1^n f$ and $P_2^n f$ admit the following representations:

$$P_1^n f = \sum_{i=1}^{l_1} a_i(f) q_{\alpha^n(i)} + A_n f$$

$$P_2^n f = \sum_{i=1}^{l_2} b_i(f) r_{\beta^n(i)} + B_n f$$

where α and β are permutations of the integers $\{1,2,\ldots,l_1\}$ and $\{1,2,\ldots,l_2\}$, respectively. Let $r \le l_1!$ denote the period of $\sum_{i=1}^{l_1} a_i(f)q_{\alpha^n(i)}$. Let

$$P_\lambda = \lambda P_1 + (1-\lambda)P_2, \quad 0 < \lambda < 1$$

be constrictive. Then

$$P_\lambda^n f = \sum_{i=0}^{l} c_i(f) s_{\gamma^n(i)} + C_n(f).$$

In this section we shall prove that $l \le \min(l_1, l_2)$. Let

$$\mathcal{P}_1 = \bigcup_{i=1}^{l_1} \{\operatorname{supp} q_i\} \text{ and } \mathcal{P}_2 = \bigcup_{i=1}^{l_2} \{\operatorname{supp} r_i\}.$$

3.3.1 Lemma

Assume P_λ is constrictive. Then the support of every density s_i must intersects both $\mathcal{P}_1$ and $\mathcal{P}_2$ on a set of positive m measure, i.e., $m(\operatorname{supp} s_i \cap \mathcal{P}_j) \neq 0$, $i = 1,2,\ldots,l$, $j = 1,2$.

Proof:

We need only prove this for $j = 1$ and for any s_i, say s_1. Since P_λ is constrictive, there exists an integer t such that $P_\lambda^{kt} s_1 = s_1$ for $k = 1,2,\ldots$. From the expansion

$$P_\lambda^n f = \lambda^n P_1^n f + \lambda^{n-1}(1-\lambda)[P_1^{n-1}P_2 + P_1^{n-2}P_2P_1 + \cdots + P_2P_1^{n-1}]f$$
$$+ \lambda^{n-2}(1-\lambda)^2[P_1^{n-2}P_2^2 + \cdots + P_2^2P_1^{n-2}]f + \cdots + (1-\lambda)^n P_2^n f,$$

we have that

$$s_1 = P_\lambda^{kt} s_1 = \lambda^{kt}[\sum_{i=1}^{l_1} a_i(s_1) q_{\alpha^{kt}(i)} + A_{kt}(s_1)]$$
$$+ [\text{terms involving } P_1 \text{ and } P_2].$$

Hence

$$\|s_1 - \lambda^{kt} \sum_{i=1}^{l_1} a_i(s_1) q_{\alpha^{kt}(i)}\| \le \lambda^{kt} \|A_{kt}(s_1)\| + \|\text{terms involving } P_1 \text{ and } P_2\|.$$

The right hand side of the last inequality is bounded by $1 - \lambda^{kt}$. Let us choose $k = k_0$ sufficiently large finite, such that

$$\lambda_{k_0 t}\|A_{k_0 t}(s_1)\|_1 + \|\text{terms involving } P_1 \text{ and } P_2\|_1 < 1$$

then

$$\|s_1 - \lambda^{k_0 t}\sum_{i=1}^{J_1} a_i(s_1) q_{\alpha^{k_0 t}(i)}\|_1 < 1$$

If therefore follows that for some q_j,

$$m(\operatorname{supp} s_1 \cap \operatorname{supp} q_j) \neq 0.$$

From Lemma 3.3.1, we know that the support of any s_i must intersect the support of some q_i on a set of positive m measure. If $l > l_1$, then at least two s_i's, s_1 and s_2 say, must intersect the support of a common q_i, say q_1, on a set of positive m measure. Let

$$D_1 = \operatorname{supp} q_1 \cap \operatorname{supp} s_1,$$

$$D_2 = \operatorname{supp} q_1 \cap \operatorname{supp} s_2.$$

Then $m(D_1) > 0$ and $m(D_2) > 0$. Let $s_1' = s_1|_{D_1}$ Then.

$$\operatorname{supp} s_1' \subset \operatorname{supp} q_1.$$

From [1], it follows that for all n,

$$\operatorname{supp} P_1^n s_1' \subset \operatorname{supp} P_1^n q_1.$$

The same argument applies to $s_2' = s_2|_{D_2}$. Since $\operatorname{supp} s_1$ and $\operatorname{supp} s_2$ are disjoint, $\operatorname{supp} s_1'$ and $\operatorname{supp} s_2'$ are disjoint. Therefore

$$\|s_1' - s_2'\|_1 = \|s_1'\|_1 + \|s_2'\|_1.$$

Since s_1' and s_2' are equal to 0 on the support of q_j, $j \neq 1$, and the period of the limit sequence $\sum_{i=1}^{l_1} a_i(f) q_{\alpha^n(i)}$ is t we have

$$P_1^{kt} s_1' = a_1(s_1')q_1 + A_{kt}(s_1')$$

and

$$P_1^{kt} s_2' = a_1(s_2')q_1 + A_{kt}(s_2').$$

Let us assume that $\|s_1'\|_1 \geq \|s_2'\|_1$. Then let us choose k large enough so that

$$\|A_{kt}(s_i')\|_1 < \frac{1}{2}\|s_2'\|_1, \qquad i = 1, 2.$$

Therefore

$$\begin{aligned}\|P_1^{kt}(s_1' - s_2')\|_1 &\leq |a_1(s_1') - a_1(s_2')|\|q_1\| + \|A_{kt}(s_1')\|_1 + \|A_{kt}(s_2')\|_1 \\ &< |a_1(s_1') - a_1(s_2')| + \|s_2'\|_1\end{aligned}$$

But the right hand side will be less than $|a_1(s_1')| + \|s_2'\|_1$ or $|a_1(s_2')| + \|s_2'\|_1$. Since $|a_1(s_i')| \leq \|a_1\|\|s_1'\|_1$ and $\|a_1\| < 1$, we get $|a_1(s_1')| + \|s_2'\|_1 \leq \|s_1'\| + \|s_2'\|$ or $|a_1(s_2')| + \|s_2'\|_1 \leq \|s_2'\| + \|s_2'\| \leq \|s_1'\| + \|s_2'\|$. In either case

$$\|P_1^{kt}(s_1' - s_2')\|_1 < \|s_1'\|_1 + \|s_2'\|_1 = \|s_1' - s_2'\|_1.$$

We can now prove:

3.3.2 Theorem

Let P_1 and P_2 be constrictive Markov operators such that P_1, P_2 have l_1, l_2 densities in their spectral representations, respectively. Then, if

$$P_\lambda = \lambda P_1 - (1 - \lambda)P_2$$

is constrictive, its representation has l densities, where

$$l \leq \min(l_1, l_2).$$

Proof:

We shall prove that for $l > l_1$ we get a contradiction. If $l > l_1$, there exists q_1, s_1 and s_2 as above, such that

$$m(\operatorname{supp} s_i \cap \operatorname{supp} q_1) > 0, \qquad i = 1, 2.$$

Write

$$s_1 = (s_1 - s_1') + s_1'$$
$$s_2 = (s_2 - s_2') + s_2'$$

where s_1' and s_2' are defined above. Since P_λ is constrictive $\operatorname{supp} s_1' \cap \operatorname{supp} s_2 = \emptyset$. Therefore, $\|s_1 - s_2\| = 2$. Let t' be the period of the finite sequence in the representation of P_λ. Since s_1 and s_2 are fixed under $P_\lambda^{k't'}$,

$$\|P_\lambda^{k't'}(s_1 - s_2)\| = 2 \qquad \text{for } k' = 1, 2, \ldots.$$

But now,

$$P_\lambda^{k't'}(s_1 - s_2) = \lambda^{k't'} P_1^{k't'}(s_1 - s_2) + \text{ [terms involving } P_1 \text{ and } P_2].$$

From the representation of $P_\lambda^n f$ in Lemma 3.3.1, it can be shown that

$$\|\text{terms involving } P_1 \text{ and } P_2\| \le 2(1 - \lambda^{k't'}).$$

Choose k' so that $k't' > k_0 t$ where k_0 satisfies

$$\lambda^{k_0 t}\|Q_{k_0 t}(s_1)\| + \|\text{terms involving } P_1 \text{ and } P_2\| < 1.$$

Then

$$\begin{aligned}
2 = \|P_\lambda^{k't'}(s_1 - s_2)\| &\le \lambda^{k't'}\|P_1^{k't'}(s_1 - s_2)\| + 2(1 - \lambda^{k't'}) \\
&= \lambda^{k't'}\|P_1^{k't' - kt} P_1^{kt}[(s_1 - s_1') + s_1' - (s_2 - s_2') - s_2']\| + 2(1 - \lambda^{k't'}) \\
&\le \lambda^{k't'}[\|P_1^{kt}(s_1 - s_1')\| + \|P_1^{kt}(s_2 - s_2')\| + \|P_1^{kt}(s_1' - s_2')\|] + 2(1 - \lambda^{k't'})
\end{aligned}$$

Since $\|P_1^{kt}(s_1' - s_2')\| < \|s_1' - s_2'\|$,

$$2 < \lambda^{k't'}[\|s_1 - s_1'\| + \|s_2 - s_2'\| + \|s_1' - s_2'\|] + 2(1 - \lambda^{k't'}) \quad ,$$

$$= \lambda^{k't'}\|s_1 - s_2\| + 2(1 - \lambda^{k't'}) = 2$$

since, s_1 and s_2 have disjoint support and $s_1' = s_1|_{A_1}$ and $s_2' = s_2|_{A_2}$ have disjoint support. Thus we have a contradiction. Repeating the entire argument with P_1 and P_2 interchanged yields the desired results.

3.3.3 Corollary

If either P_1 or P_2 is exact, then P_λ is exact.

Proof:

$\min(l_1, l_2) = 1$.

Example

If P_1 and P_2 are Frobenius-Perron operators associated with maps τ_1, τ_2 of the interval $[0,1]$, and if either τ_1 or τ_2 is exact, then the random map $(\tau_1, \tau_2, \lambda, 1-\lambda)$, whose Markov operator is

$$P_\lambda = \lambda P_1 + (1-\lambda)P_2$$

is exact provided P_λ admits a fixed point. In this case, the random map has a unique absolutely continuous measure. This generalizes the examples on p. 821 of [3].

3.3.4 Corollary

If P_1 is an integral operator with kernel $K(x,y) \leq g(x)$ where $g \in L_1$ and P_2 is any Markov operator. Then

$$P_\lambda = \lambda P_1 + (1-\lambda)P_2 \qquad 0 < \lambda < 1,$$

is a constrictive operator and the number of densities in its spectral representation is bounded by the number of densities in the spectral representation of P_1.

Proof:

By example ii) in section 2., P_λ is constrictive. The result follows from the proof of Theorem 3.3.2.

3.4 Spectral Representation and Ergodic Decomposition

Let P be any constrictive Markov operator. Then

$$Pf = \sum_{i=1}^{l} a_i(f)g_i + Q(f).$$

Let $f^* \in D$ be a fixed point of P, and let b be the period of $\sum_{i=1}^{l} a_i(f)g_i$. Then $f^* = P^{kb}f^*$ implies that

$$f^* = \sum_{i=1}^{l} a_i(f^*)g_i + Q_{kb}(f^*)$$

for all k. But $\|Q_{kb}(f)\| \to 0$ as $k \to \infty$. Since f^* and $\sum_{i=1}^{l} a_i(f^*)g_i$ are fixed functions independent of k, $Q_{kb}(f) = 0$. Thus

$$f^* = \sum_{i=1}^{l} a_i(f^*)g_i.$$

From this it follows that $\operatorname{supp} f^* \subset \bigcup_{i=1}^{l} \operatorname{supp} g_i$ for every $f^* \in D$ fixed under P. Hence the maximum number of independent fixed density functions must be less than or equal to l. That is, the number of ergodic components, n, of P is bounded by the number of densities in the spectral representation, i.e, $n \leq l$.

Now, if P_λ, P_1, P_2 are constrictive, where

$$P_\lambda = \lambda P_1 + (1-\lambda)P_2$$

it follows from Theorem that $l \leq \min(l_1, l_2)$, where l is the number of densities in the spectral representation of P_λ.

In [3], it is shown that if $\tau : [0,1] \mapsto [0,1]$ is piecewise continuous, piecewise $\mathbf{C}^1$ with finitely many discontinuities n, Then τ has at most n absolutely continuous invariant measures whose densities $f_1, f_2, \ldots, f_n$ are disjoint. Let P_τ be the Frobenius-Perron operator of τ. then $P_\tau f_i = f_i$, $i = 1, 2, \ldots, n$. Under the assumptions that $|\tau'(x)| \geq \beta > 1$ and $|\tau''(x)|/|\tau'(x)|^2 \leq c < \infty$, it can be shown that P_τ is constrictive [3]. Hence by the spectral representation Theorem, there exists l densities $\{g_1, g_2, \ldots, g_n\}$ such that $P_\tau g_i = g_{\alpha(i)}$, where $\alpha : 1, 2, \ldots, l \mapsto 1, 2, \ldots, l$ is a permutation.

3.4.1 Theorem

Let τ_1, τ_2 be maps as in the foregoing paragraph and let P_1, P_2 be their Frobenius-Perron operators. Then

$$P_\lambda = \lambda P_1 + (1 - \lambda) P_2$$

is constrictive. If the permutations in the spectral representations of P_1 and P_2 are cyclical, we have

$$n \leq \min(l_1, l_2) = \min(n_1, n_2),$$

where n_i is the number of independent densities in the ergodic decomposition of τ_i, $i = 1, 2$.

Proof:

Consider τ_1 and let $P_1 = P_{\tau_1}$. Let $\alpha_1 : 1, 2, \ldots, l_1 \mapsto 1, 2, \ldots, l_1$ be a cyclical permutation. (This is equivalent to P being ergodic [?, Thm 5.5.1]) Then $P_1^{l_1} g_i = g_i$, where $\{g_1, g_2, \ldots, g_{l_1}\}$ are the densities of the spectral representation of P_1. But by the ergodic decomposition Theorem [3], there can be at most n_1 such densities.

Hence $l_1 \leq n_1$. But $l_1 \geq n_1$. Thus $n_1 = l_1$ and similarly $n_2 = l_2$. It therefore follows that

$$n \leq \min(l_1, l_2) = \min(n_1, n_2).$$

REFERENCES

[1] LASOTA, A.; LI,T. Y.; YORKE, J. A. "Asymptotic Periodicity for the Iterates of Markov Operators", Trans. Amer. Math. Soc. 286 (1984), no 2, 751-764.

[2] KOMORNIK, J. "Asymptotic Periodicity for the Iterates of a Weakly Constrictive Markov Operators", Tokohu Math J.(2) 38 (1986), no 1, 15-27.

[3] PELIKAN, S. "Invariant Densities for Random Maps of the Interval", Trans. Amer. Math. Soc. 281 (1984), no 2, 813-825.

[4] LASOTA, A.; YORKE, J. A. "On the Existence of In variant Measure for Piecewise Monotonic Transformation" Trans. Amer. Math. Soc. 281 (1984), no 2, 813-825.

[5] LASOTA, A.; MACKAY, M. C. "Probabilistic Properties of Deterministic Systems", Cambridge University Press, Cambridge-New York (1985).

[6] KRENGEL, U. "Ergodic Theorems" Walter De Gruyter & Co., Berlin-New York (1985)

[7] DUNFORD, N.; SCHWARTZ, J. T. "Linear Operators Part I", Interscience Publishers Inc., New York, (1964)