STRONG ERGODIC THEOREMS FOR MARKOV PROCESSES

S. HOROWITZ[1]

1. **Definitions and notation.** A Markov process is defined to be a quadruple (X, Σ, μ, P) where (X, Σ, μ) is a σ-finite measure space with positive measure μ and where P is an operator on $L_1(\mu)$ satisfying

(i) P is a contraction: $\|P\| \leqq 1$.

(ii) P is positive; if $0 \leqq u \in L_1(\mu)$ then $uP \geqq 0$. The operator adjoint to P is defined on $L_\infty(\mu)$. It will also be denoted by P but will be written to the left of its variable. Thus $\langle uP, f\rangle = \langle u, Pf\rangle$ for $u \in L_1(\mu)$, $f \in L_\infty(\mu)$.

The operator P acts on the space of the signed measures absolutely continuous with respect to μ in the following form:

$$\lambda P(A) = \int P1_A(x)\lambda(dx) \quad \text{where} \quad \lambda \prec \mu. \tag{1.1}$$

Equation (1.1) will occasionally be used for σ-finite positive measures.

The operator P is called *ergodic* if:

$$P1_A = 1_A \Rightarrow \mu(A) = 0 \quad \text{or} \quad \mu(A^c) = 0. \tag{1.2}$$

The operator P is said to be *conservative* if:

$$\mu(A) > 0 \Rightarrow \sum_{n=1}^{\infty} p^n 1_A(x) = \infty \quad \text{a.e.} \tag{1.3}$$

Throughout this paper, P is assumed to be an ergodic and conservative operator, and μ is assumed to be an invariant measure (i.e., $\mu P = \mu$).

It is well known that

$$\|Pf\|_1 \leqq \|P|f|\|_1 = \int P|f|\,\mu(dx) = \int |f|\,\mu P(dx)$$
$$= \int |f|\,\mu(dx) = \|f\|_1,$$

Received by the editors November 8, 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.

hence P and its adjoint are both contractions on $L_1(\mu)$. Thus by the Riesz Convexity Theorem the operator P and its adjoint are contractions on $L_p(\mu)$ for every $1 \leqq p \leqq \infty$.

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

$$K = \{f \mid f \in L_2(\mu), \|p^n f\|_2 = \|p^{*n} f\|_2 = \|f\|_2, \forall n\}, \tag{1.4}$$

$$\Sigma_1 = \text{the } \sigma\text{-field generated by sets } A \text{ with } l_A \in K. \tag{1.5}$$

In [1, Chapter VIII] 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 $f \in K$ if $f \in L_2(X, \Sigma, \mu)$ and is Σ_1 measurable.

(d) If $A \in \Sigma_1$ and $\mu(A) < \infty$ than $P^n 1_A$ and $p^{*n} 1_A$ are characteristic functions of sets in Σ_1.

2. **A strong ergodic theorem for Harris's processes.** Let $P^n = Q_n + R_n$ where Q_n is an integral operator with the kernel $q_n(x, y)$, and if K is any integral operator such that $0 \leqq K \leqq R_n$ then $K = 0$. (See [1, Chapter V]).

DEFINITION. *(X, Σ, μ, P) is said to be a Harris process if $Q_n > 0$ for some integer n.*

Throughout this section P is assumed to be a Harris process and μ is assumed to be a probability measure, that means $\mu(X) = 1$. Theorem D of [1, Chapter V] says that if P is a Harris process then Σ_1 is an atomic field. Let W be an atom then $\Sigma_1 = \{W \cup PW \cup \cdots \cup P^{d-1}W\}$ and $P^d W = W$, because of the assumption that P is ergodic and conservative. The integer d is called the order of the process.

Let S be the projection on K, defined in (1.4):

$$Sf = E(f/\Sigma_1). \tag{2.1}$$

Since K is invariant under P and P^* and $P^d f = P^{*d} f = f$ for all $f \in K$:

$$PS = SP, \qquad P^d S = SP^d = S. \tag{2.2}$$

Hence for all $f \in L_p(\mu)$ we have $P^n(I - S)f$ converges weakly to 0 for all $1 \leqq p \leqq \infty$, because $L_p(\mu) \subset L_2(\mu)$ if $2 \leqq p \leqq \infty$, and $L_2(\mu)$ is dense in $L_p(\mu)$ if $1 \leqq p \leqq 2$ ($(I - S)f$ is orthogonal to K). By Theorem E, Chapter V of [1] we have:

$$R_n 1(x) \xrightarrow[n\to\infty]{} 0 \quad \text{and} \quad R_n^* 1(x) \xrightarrow[n\to\infty]{} 0 \quad \text{a.e.} \tag{2.3}$$

Let us consider P as a mapping from $L_p(\mu)$ into $L_q(\mu)$ $(p \geqq q)$. We denote by $\|\cdot\|_{p,q}$ the norm of this operator.

THEOREM 2.1. *For all* $1 \leqq q < p \leqq \infty$ *we have*

$$\|P^n - P^n S\|_{p,q} \xrightarrow[n\to\infty]{} 0 \tag{2.4}$$

or equivalently (*by* (2.2)):

$$\|P^{nd+r} - SP^r\|_{p,q} \xrightarrow[n\to\infty]{} 0. \tag{2.5}$$

We shall first prove the following lemmas:

LEMMA 2.1. *For all* $1 \leqq q < p \leqq \infty$ *we have*

$$\|R_n\|_{p,q} \xrightarrow[n\to\infty]{} 0. \tag{2.6}$$

PROOF. Let us first assume that $p = \infty$, and then for all $\|f\|_\infty \leqq 1$ we have, for any $1 \leqq q < \infty$:

$$\|R_n f\|_q = \left[\int |R_n f|^q \mu(dx)\right]^{1/q} \leqq \left[\int (R_n |f|)^q \mu(dx)\right]^{1/q}$$
$$\leqq \left[\int (R_n 1)^q \mu(dx)\right]^{1/q} \xrightarrow[n\to\infty]{} 0$$

by the Dominated Convergence Theorem. Hence

$$\|R_n\|_{\infty,q} \xrightarrow[n\to\infty]{} 0.$$

In a similar way, it can be shown: $\|R_n^*\|_{\infty,q} \to 0$. Now let us assume $q = 1$, $1 < p < \infty$. Given a function f with $\|f\|_p \leqq 1$, let us define $g_n(x) = \operatorname{sgn} R_n f(x)$, we have $R_n f \cdot g_n = |R_n f|$. Hence,

$$\|R_n f\|_1 = \int R_n f \cdot g_n \mu(dx) = \int f \cdot R_n^* g_n \mu(dx) \leqq \|f\|_p \cdot \|R_n^* g_n\|_r$$

where $1/p + 1/r = 1$. But $\|g_n\|_\infty = 1$, hence $\|R_n^* g_n\|_r \leqq \|R_n^*\|_{\infty,r}$, and by the former result we have:

$$\|R_n\|_{p,1} = \sup_{\|f\|_p \leqq 1} \|R_n f\|_1 \leqq \|R_n^*\|_{\infty,r} \xrightarrow[n\to\infty]{} 0.$$

By the Riesz Convexity Theorem we have: $\|R_n\|_{p,q} \to_{n\to\infty} 0$ for all $1 \leqq q < p \leqq \infty$.

LEMMA 2.2. *Let T be the operator T from $L_p(\mu)$ into $L_q(\mu)$ $(\mu(X) = 1)$, $1 \leqq q < p \leqq \infty$ given by $Tf(x) = \int K(x,y) f(y)\mu(dy)$, such that $\int K(x,y)\mu(dx) \leqq 1$ and $\int K(x,y)\mu(dy) \leqq 1$ then T is a compact operator.*

PROOF. Let us first assume $p=\infty$, $q=1$. Let $\{f_n\}\subset L_\infty(\mu)$ ($\|f_n\|_\infty \leqq 1$) a sequence that weakly* converges to 0. It is given that $K(x, \cdot) \in L_1(\mu)$, hence for each x:

$$1 \geqq \left| \int K(x, y) f_n(y)\mu(dy) \right| \xrightarrow[n\to\infty]{} 0,$$

and by the Dominated Convergence Theorem we have

$$\|Tf_n\|_1 \xrightarrow[n\to\infty]{} 0,$$

hence T is a compact operator from $L_\infty(\mu)$ into $L_1(\mu)$.

Let us denote $B_\infty = \{f \mid \|f\|_\infty \leqq 1\}$, $\mathrm{Cl}(TB_\infty)$ is a compact set in $L_1(\mu)$ but a sequence of functions that is uniformly bounded converges in $L_q(\mu)$ ($1 \leqq q < \infty$) if and only if it converges in measure, hence $\mathrm{Ce}(TB_\infty)$ is a compact set in $L_q(\mu)$ for all $1 \leqq q < \infty$, and T is a compact operator from $L_\infty(\mu)$ into $L_q(\mu)$. In a similar way it can be shown that also T^* is a compact operator from $L_\infty(\mu)$ into $L_q(\mu)$ ($1 \leqq q < \infty$). Now let us assume $q=1$, $1<p<\infty$. Let $\{f_n\}\subset L_p(\mu)$ be a sequence that converges weakly to 0 in $L_p(\mu)$. Let us define the sequence $g_n(x) = \mathrm{sgn}\ Tf_n(x)$, then we have

$$\|Tf_n\|_1 = \int Tf_n \cdot g_n \mu(dx) = \int f_n \cdot T^* g_n \mu(dx).$$

By the former result $\{T^*g_n\}$ is a compact set in $L_r(\mu)$, where $1/r + 1/p = 1$, because $\|g_n\|_\infty = 1$. There is no loss of generality in assuming that the sequence $\{T^*g_n\}$ converges in $L_r(\mu)$, say to g. Hence

$$\begin{aligned}\|Tf_n\|_1 &= \int f_n \cdot g\mu(dx) + \int f_n(T^*g_n - g)\mu(dx) \\ &\leqq \langle f_n, g\rangle + \|f_n\|_p \cdot \|T^*g_n - g\|_r\end{aligned}$$

and both expressions of the right side tend to 0, the first because $\{f_n\}$ converges weakly to 0 in $L_p(\mu)$, and the second because

$$\|T^*g_n - g\|_r \xrightarrow[n\to\infty]{} 0$$

and $\|f_n\|_p$ is bounded. Hence T is a compact operator from $L_\infty(\mu)$ into $L_q(\mu)$ for all $1 \leqq q < \infty$ and from $L_p(\mu)$ into $L_1(\mu)$ for all $1 < p \leqq \infty$. Hence by the Lions-Peetre Interpolation Theorem [**2**, Chapter V]; T is a compact operator from $L_p(\mu)$ into $L_q(\mu)$ for all $1 \leqq q < p \leqq \infty$.

PROOF OF THE THEOREM. The operator Q_n fulfills the conditions of Lemma 2, hence it is a compact operator from $L_p(\mu)$ into $L_q(\mu)$ for all $1 \leqq q < p \leqq \infty$. The sequence $(P^n - P^nS)f$ converges weakly to 0 in $L_p(\mu)$ for each $f \in L_p(\mu)$ and for all $1 \leqq p < \infty$. According to Lemma 2.1, for all $\epsilon > 0$, we have for sufficiently large k, $\|R_k\|_{\infty,q} \leqq \epsilon$ where $1 \leqq q < \infty$. Hence for each $f \in L_\infty(\mu)$: $\|P^k(P^n - P^nS)f\|_q \leqq \|Q_k(P^n - P^nS)f\|_q + \|R_k(P^n - P^nS)f\|_q$. The first expression in right side tends to zero, because Q_k is a compact operator from $L_\infty(\mu)$ to $L_q(\mu)$, while the second is less than $\epsilon \cdot \|f\|_q$. Hence

$$\|(P^n - P^nS)f\|_q \xrightarrow[n\to\infty]{} 0 \qquad \text{for all } 1 \leqq q < \infty. \tag{2.6}$$

But $L_\infty(\mu)$ is dense in $L_q(\mu)$, hence (2.6) is true for each $f \in L_q(\mu)$. On the other hand we have

$$\|(P^n - P^nS)P^kf\|_q \leqq \|R_kf\|_q + \|(P^n - P^nS)Q_kf\|_q.$$

For all $1 \leqq q < p \leqq \infty$, we have for all $\epsilon > 0$ and for sufficiently large k, $\|R_k\|_{p,q} \leqq \epsilon$.

Let us denote $B_p = \{f \mid \|f\|_p \leqq 1\}$. Hence for each $f \in B_p$ we have $\|(P^n - P^nS)P^kf\|_q \leqq \|(P^n - P^nS)Q_kf\|_q + \epsilon$.
But $\mathrm{Cl}(Q_kB_p)$ is a compact set is $L_q(\mu)$, hence $\|P^n(I-S)g\|_q$ where $g = Q_kf$ is a monotone sequence of real functions defined on the compact set $\mathrm{Cl}(Q_kB_p)$, and converges to 0.

By Dini's Theorem this sequence converges uniformly; that means, for every $\epsilon > 0$ we have for sufficiently large n for all $f \in B_p$: $\|P^n - P^nS)Q_kf\| \leqq \epsilon$.

Hence for n and k sufficiently large we have for all $f \in B_p$: $\|(P^{n+k} - P^{n+k}S)f\|_q \leqq 2\epsilon$. In other words,

$$\|P^n - P^nS\|_{p,q} \xrightarrow[n\to\infty]{} 0.$$

3. **Strong ergodic theorems for special cases.** Let (X, Σ, μ, P) be any Markov process as defined in § 1. Let us define:

$$\Sigma_d = \{P^n 1_A = 1_{B_n}, \ \forall_n\}. \tag{3.1}$$

It can be shown (see [1]) that Σ_d is a field, and in the case that μ is finite we have

$$\Sigma_1 = \bigcap_{n=\infty}^{\infty} P^n\Sigma_d. \tag{3.2}$$

In the case that μ is infinite and $\bigcap_{n=1}^\infty P^n\Sigma_d$ is an atomic field, $\Sigma_1 = \varnothing$. If the operator P is induced to be the pointwise transformation T then $\Sigma_d = \Sigma$, denote $\Sigma_n = T^{-n}\Sigma$ and then $\bigcap_{n=0}^\infty P^n\Sigma_d = \bigcap_{n=0}^\infty \Sigma_n$.

THEOREM 3.1. *Let P be induced by the pointwise transformation T and $\Sigma_1=\{X, \varnothing\}$ (and hence μ is finite, assume $\mu(X)=1$), let $\nu\prec\mu$ be any probability measure, then*

$$\|\nu P^n - \mu\| \xrightarrow[n\to\infty]{} 0. \tag{3.3}$$

PROOF. Let $u=d\nu/d\mu$. Let us first assume that the function u is bounded. Then we have

$$\|\nu P^n - \mu\| = \sup_{A\in\Sigma} | \nu(T^{-n}A) - \mu(A) | ,$$

but

$$\begin{aligned} | \nu(T^{-n}A) - \mu(A) | &= \left| \int_{T^{-n}A} (u - 1)\mu(dx) \right| \\ &= \left| \int_{T^{-n}A} (E(u/\Sigma_n) - 1)\mu(dx) \right| \\ &\leqq \int | E(u/\Sigma_n) - 1 | \mu(dx) \quad (\mu(T^{-n}A) = \mu(A)). \end{aligned}$$

By the Backward Martingale Convergence Theorem we have

$$E(\mu/\Sigma_n) = u\mu(dx) = 1,$$

and by the Dominated Convergence Theorem

$$\int | E(\mu/\Sigma_n) - 1 | \mu(dx) \xrightarrow[n\to\infty]{} 0.$$

But the set of measures with bounded Radon-Nikodym derivatives is dense in the space of the measures absolutely continuous with respect to μ. Hence for every probability measure $\nu\prec\mu$ we have

$$\|\nu P^n - \mu\| \xrightarrow[n\to\infty]{} 0.$$

For the following theorem, let K be as in (1.4). μ may be finite or infinite.

THEOREM 3.2. *Let P be a normal operator on $L_2(\mu)$. Then for each $f\perp K$ we have*

$$\|P^n f\|_2 \to 0. \tag{3.4}$$

PROOF. Let $E(\cdot)$ be the spectral measure with respect to P. Denote $C=\{\lambda | |\lambda| =1\}$, $D=\{\lambda| |\lambda| <1\}$. It is clear that $E(C)L_2(\mu)=K$,

$E(D)L_2(\mu)=K^{\perp}$. Hence if $f\in K^{\perp}$ then

$$\|P^n f\|_2^2=\int_D |\lambda|^n \|E(d\lambda)f\|^2 \xrightarrow[n\to\infty]{} 0.$$

EXAMPLE. Let T_1 and T_2 two pointwise invertible measure preserving transformations so that $T_1T_2=T_2T_1$, and for each set A with $\mu(A)>0$ and $\mu(A^C)>0$ we have $T_1^{-1}A\neq T_2^{-1}A$. (For example $T_1=I$ and T_2 is any invertible ergodic transformation.) Define $P=\alpha T_1+\beta T_2$ where $0<\alpha,\beta<1$ and $\alpha+\beta=1$. If $P1_A=1_B$ then $T_1^{-1}A=T_2^{-1}A=B$ and $\mu(A)=0$ or $\mu(A^C)=0$. Hence Σ_d is trivial. Therefore if $\mu(X)=\infty$ then $K=\{0\}$ and if $\mu(X)=1$ then $K=\{\text{const}\}$. But P is a normal operator, and by Theorem 3.2 we have for each $f\in L_2(\mu)$, $\|P^n f\|_2\to_{n\to\infty} 0$ in the former case and

$$P^n f \xrightarrow[n\to\infty]{L_2} \int f\mu(dx)$$

in the later case.

REMARK. Theorem 3.2 in the case that the measure μ is finite, was proved, by different methods, in [3]. The proof fails in the case that the measure μ is infinite.

REFERENCES

1. S. R. Foguel, *The ergodic theory of Markov processes*, Van Nostrand, New York (to appear).

2. J. J. Lions and J. Peetre, *Sur une class d'espaces d'interpolations*, Inst. Hautes Études Sci. Publ. Math. **19** (1964), 5–68.

3. M. Rosenblatt. *Transition probability operators*, Fifth Berkeley Sympos. Univ. of California, Berkeley, Calif. Vol. II, 1966, pp. 473–483.

HEBREW UNIVERSITY OF JERUSALEM