SOME LIMIT THEOREMS FOR MARKOV PROCESSES

BY
S. HOROWITZ(1)

The aim of this paper is to prove some limit theorems for Markov processes using only functional analytic methods. Some of our results were proved in [7], [8] and [5] by probabilistic methods. We prove in the Appendix a theorem on Markov processes that have no finite invariant measure.

1. **Definitions and notations.** Let (X, Σ, m) be a measure space, such that m is a probability measure. Let $P(x,A)$ be a Markov transformation, i.e. a function on $X \times \Sigma$ such that, for each $x \in X$, $P(x,\cdot)$ is a probability measure and for each $A \in \Sigma$, $P(.\,,A)$ is a measurable function. A Markov transformation induces an operator on $B(X,\Sigma)$, the space of the bounded and measurable function, and on $M(X,\Sigma)$ the space of the signed measures, by:

$$(1.1) \qquad (Pf)(x) = \int f(y)P(x,dy)$$

$$(1.2) \qquad (\nu P)(A) = \int P(x,A)\nu(dx).$$

Thus, if 1_A denotes the characteristic function of $A \in \Sigma$ and δ_x the Dirac measure at x then

$$(P1_A)(x) = P(x,A), \quad (\delta_x P)(A) = P(x,A).$$

Eq. (1.2) will occasionally be used for σ-finite positive measures.

The two operators are related by

$$(1.3) \qquad \int (Pf)(x)\nu(dx) = \int f(x)(\nu P)(dx).$$

The iterates of P are defined inductively by

$$(1.4) \qquad P^n(x,A) = \int P^{n-k}(x,dy)P^k(y,A), \qquad 0 < k < n.$$

The definition corresponds to the notion of powers of the operator P considered either on bounded measurable functions or on signed measures. The measure

Received February 14, 1968.

(1) This paper is a part of the author's Ph.D. thesis to be submitted to the Hebrew University of Jerusalem. The author wishes to express his thanks to Professor S. R. Foguel for much valuable advice and encouragement.

m is assumed to satisfy

$$m \succ mP \tag{1.5}$$

(mP is absolutely continuous with respect to m). Hence if $m(A) = 0$ then $P(x, A) = 0$ a.e. with respect to m (a.e. m).

We shall also define the operator I_A, for $A \in \Sigma$, by

$$I_A f(x) = 1_A(x) \cdot f(x) \tag{1.6}$$

$$\nu I_A(B) = \nu(B \cap A). \tag{1.7}$$

DEFINITION 1. *The process (X, Σ, m, P) is said to be conservative if for every $A \in \Sigma$ with $m(A) > 0$, $\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = 1$ is satisfied* a.e. *m on A.*

The process is called ergodic if $P(x, A) = 1_A(x)$ implies $m(A) = 0$ or $m(A) = 1$.

It can be shown that if the process is conservative and ergodic then $\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = 1$ a.e. m. for every A with $m(A) > 0$ (see, for example, [2] Theorem 2.3).

REMARK. $\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x)$ is the probability that x enters A at least once.

2. Processes with an invariant measure.

In the rest of this paper we shall assume the following:

ASSUMPTION 1. *The process is conservative and ergodic and there exists a σ finite measure μ which is equivalent to m, and $\mu = \mu P$.*

It is easy to see that P is well defined as an operator on $L_p(X, \Sigma, \mu)$ for every $1 \leqq p \leqq \infty$.

If $|f(x)| < M$ then $|Pf(x)| < M$, hence $\|P\|_\infty \leqq 1$. On the other hand, $\|Pf\|_1 \leqq \|P|f|\|_1 = \int P|f|\mu(dx) = \int |f| \mu P(dx) = \int |f| \mu(dx) = \|f\|_1$, hence $\|P\|_1 \leqq 1$.

Thus by the Riesz Convexity Theorem the operator P is a contraction on $L_p(x, \Sigma, \mu)$ for every $1 \leqq p \leqq \infty$. Let us now consider the action of P on the signed measures. It is easy to see that if $\nu \prec \mu$ then also $\nu P \prec \mu$, or P leaves the subspace, consisting of signed measures that are weaker than μ, invariant. If $\nu \prec \mu$ then $d\nu = f d\mu$ where $f \in L_1(x, \Sigma, \mu)$ is the Radon-Nikodim derivative of ν with respect to μ.

Let us denote:

$$fP^n = g \quad \textit{iff whenever } d\nu = f d\mu \textit{ then } g = \frac{d\nu P^n}{d\mu}. \tag{2.1}$$

This is the adjoint operator of P, i.e. $P^* f = fP$. Because of assumption 1, it is clear:

$$P^* 1 = 1P = 1, \tag{2.2}$$

so it is clear that P^* is also a contraction on $L_p(X,\Sigma,\mu)$ for every $1 \leqq p \leqq \infty$. Notice that P^* is defined as an operator on $L_p(X,\Sigma,\mu)$ and need not be induced by a Markov transformation.

3. **P as an operator on $L_2(X,\Sigma,\mu)$**

Let us consider P as an operator on $L_2(X,\Sigma,\mu)$; we denote

(3.1) $$K = \{f \mid f \in L_2(\mu),\ \|P^n f\| = \|P^{*}f\| = \|f\|, \forall n\}$$

(3.2) $$\Sigma_1 = \text{the } \sigma\text{-field generated by sets } A \text{ with } 1_A \in K.$$

In [3] the following results are proved:

(a) K is invariant under P and P^*, and P restricted to K is a unitary operator.

(b) If $f \perp K$ then weak $\lim P^n f =$ weak $\lim P^{*n} f = 0$.

(c) $K = L_2(X,\Sigma_1,\mu)$ equivalently $f \in K$ iff $f \in L_2(X,\Sigma,\mu)$ and is Σ_1 measurable.

(d) If $A \in \Sigma_1$ and $\mu(A) < \infty$ then $P1_A$ and P^*1_A are both characteristic functions of sets in Σ_1.

Assumption 2. *The set Σ_1 is atomic.*

If $\mu(X) = \infty$ then $\Sigma_1 = \phi$, if $\mu(X) < \infty$ then $\Sigma_1 = \{W \cup PW \cup \cdots \cup P^{k-1}W\}$ and $P^k W = W$, because of the assumption that P is ergodic and conservative. The integer k is called the order of W.

The following theorem is a simple consequence of theorem 8 of [3].

Theorem 1. *Let $\nu \prec \mu$, be a finite measure; then*

(a) *If μ is an infinite measure then for every set A with $\mu(A) < \infty$, $\lim_{n\to\infty}(\nu P^n)(A) = 0$.*

(b) *If μ is a probability measure and $A \subset W$, where $\Sigma_1 = \{W \cup PW \cup \times P^{k-1}W\}$ then*

$$\lim_{n\to\infty} (\nu P^{nk+r})(A) = k\mu(A)(\nu P^r)(W).$$

Remark. Theorem 1 remains true if we replace P by P^*.

4. **Markov processes satisfying Harris' condition.** Let (X,Σ,m,P) be a Markov process as in Section 1.

Definition 2. *The process is said to satisfy Harris' condition if $m(A) > 0$ implies*

(4.1) $$\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = 1 \quad \textit{for all } x \in X.$$

It is well known (see, for example, [2], [5], [7], [8]) that Harris' condition implies Assumption 1. Let us denote

$$P^n(x, .) = Q_n(x, .) + R_n(x, .) \tag{4.2}$$

$$Q_n(x, .) \succ m, \; R_n(x, .) \perp m$$

$$\phi_n(x, y) = \frac{dQ_n(x,.)}{d\mu} \text{ where } \mu \sim m, \; \mu = \mu P.$$

We shall assume that Σ is separable, then $\phi_n(x,y)$ is $\Sigma \times \Sigma$ measurable.

If Harris' condition is satisfied then for each x, and for each set A with $\mu(A) > 0$ there is an integer n such that $Q_n(x,A) > 0$.

Because if there is an x and a set A with $\mu(A) > 0$ and $Q_n(x,A) = 0$ for all n. then $P^n(x,A) = R_n(x,A)$. Let $F_n = \operatorname{supp} R_n(x, .)$, $F = \cup_{n=1}^{\infty} F_n$, $\mu(F) = 0$. hence $\sum_{n=1}^{\infty} P^n(x, A - F) = 0$. But $\mu(A - F) > 0$, and this contradicts Harris' condition.

Theorem 6 of [3] says that if A is in the non-atomic part of Σ_1, then $\mu\{x \mid Q_n(x,A) > 0\} = 0$ for every n, therefore Harris' condition implies that Σ_1 is atomic.

In the following lemma we shall give a condition that is equivalent to Harris'.

LEMMA. *The process (X, Σ, m, P) satisfies Harris' condition if and only if for every set N with $m(N) = 0$*

$$\lim_{n\to\infty} P^n(x,N) = 0 \quad \textit{for all } x \in X. \tag{4.3}$$

We shall first prove two propositions:

PROPOSITION 1. *For every integer n and for every set A,*

$$\sum_{k=0}^{n} (I_{A^c}P)^k 1_A(x) + (I_{A^c}P)^{n+1} 1(x) = 1 \tag{4.4}$$

Proof. By induction. For $n = 0$: $1_A + I_{A^c}P1 = 1_A + 1_{A^c} = 1$. Assume for n, we shall prove for $n+1$:

$$\sum_{k=1}^{n+1} (I_{A^c}P)^k 1_A + (I_{A^c}P)^{n+2} 1 = \sum_{k=0}^{n} (I_{A^c}P)^k 1_A + (I_{A^c}P)^{n+1} 1_A + (I_{A^c}P)^{n+1} I_{A^c}P1$$

$$= \sum_{k=0}^{n} (I_{A^c}P)^k 1_A + (I_{A^c}P)^{n+1}(1_A + 1_{A^c}) = \sum_{k=0}^{n} (I_{A^c}P)^k 1_A + (I_{A^c}P)^{n+1} 1 = 1.$$

PROPOSITION 2. *For every $x \in X$ and for every set A, the sequence $\sum_{n=0}^{\infty} P^k (I_{A^c}P)^n 1_A$ is decreasing, and therefore the limit*

$$\lim_{k\to\infty} \sum_{n=0}^{\infty} P^k (I_{A^c}P)^n 1_A(x)$$

exists.

Proof.

$$\sum_{n=0}^{\infty} P^{k+1}(I_{A^c}P)^n 1_A(x) = \sum_{n=0}^{\infty} P^k I_{A^c} P(I_{A^c}P)^n 1_A(x)$$

$$+ \sum_{n=0}^{\infty} P^k I_A P(I_{A^c}P)^n 1_A(x) = \sum_{n=1}^{\infty} P^k(I_{A^c}P)^n 1_A(x)$$

$$+ P^k I_A \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = \sum_{n=0}^{\infty} P^k(I_{A^c}P)^n 1_A(x) - P^k 1_A(x)$$

$$+ P^k I_A \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = \sum_{n=0}^{\infty} P^k(I_{A^c}P)^n 1_A(x)$$

$$- P^k I_A\left(1 - \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x)\right) \leqq \sum_{n=0}^{\infty} P^k(I_{A^c}P)^n 1_A(x),$$

Because $1 - \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) \geqq 0$.

REMARK.

$\lim_{k\to\infty} \sum_{n=0}^{\infty} P^k(I_{A^c}P)^n 1_A(x)$ is the probability that x enters A infinitely many times.

Proof of the Lemma.

(a) Assume Harris' condition is satisfied. If N is a set with $m(N) = 0$, let us denote $F = \{x \mid \sum_{n=1}^{\infty} P^n(x,N) > 0\}$, then, by $m \succ mP$, $m(F) = 0$. But

$$P^n(x,N) = (I_F P)^n(x,N).$$

This can be proved inductively, assume $P^n 1_N = (I_F P)^n 1_N$, and then:

$$P^{n+1}1_N = PP^n 1_N = P(I_F P)^n 1_N = (I_F P)^{n+1} 1_N + (I_{F^c}P)(I_F P)^n 1_N.$$

but $(I_{F^c}P)(I_F P)^n 1_N \leqq I_{F^c} P^{n+1} 1_N = 0$, hence $P^{n+1}1_N = (I_F P)^{n+1} 1_N$. By Proposition 1 $\sum_{k=0}^{n-1} (I_F P)^k 1_{F^c}(x) + (I_F P)^n 1(x) = 1$. Let n tend to ∞, then by Harris' condition,

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} (I_F P)^k 1_{F^c}(x) = \sum_{k=0}^{\infty} (I_F P)^k 1_{F^c}(x) = 1_{F^c}(x)$$

$$+ I_F \sum_{k=0}^{\infty} P(I_F P)^k 1_{F^c}(x) = 1_{F^c}(x) + 1_F(x) = 1.$$

Hence:

$$\lim_{n\to\infty} P^n(x,N) = \lim_{n\to\infty} (I_F P)^n(x,N) \leqq \lim_{n\to\infty} (I_F P)^n 1(x) = 0.$$

(b) Assume (4.3). By Assumption 1 the process is conservative and ergodic. Therefore for every $A \in \Sigma$, with $m(A) > 0$, there exists a set N with $m(N) = 0$ so that for every $x \in N^c$, $\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) = 1$, ($N$ may depend on A). We shall prove that $N = \varnothing$. Assume the contrary, take $x \in N$ then

$$\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) < 1.$$

But

$$\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) \geqq \lim_{k\to\infty} \sum_{n=0}^{\infty} P^k(I_{A^c}P)^n 1_A(x)$$

$$= \lim_{k\to\infty} P^k(I_{N^c}+I_N)\sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x) \geqq \lim_{k\to\infty} P^k I_{N^c} \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(x)$$

$$= \lim_{k\to\infty} \int_{N^c} P^k(x,dy) \sum_{n=0}^{\infty} P(I_{A^c}P)^n 1_A(y) = \lim_{k\to\infty} P^k(x,N^c) = 1,$$

by (4.3). Hence $x \notin N$, a contradiction. Therefore $N = \emptyset$.

REMARK. The "only if" part of our lemma is lemma 2.4 of Jain [5].

DEFINITION 3. The process (X, Σ, m, P) is said to satisfy *Doeblin's condition* if there exists an integer d such that if $m(N) = 0$ then $\sup\{P^d(x,N) \mid x \in X\} < 1$.

Let us put in theorem 10 of [3] $\mu = \delta_x$, $\delta_x P^n = \tau_n + \sigma_n$ where $\tau_n \prec m$, $\sigma_n \perp m$, then if $m(N) = 0$,

$$\lim_{n\to\infty} P^n(x,N) = \lim_{n\to\infty} \sigma_n(N) \leqq \lim_{n\to\infty} \sigma_n(X) = 0.$$

Hence Doeblin's condition implies Harris' condition. On the other hand, in [6] there is an example that satisfies Harris' condition but not Doeblin's condition.

REMARK. There is no loss generality in assuming that the process is ergodic: if (4.3) is satisfied then $P(x,A) = 1_A(x)$ implies $m(A) > 0$. Hence $X = \bigcup_j A_j$ where each A_j is ergodic.

THEOREM 2. *Let ν be a finite measure, let P satisfy Harris' condition, and $\nu P^n = \tau_n + \sigma_n$ where $\tau_n \prec m$, $\sigma_n \perp m$, then $\lim_{n\to\infty} \sigma_n(X) = 0$.*

Proof. Let $R_n(x,.)$ as in (4.2). Let us first prove:

(4.5) $$\lim_{n\to\infty} R_n(x,X) = 0 \quad \text{for } all \ x \in X.$$

Let $F_n = \operatorname{supp} R_n(x,.)$, ($F_n$ depends on x) $F = \bigcup_{n=1}^{\infty} F_n$ then $m(F) = 0$, and by (4.3)

$$\lim_{n\to\infty} R_n(x,X) = \lim_{n\to\infty} R_n(x,F) = \lim_{n\to\infty} P^n(x,F) = 0.$$

Let ν be any measure, then,

$$\nu P^n(A) = \int Q_n(x,A)\nu(dx) + \int R_n(x,A)\nu(dx) = \int_A \int \phi_n(x,y)\nu(dx)\mu(dy)$$
$$+ \int R_n(x,A)\nu(dx)$$

so, $\nu Q_n \prec \mu$ (or $\nu Q_n \prec m$). Hence $\sigma_n(X) \leqq \nu R_n(X)$ and by (4.5) and by the dominated convergence theorem we have:

$$\lim_{n\to\infty} \sigma_n(X) \leqq \lim_{n\to\infty} \nu R_n(X) = 0.$$

THEOREM 3. *Assume that P satisfies Harris' condition. Let μ be the invariant measure of Assumption* 1.
Let ν be any finite measure. Then:
(a) *If $\mu(X) = \infty$ then for every $A \in \Sigma$ with $\mu(A) < \infty$, $\lim_{n\to\infty} \nu P^n(A) = 0$.*
(b) *If $\mu(X) = 1$ and $\Sigma_1 = \{W \cup PW \cup \cdots \cup P^{k-1}W\}$ then for every $A \subset W$, $\lim_{n\to\infty} \nu P^{nk+r} = k \cdot \mu(A)(\nu P^r)(W)$.*

Proof. (a) If μ is infinite, let $\nu P^n = \tau_n + \sigma_n$ where $\tau_n \prec \mu$, $\sigma_n \perp \mu$. For each $\varepsilon > 0$ we can choose an integer n_0 such that $\sigma_{n_0}(X) < \varepsilon$, by Theorem 2. Hence, for every set A with $\mu(A) < \infty$:

$$\begin{aligned} \nu P^n(A) &= \tau_{n_0} P^{n-n_0}(A) + \sigma_{n_0} P^{n-n_0}(A) \leqq \tau_{n_0} P^{n-n_0}(A) + \sigma_{n_0} P^{n-n_0}(X) \\ &\leqq \tau_{n_0} P^{n-n_0}(A) + \sigma_{n_0}(X) < \tau_{n_0} P^{n-n_0}(A) + \varepsilon. \end{aligned}$$

But $\lim_{n\to\infty} \tau_{n_0} P^{n-n_0}(A) = 0$, by Theorem 1, and ε is arbitrary, therefore $\lim_{n\to\infty} \nu P^n(A) = 0$.

(b) If μ is a probability measure, let $\nu P^n = \tau_n + \sigma_n$ where $\tau_n \prec \mu$, $\sigma_n \perp \mu$. For each $\varepsilon > 0$ we can choose an integer n_0 such that $\sigma_{n_0}(X) < \varepsilon$, and $\tau_{n_0}(X) > \nu(X) - \varepsilon$. Let us first assume that Σ_1 is trivial and $\nu(X) = 1$. Then: $\lim_{n\to\infty} \tau_{n_0} P^{n-n_0}(A) = \mu(A) \cdot \tau_{n_0}(X)$ by Theorem 1. Hence, for every n sufficiently large,

$$\mu(A)(1 - 2\varepsilon) \leqq \tau_{n_0} P^{n-n_0}(A) \leqq \mu(A) + \varepsilon.$$

Also, for all n, $\sigma_{n_0} P^{n-n_0}(A) \leqq \sigma_{n_0} P^{n-n_0}(X) \leqq \sigma_{n_0}(X) < \varepsilon$. Hence:

$$\mu(A)(1 - 2\varepsilon) \leqq \nu P^n(A) = \tau_{n_0} P^{n-n_0}(A) + \sigma_{n_0} P^{n-n_0}(A) \leqq \mu(A) + 2\varepsilon.$$

But ε is arbitrary, therefore $\lim_{n\to\infty} \nu P^n(A) = \mu(A)$. The generalization for Σ_1 of order k is obvious.

If we choose $\nu = \delta_x$ we get:

COROLLARY. *Let P, μ, Σ_1 as in Theorem 3, then for every $x \in X$:*
(a) *If μ is infinite then $\mu(A) < \infty$ implies $\lim_{n\to\infty} P^n(x, A) = 0$.*
(b) *If μ is a probability measure then $A \subset W$ implies*

$$\lim_{n\to\infty} P^{nk+r}(x, A) = k\mu(A) \cdot P^r(x, W).$$

REMARK. Part (a) of this corollary is Theorem 2.5 of Jain [5]. Part (b) appears, for example, in [8].

THEOREM 4. *Let P, Σ_1, μ be as in Theorem 3, part (b), then for every $A \subset W$,*

$$\lim_{n\to\infty} P^{*nk+r}1_A(x) = k\mu(A)\cdot P^{*r}1_W(x) \text{ a.e. } \mu.$$

Proof. Let us first assume that Σ_1 is trivial. Let $P^n(x,A) = Q_n(x,A) + R_n(x,A)$ as in (4.2). Let Q_n and R_n be the operators that are induced by $Q_n(x,\,.\,)$ and $R_n(\cdot\,,\cdot)$ respectively. For all $x \in X$, $\lim_{n\to\infty} R_n(x,X) = 0$ by (4.5). By the dominated convergence theorem we have $\lim_{n\to\infty}\int R_n^* 1\mu(dx) = \lim_{n\to\infty}\int R_n 1\mu(dx) = 0$, where R_n^* is the adjoint of R_n. Hence we can find a sequence of integers $\{n_k\}$ such that $\lim_{k\to\infty} R_{n_k}^* 1(x) = 0$ a.e. μ. Hence for every $x \in X$ that is not in an exceptional nil set, there can be found an integer n_{k_0} such that $R_{n_{k_0}}^* 1(x) < \varepsilon$.

Let us write $P^{*n} = Q_n^* + R_n^*$.

Q_n is an integral operator with the kernel $\phi_n(x,y)$, and therefore Q_n^* is also an integral operator with the kernel $\phi_n(y,x)$. Hence:

$$P^{*n}1_A(x) = Q_{n_{k_0}}^* P^{*n-n_{k_0}} 1_A(x) + R_{n_{k_0}}^* P^{*n-n_{k_0}} 1_A(x).$$

But

$$R_{n_{k_0}}^* P^{*n-n_{k_0}} 1_A(x) \leqq R_{n_{k_0}}^* P^{*n-n_{k_0}} 1(x) = R_{n_{k_0}}^* 1(x) < \varepsilon.$$

Hence

$$Q_{n_{k_0}}^* P^{*n-n_{k_0}} 1_A(x) \leqq P^{*n}1_A(x) \leqq Q_{n_{k_0}}^* P^{*n-n_{k_0}} 1_A(x) + \varepsilon.$$

Denote:

$$\delta_x Q_{n_{k_0}}^*(A) = Q_{n_{k_0}}^* 1_A(x) = \int_A \phi_{n_{k_0}}(y,x)\mu(dy).$$

$\delta_x Q_{n_{k_0}}^*$ is a measure absolutely continuous with respect to μ, and $\delta_x Q_{n_k}^*(X) > 1-\varepsilon$. Hence $\lim_{n\to\infty}\delta_x Q_{n_{k_0}}^* P^{*n-n_{k_0}}(A) = \mu(A)\cdot\delta_x Q_{n_{k_0}}^*(X)$, by Theorem 1. Therefore, for every n sufficiently large we have:

$$\mu(A)(1-2\varepsilon) \leqq \delta_x Q_{n_{k_0}}^* P^{*n-n_{k_0}}(A) \leqq \mu(A) + \varepsilon.$$

Hence

$$\mu(A)(1-2\varepsilon) \leqq \delta_x Q_{n_{k_0}}^* P^{*n-n_{k_0}}(A) \leqq P^{*n}1_A(x) \leqq \delta_x Q_{n_{k_0}}^* P^{*n-n_{k_0}}(A) + \varepsilon$$
$$\leqq \mu(A) + 2\varepsilon.$$

But ε is arbitrary, therefore $\lim_{n\to\infty} P^{*n}1_A(x) = \mu(A)$. The generalization for Σ_1 of order k is obvious.

THEOREM 5. *Let P, Σ_1, μ be as in Theorem 4. Let v be any finite measure supported on W, then*

$$\text{(4.6)} \qquad \left\| vP^{nk+r} - k\cdot vP^r(W)\cdot\mu\, I_{P^{k-r}W}\right\| \underset{n\to\infty}{\longrightarrow} 0$$

Proof. Let us first assume that Σ_1 is trivial.

(a) If $\nu \prec \mu$ and $f = \frac{d\nu}{d\mu}$, we shall prove:

$$fP^n \xrightarrow[n\to\infty]{L_1} \int f d\mu \tag{4.7}$$

what is equivalent to (4.6).

For every characteristic function 1_A we have, by Theorem 4, $\lim_{n\to\infty} 1_A P^n(x) = \lim_{n\to\infty} P^{*n} 1_A(x) = \mu(A)$ a.e. μ. By the dominated convergence theorem $1_A P^n(x) \xrightarrow[n\to\infty]{L_1} \mu(A)$. But the span of the set of characteristic functions is dense in $L_1(\mu)$, hence for every $f \in L_1(\mu)$ we have

$$fP^n \xrightarrow[n\to\infty]{L_1} \int f d\mu .$$

(b) Let ν be any finite measure.

Denote $\nu P^n = \tau_n + \sigma_n$ where $\tau_n \prec \mu$, $\sigma_n \perp \mu$. For each $\varepsilon > 0$, choose an integer n_0 such that

$$\sigma_{n_0}(X) < \varepsilon, \quad \tau_{n_0}(X) > \nu(X) - \varepsilon .$$

Hence

$$\begin{aligned} \| \nu P^n - \nu(X)\mu \| &= \| \tau_{n_0} P^{n-n_0} + \sigma_{n_0} P^{n-n_0} - \nu(X)\cdot\mu \| \leqq \| \tau_{n_0} P^{n-n_0} - \tau_{n_0}(X)\cdot\mu \| \\ &+ \| (\tau_{n_0}(X) - \nu(X))\mu \| + \| \sigma_{n_0} P^{n-n_0} \| \leqq \| \tau_{n_0} P^{n-n_0} - \tau_{n_0}(X)\cdot\mu \| + \\ &+ (\nu(X) - \tau_{n_0}(X)) \| \mu \| + \| \sigma_{n_0} \| \leqq \| \tau_{n_0} P^{n-n_0} - \tau_{n_0}(X)\cdot\mu \| + 2\varepsilon . \end{aligned}$$

By (4.7) we have $\| \tau_{n_0} P^{n-n_0} - \tau_{n_0}(X)\cdot\mu \| \underset{n\to\infty}{\to} 0$ and ε is arbitrary, therefore $\| \nu P^n - \nu(X)\cdot\mu \| \underset{n\to\infty}{\to} 0$. The generalization to the case where Σ_1 is of order k, is obvious.

Remark. Our theorem 5 was first proved by Orey in [8] Theorem 3.1. His proof was complicated. Another proof was given by Jamison and Orey in [7]. Their proof is by probabilistic methods. Our analytical proof seems more simple.

5. **Strong mixing in $L_1(\mu)$.** Consider the Markov process (X, Σ, μ, P) where $\mu P = \mu$, and μ is a probability measure.

Definition 4. (a) P is *strong mixing in* $L_1(\mu)$ if for every probability measure $\nu \prec \mu$,

$$\| \nu P^n - \mu \| \xrightarrow[n\to\infty]{} 0, \text{ or equivalently} \tag{5.1}$$

$$fP^n \xrightarrow[n\to\infty]{L_1} \int f d\mu \text{ for every } f \in L_1(\mu). \tag{5.2}$$

(b) P is strong mixing pointwise if for every $f \in L_\infty(\mu)$, $\lim_{n\to\infty} P^n f(x) = \int f du$ a.e. μ.

In §4 we saw that if P satisfies Harris' condition then P and P^* are strong mixing in $L_1(\mu)$ and pointwise. It is clear that a necessary condition to strong mixing in $L_1(\mu)$ is that Σ_1 is trivial. But this condition is not sufficient. Furthermore there is no symmetry between P and P^* with respect to this property, as we can see from the following example.

EXAMPLE. Consider the pointwise transformation on the unit interval $[0, 1]$, $Tx = 2x(\mathrm{mod}\, 1)$. It induces the operator

$$Pf(x) = \begin{cases} f(2x) & 0 \leqq x \leqq \frac{1}{2} \\ f(2x-1) & \frac{1}{2} < x \leqq 1. \end{cases}$$

A simple calculation shows that the adjoint of P is

$$P^*f(x) = \frac{f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right)}{2}.$$

We shall prove that the space K, defined in (3.1), contains only the constants, and hence Σ_1 is trivial.

It is easy to see that

$$P^*f^n(x) = \frac{1}{2^n} \sum_{k=0}^{2^n-1} f\left(\frac{x+k}{2^n}\right).$$

Let f be Riemann integrable. Then $P^{*n}f(x)$ is the Riemann sum, hence $P^{*n}f(x) \underset{n\to\infty}{\to} \int f\mu(dx)$ for all x. In particular if $f \perp 1$ then $P^{*n}f(x) \underset{n\to\infty}{\to} 0$. By the dominated convergence theorem, we have, for every function $f \perp 1$ that is bounded and Riemann-integrable, $\| P^{*n}f \| \underset{n\to\infty}{\to} 0$. But such functions are dense in $L_1(\mu)$. Hence $K = \{\text{const}\}$, and Σ_1 *is trivial*.

We shall now show that P^* *is not strong mixing in* $L_1(\mu)$. Let $f \in L_1(\mu)$ and $f \perp 1$. $fP^{*n} = P^n f$, but P is an isometry in $L_1(\mu)$, i.e. $\| P^n f \|_1 = \| f \|_1$, hence $fP^{*n} \underset{n\to\infty}{\overset{L_1}{\nrightarrow}} 0$, and P^* is not strong mixing in $L_1(\mu)$.

On the other hand, P *is strong mixing in* $L_1(\mu)$. Let $f \in L_1(\mu)$ and be bounded and Riemann-integrable. Then:

$$\lim_{n\to\infty} fP^n(x) = \lim_{n\to\infty} P^{*n}f(x) = \int f\mu(dx) \quad \text{for all } x.$$

By the dominated convergence theorem, $fP^n \underset{n\to\infty}{\overset{L_1}{\to}} \int f\mu(dx)$. But such functions are dense in $L_1(\mu)$. Hence, for every $f \in L_1(\mu), fP^n \underset{n\to\infty}{\overset{L_1}{\to}} \int f\mu(dx)$, and P is strong mixing in $L_1(\mu)$.

Appendix

Let (X, Σ, m, P) be a Markov process. m is a probability measure and $m \succ mP$. A is called an invariant set if $P1_A = 1_A$ and $m(A) > 0$. We denote Σ_i the collection of the invariant sets. If P is conservative then Σ_i is a σ-field.

Y. Ito proved, in [1], the following theorem:

A necessary and sufficient condition for the existence of a probability measure μ, so that $m \sim \mu$ and $\mu P = \mu$, is that for every A with $m(A) > 0$, we have

$$\text{(A.1)} \qquad \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=1}^{n} P^k(x, A) > 0 \text{ for every } x \in F \text{ where } m(F) > 0.$$

(F depends on A).

THEOREM. *If there is no probability measure μ so that $\mu \prec m$ and $\mu P = \mu$, then there is a decomposition*

$$\text{(A.2)} \qquad X = \bigcup_{j=1}^{\infty} X_j \text{ so that } \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} P^k(x, X_j) = 0 \quad \text{a.e. } m.$$

Proof. Assume that P is conservative (on the dissipative part the theorem is trivial). We assume that there is not a probability measure μ such that $\mu \prec m$ and $\mu P = \mu$. Hence, by Ito's theorem there is a set A, with $m(A) > 0$ and $\lim_{n\to\infty} 1/n \sum_{k=1}^{n} P^k(x, A) = 0$, a.e. m. Let us denote $A_n = \operatorname{supp} P^n(x, A)$, $\tilde{A} = \bigcup_{n=1}^{\infty} A_n$. It is known (see, for example, [2]) that $\tilde{A} \in \Sigma_i$. Let us also denote $A_n^i = \{x \mid P^n(x, A) \geqq 1/i\}$. Clearly $\bigcup_{i=1}^{\infty} A_n^i = A_n$. But

$$0 = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^{k+n}(x, A) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k P^n 1_A(x) \geqq$$

$$\geqq \frac{1}{i} \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} P^k(x, A_n^i), \left(P^n 1_A \leqq \frac{1}{i} 1_{A_i^n} \right).$$

Hence: $\lim_{N\to\infty} 1/N \sum_{k=1}^{N} P^k(x, A_n^i) = 0$ a.e. for all A_n^i, and for the invariant set $\tilde{A}$ there is such a decomposition $\tilde{A} = \bigcup_{i,n} A_n^i$. Consider $X - \tilde{A}$. It is an invariant set. Therefore we can consider the process on $X - \tilde{A}$, and find as before, $\tilde{\tilde{A}} \in X - \tilde{A}$ and $\tilde{\tilde{A}} = \bigcup_{j=1}^{\infty} A_j$ so that $\lim_{n\to\infty} 1/n \sum_{k=1}^{n} P^k(x, A_j) = 0$ for every j.

Let $\mathscr{F}$ be the collection of all sets A such that (i) $A \in \Sigma_i$, (ii) $A = \bigcup_{j=1}^{\infty} A_j$, $\lim_{n\to\infty} 1/n \sum_{k=1}^{n} P^k(x, A_j) = 0$ a.e. for every j. Let $\alpha = \sup_{A \in \mathscr{F}} m(A)$, we shall prove $\alpha = 1$. There is a sequence $\{A_i\} \subset \mathscr{F}$ so that $m(A_i) \nearrow \alpha$. It is clear that $\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$ and hence $m(\bigcup_{i=1}^{\infty} A_i) = \alpha$. If $\alpha < 1$ then $m(X - \bigcup_{i=1}^{\infty} A_i) > 0$ and clearly $X - \bigcup_{i=1}^{\infty} A_i \in \Sigma_i$ and we can consider the process on it and find, as before, $E \subset X - \bigcup_{i=1}^{\infty} A_i$ with $m(E) > 0$ so that $E = \bigcup_{j=1}^{\infty} E_j$ and $\lim_{n\to\infty} 1/n \sum_{k=1}^{n} P^k(x, E_j) = 0$ a.e. for all j. Hence $\bigcup_{i=1}^{\infty} A_i \cup E \in \mathscr{F}$ and $m(\bigcup_{i=1}^{\infty} A_i \cup E) > \alpha$. A contradiction. Hence $\alpha = 1$, and $X = \bigcup_{i=1}^{\infty} A_i$ and $A_i = \bigcup_{j=1}^{\infty} A_{ij}$ and $\lim_{n\to\infty} 1/n \sum_{k=1}^{n} P^k(x, A_{ij}) = 0$ a.e. for all j. So the theorem is proved.

REMARK. A theorem of this kind was proved by Dean and Sucheston in [1], Theorem 2. They proved that if there is *no* probabilistic measure μ so that $\mu \prec m$ and $\mu P = \mu$, then there is a decomposition $X = \bigcup_{j=1}^{\infty} X_j$ so that:

$$\lim_{n\to\infty} \sup_i \frac{1}{n} \sum_{k=1}^{n} mP^{k+i}(X_j) = 0 \quad \text{for all } j.$$

REFERENCES

1. D. W. Dean and L. Sucheston, *On invariant measures for operators*, Z. Wahrscheinlichkeitstheorie, **6** (1966), 1–9.

2. J. Feldman, *Subinvariant measures for Markov operators*, Duke Math. J. **22** (1962), 71–98.

3. S. R. Foguel, *The ergodic theorem for Markov processes*, Israel J. Math. **4** (1966), 11–22.

4. Y. Ito, *Invariant measures for Markov processes*, Trans. Amer. Math. Soc. **110** (1964), 152–184.

5. N. C. Jain, *Some limit theorems for a general Markov process*, Z. Wahrscheinlichkeitstheorie **6** (1966), 206–223.

6. N. C. Jain and B. Jamison, *Contributions to Doeblin's theory of Markov processes*, Z. Wahrscheinlichkeitstheorie **8** (1967), 19–40.

7. B. Jamison and S. Orey, *Markov chains recurrent in the sense of Harris*, Z. Wahrscheinlichtkeitstheorie **8** (1967), 41–48.

8. S. Orey, *Recurrent Markov chains*, Pacific J. Math. **9** (1959), 805–827.

THE HEBREW UNIVERSITY OF JERUSALEM