J. Math. Soc. Japan
Vol. 25, No. 2, 1973

Ergodic theorems and weak mixing for Markov processes

By Ryotaro SATO

(Received Feb. 9, 1972)

§ 1. Definitions and notation.

A Markov process is defined to be a quadruple $(\Omega, \mathcal{B}, m, P)$ where $(\Omega, \mathcal{B}, m)$ is a σ-finite measure space with positive measure m and where P is a positive linear contraction on $L^1(\Omega)$. P will be written to the right of its variable, and the adjoint in $L^\infty(\Omega)$ 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 $u \in L^1(\Omega)$ and $f \in L^\infty(\Omega)$. A σ-finite positive measure λ on $(\Omega, \mathcal{B})$ absolutely continuous with respect to m is called *subinvariant* if $\int P1_A(\omega)\lambda(d\omega) \leqq \lambda(A)$ for any $A \in \mathcal{B}$ and *invariant* if $\int P1_A(\omega)\lambda(d\omega) = \lambda(A)$ for any $A \in \mathcal{B}$. Throughout this paper m is assumed to be either an infinite subinvariant measure or a finite invariant measure. It is well known that P on $L^\infty(\Omega)$ is also a linear contraction on $L^1(\Omega)$ and hence it may be considered to be a linear contraction on each $L^p(\Omega)$ with $1 \leqq p \leqq \infty$ by the Riesz convexity theorem. The adjoint process of $(\Omega, \mathcal{B}, m, P)$ will be denoted by $(\Omega, \mathcal{B}, m, P^*)$; its properties are studied in [4, Chapter VII].

The process $(\Omega, \mathcal{B}, m, P)$ is called

1) *ergodic*, if $P1_A = 1_A$ implies $m(A) = 0$ or $m(\Omega - A) = 0$;

2) *weakly mixing*, if

$$L^2(\Omega) \ominus \left\{ f \in L^2(\Omega)\,;\ \lim_n \frac{1}{n} \sum_{i=0}^{n-1} |\langle P^i f, f\rangle| = 0 \right\}$$

is so small as to contain nothing more than the constant functions;

3) *strongly mixing*, if

$$L^2(\Omega) \ominus \{ f \in L^2(\Omega)\,;\ \lim_n \langle P^n f, f\rangle = 0 \}$$

is so small as to contain nothing more than the constant functions.

We note that our definition of strong mixing is due to Foguel [4] and coincides with the notion of "mixing" proposed by Lin [7].

§ 2. Results.

THEOREM 1. *If* $1 \leqq p < \infty$, $f \in L^p(\Omega)$, *and* $k_1, k_2, \cdots$ *is a uniform sequence (in the sense of Brunel and Keane* [2]*), then the limit*

$$\tilde{f}(\omega) = \lim_n \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f(\omega) \tag{1}$$

exists and is finite almost everywhere. In particular, if $1 < p < \infty$ *then*

$$\lim_n \left\| \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f - \tilde{f} \right\|_p = 0 . \tag{2}$$

THEOREM 2. *The process* $(\Omega, \mathcal{B}, m, P)$ *is weakly mixing if and only if for any* $f \in L^1(\Omega)$ *and any uniform sequence* $k_1, k_2, \cdots$,

$$\tilde{f}(\omega) = \frac{1}{m(\Omega)} \int f \, dm \qquad \textit{almost everywhere.} \tag{3}$$

To prove the above theorems, we require the following lemmas. The first lemma is an extension of [9, Theorem 1] to Markov processes.

LEMMA 1. *If* m *is an infinite subinvariant measure then* Ω *is decomposed into three disjoint measurable sets* Ω_0, Ω_+ *and* Ω_{++} *such that*

i) $P1_{\Omega_0} \leqq 1_{\Omega_0}$ *and for any* $A \in \mathcal{B}$ *with* $A \subset \Omega_0$ *and* $m(A) < \infty$,

$$\lim_n \langle P^n 1_A, 1_A \rangle = 0 ;$$

ii) $P1_{\Omega_+} = 1_{\Omega_+}$ *and for any* $A \in \mathcal{B}$ *with* $A \subset \Omega_+$ *and* $0 < m(A) < \infty$,

$$\limsup_n \langle P^n 1_A, 1_A \rangle \neq 0$$

but

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P^i 1_A, 1_A \rangle = 0 ;$$

iii) $P1_{\Omega_{++}} = 1_{\Omega_{++}}$, Ω_{++} *is a union of countably many sets* $A_n \in \mathcal{B}$ *with* $m(A_n) < \infty$ *and* $P1_{A_n} = 1_{A_n}$, *and for any* $A \in \mathcal{B}$ *with* $A \subset \Omega_{++}$ *and* $m(A) > 0$,

$$\limsup_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P^i 1_A, 1_A \rangle \neq 0 .$$

PROOF. Let $J = \{f \in L^2(\Omega) ; Pf = f\}$ and $K = \{f \in L^2(\Omega) ;\ \|P^n f\|_2 = \|P^{*n} f\|_2 = \|f\|_2$ for $n = 1, 2, \cdots\}$, and define

$$\Omega_{++} = \text{ess sup} \{A \in \mathcal{B} ;\ 1_A \in J\} ,$$

$$M = \text{ess sup} \{A \in \mathcal{B} ;\ 1_A \in K\} .$$

Then $P1_{\Omega_{++}} = P^* 1_{\Omega_{++}} = 1_{\Omega_{++}}$, $P1_M = P^* 1_M = 1_M$, and for any $A \in \mathcal{B}$ with $1_A \in K$, $P1_A$ and $P^* 1_A$ are characteristic functions of sets and $PP^* 1_A = P^* P 1_A = 1_A$;

moreover J and K are generated by $\{1_A; 1_A \in J\}$ and $\{1_A; 1_A \in K\}$, respectively (cf. [**4**, pp. 87-88]). Therefore a slightly modified argument of [**6**, p. 155] shows that $M-\Omega_{++}$ is decomposed into two disjoint measurable sets Ω_+ and M_0 such that

a) $P1_{\Omega_+} = P^*1_{\Omega_+} = 1_{\Omega_+}$ and $P1_{M_0} = P^*1_{M_0} = 1_{M_0}$;

b) for any $A \in \mathcal{B}$ with $A \subset \Omega_+$ and $0 \neq 1_A \in K$,

$$\limsup_n \langle P^n 1_A, 1_A \rangle \neq 0$$

but

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P^i 1_A, 1_A \rangle = 0$$

(the last equality follows from the fact that $1_A \perp J$ and hence $\lim_n \|1/n \sum_{i=0}^{n-1} P^i 1_A\|_2 = 0$ by the mean ergodic theorem (cf. [**11**, pp. 213-214]));

c) for any $A \in \mathcal{B}$ with $A \subset M_0$ and $1_A \in K$,

$$\lim_n \langle P^n 1_A, 1_A \rangle = 0 .$$

Define $\Omega_0 = \Omega - (\Omega_+ \cup \Omega_{++})$. Then it may be readily seen that Ω_0, Ω_+ and Ω_{++} are the desired decomposition of Ω. The proof is complete.

Let $k_1, k_2, \cdots$ be a uniform sequence, and let $(X, \mathcal{X}, \mu, \varphi)$ and y, Y be the apparatus connected with this sequence. Φ will denote the operator on $L^p(X)$, $1 \leqq p \leqq \infty$, induced by φ. Taking $(\Omega', \mathcal{B}', m')$ to be the direct product of $(\Omega, \mathcal{B}, m)$ and $(X, \mathcal{X}, \mu)$ and P' the direct product of P and Φ, it follows easily that P' is a positive linear contraction on each $L^p(\Omega')$ with $1 \leqq p \leqq \infty$.

LEMMA 2. *Let m be a finite invariant measure. If the process $(\Omega, \mathcal{B}, m, P)$ is ergodic and if P and Φ have no common eigenvalues other than* 1 *as operators on $L^2(\Omega)$ and $L^2(X)$, respectively, then the process $(\Omega', \mathcal{B}', m', P')$ is ergodic.*

PROOF. Without loss of generality it may be assumed that $m(\Omega) = 1$. Let $f \in L^2(\Omega)$ and $g \in L^2(X)$. If $f \perp \{h \in L^2(\Omega); Ph = ch$ for some constant c with $|c| = 1\}$ and $\langle g, 1 \rangle = 0$, then

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} |\langle P^i f, f \rangle| = 0$$

and

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle \Phi^i g, g \rangle = 0 .$$

Hence

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P^i f, f \rangle \langle \Phi^i g, g \rangle = 0 . \tag{4}$$

Similarly, if $\langle f, 1 \rangle = 0$ and $g \perp \{h \in L^2(X); \Phi h = dh$ for some constant d with $|d| = 1\}$, then (4) holds. Next suppose that $Pf = cf$, $\Phi g = dg$, $|c| = |d| = 1$ and

$c \neq d$. Then, since $cd \neq 1$ by hypothesis, (4) holds also. Thus an approximation argument shows that for $f \in L^2(\Omega)$ and $g \in L^2(X)$,

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P^i f, f \rangle \langle \Phi^i g, g \rangle = \langle f, 1 \rangle \langle 1, g \rangle ,$$

from which it follows that if $f', g' \in L^2(\Omega')$ then

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \langle P'^i f', g' \rangle = \langle f', 1 \rangle \langle 1, g' \rangle .$$

This completes the proof of the present lemma.

We are now in a position to prove the above theorems.

PROOF OF THEOREM 1. Since $\|P\|_\infty \leqq 1$, the first half of the theorem follows from [8, Theorem 1]. Hence we prove here the second half. The method of proof is somewhat similar to that given in [9]. By Lemma 1, it is sufficient to prove that if $f \in L^p(\Omega)$ with $1 < p < \infty$ is supported on $\Omega_0 \cup \Omega_+$ then

$$\lim_n \left\| \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f \right\|_p = 0 \tag{5}$$

for any uniform sequence $k_1, k_2, \cdots$. An approximation argument then shows that it is sufficient to prove (5) for $f = 1_A$ with $A \subset \Omega_0 \cup \Omega_+$ and $m(A) < \infty$. Since $\|P\|_\infty \leqq 1$, it is also sufficient to consider the case $1 < p < 2$. It follows from Lemma 1 that there exists a subset S of the non-negative integers having density zero such that if n is restricted to be outside S, then $\lim_n \langle P^n 1_A, 1_A \rangle = 0$. But, since $\lim_n (\langle P^{n+k} 1_A, P^k 1_A \rangle - \langle P^n 1_A, 1_A \rangle) = 0$ uniformly in k (cf. [4, p. 86]), it follows that for any given $\varepsilon > 0$ there exists a positive integer $N(\varepsilon)$ such that if $n > N(\varepsilon)$ and $n \notin S$ then

$$\langle P^{n+k} 1_A, P^k 1_A \rangle < \varepsilon \qquad \text{for } k = 0, 1, \cdots . \tag{6}$$

Define $D(k, N(\varepsilon)) = \{ j \geqq 0 \,;\, |k-j| \leqq N(\varepsilon) \}$, $\delta = p-1$, and

$$\begin{cases} a_{n,k} = 1/n & \text{if } k = k_i \text{ for some } 1 \leqq i \leqq n , \\ a_{n,k} = 0 & \text{otherwise} . \end{cases}$$

Since $0 < \delta < 1$, it follows that

$$\begin{aligned} \left\| \frac{1}{n} \sum_{i=1}^{n} P^{k_i} 1_A \right\|_p^p &= \int \Big| \sum_{k=0}^{\infty} a_{n,k} P^k 1_A \Big|^p dm \\ &\leqq \sum_{k=0}^{\infty} a_{n,k} \int P^k 1_A \Big(\sum_{j \in D(k, N(\varepsilon))} a_{n,j} P^j 1_A \Big)^\delta dm \\ &\quad + \sum_{k=0}^{\infty} a_{n,k} \int P^k 1_A \Big(\sum_{\substack{j \notin D(k, N(\varepsilon)) \\ |k-j| \in S}} a_{n,j} P^j 1_A \Big)^\delta dm \\ &\quad + \sum_{k=0}^{\infty} a_{n,k} \int P^k 1_A \Big(\sum_{\substack{j \notin D(k, N(\varepsilon)) \\ |k-j| \notin S}} a_{n,j} P^j 1_A \Big)^\delta dm \end{aligned}$$

$$=\mathrm{I}(n)+\mathrm{II}(n)+\mathrm{III}(n)\,.$$

It follows easily that $\lim_n \mathrm{I}(n)=0$. Next we estimate $\mathrm{II}(n)$. It is clear that

$$\mathrm{II}(n)\leqq \sum_{k=0}^{\infty} a_{n,k}\int P^k 1_A\Big(\sum_{|k-j|\in S} a_{n,j}P^j 1_A\Big)^{\delta} dm$$

$$\leqq \sum_{k=0}^{\infty} a_{n,k}\Big(\sum_{j\in S} a_{n,k+j}\Big)^{\delta} m(A)$$

$$+\Big(\sum_{\substack{j\in S\\ j\leqq k_n}} a_{n,k_n-j}\Big)^{\delta} m(A)\,.$$

It follows from a slightly modified argument of [**5**, pp. 146-147] that

$$\lim_n \sum_{k=0}^{\infty} a_{n,k}\Big(\sum_{j\in S} a_{n,k+j}\Big)^{\delta} m(A)=0\,.$$

On the other hand,

$$\sum_{\substack{j\in S\\ j\leqq k_n}} a_{n,k_n-j}\leqq \frac{k_n}{n}\,\frac{|\{j\in S;\, j\leqq k_n\}|}{k_n}\longrightarrow 0$$

as $n\to\infty$, since the k_n/n are bounded (see [**2**]) and S has density zero. Hence $\lim_n \mathrm{II}(n)=0$. Therefore in order to complete the proof it is sufficient to prove that $\mathrm{III}(n)$ can be arbitrarily small for all n. To see this, let for any $\varepsilon_1>0$,

$$G(n, k\,;\varepsilon_1)=\{\omega\in\Omega\,;\sum_{\substack{j\notin D(k, N(\varepsilon))\\ |k-j|\notin S}} a_{n,j}P^j 1_A(\omega)>\varepsilon_1\}\,.$$

Then (6) implies that

$$\langle P^k 1_A,\, 1_{G(n,k\,;\,\varepsilon_1)}\rangle\leqq \frac{1}{\varepsilon_1}\sum_{\substack{j\notin D(k, N(\varepsilon))\\ |k-j|\notin S}} a_{n,j}\langle P^k 1_A,\, P^j 1_A\rangle\leqq \frac{\varepsilon}{\varepsilon_1}\,.$$

Thus

$$\mathrm{III}(n)\leqq \sum_{k=0}^{\infty} a_{n,k}(\langle P^k 1_A,\, 1_{G(n,k\,;\,\varepsilon_1)}\rangle+\varepsilon_1^{\delta}\, m(A))$$

$$\leqq \frac{\varepsilon}{\varepsilon_1}+\varepsilon_1^{\delta}\, m(A)\,.$$

Since the right hand side of the last inequality can be arbitrarily small, this completes the proof of Theorem 1.

PROOF OF THEOREM 2. Case I. Suppose m is a finite invariant measure. If $(\Omega, \mathcal{B}, m, P)$ is weakly mixing then, clearly, it is ergodic. Hence Lemma 2 implies that the Markov process $(\Omega', \mathcal{B}', m', P')$ is ergodic. Thus a slightly modified argument of [**2**, p. 236] (see also the proof of [**8**, Theorem 1]) is sufficient for the proof of (3), and hence we omit the details. If $(\Omega, \mathcal{B}, m, P)$ is not weakly mixing, then there exists a function $f\in L^2(\Omega)$ such that $f\neq 0$, $\langle f, 1\rangle=0$ and $Pf=cf$ for some constant c with $|c|=1$. Define, as in [**10**], a uniform sequence $k_1, k_2, \cdots$ by the following way:

$$k_1 = \min\{j \geqq 1;\ -\pi/4 < \arg(c^j) < \pi/4\},$$

$$k_n = \min\{j > k_{n-1};\ -\pi/4 < \arg(c^j) < \pi/4\}.$$

It is then clear that $\tilde{f}(\omega) = \lim_n \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f(\omega)$ is not a constant function.

Case II. Suppose m is an infinite subinvariant measure. If $(\Omega, \mathcal{B}, m, P)$ is weakly mixing, it follows easily that for any $f \in L^1(\Omega)$,

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} P^i f(\omega) = 0 \qquad \text{almost everywhere}.$$

Thus clearly (3) holds for any uniform sequence $k_1, k_2, \cdots$ and any $f \in L^1(\Omega)$. If $(\Omega, \mathcal{B}, m, P)$ is not weakly mixing, it follows from Lemma 1 that there exists a measurable set A with $0 < m(A) < \infty$ and $P1_A = 1_A$ almost everywhere. Hence for any uniform sequence $k_1, k_2, \cdots$,

$$\lim_n \frac{1}{n} \sum_{i=1}^{n} P^{k_i} 1_A(\omega) = 1_A(\omega) \qquad \text{almost everywhere}.$$

This completes the proof of Theorem 2.

From the proof of Theorem 2 we have the following result (cf. [**2**, Corollary]).

THEOREM 3. *Let $(\Omega, \mathcal{B}, m, P)$ and $k_1, k_2, \cdots$ be as in Lemma 2. Then for any $f \in L^1(\Omega)$,*

$$\tilde{f}(\omega) = \frac{1}{m(\Omega)} \int f\, dm \qquad \textit{almost everywhere}.$$

It is known [**7**] that i) if m is a finite invariant measure then $(\Omega, \mathcal{B}, m, P)$ is strongly mixing if and only if for any p with $1 \leqq p < \infty$, any $f \in L^p(\Omega)$ and any strictly increasing sequence $k_1, k_2, \cdots$ of non-negative integers,

$$\lim_n \left\| \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f - \frac{1}{m(\Omega)} \int f\, dm \right\|_p = 0; \tag{7}$$

ii) if m is an infinite subinvariant measure then $(\Omega, \mathcal{B}, m, P)$ is strongly mixing if and only if for any p with $1 < p < \infty$, any $f \in L^p(\Omega)$ and any strictly increasing sequence $k_1, k_2, \cdots$ of non-negative integers,

$$\lim_n \left\| \frac{1}{n} \sum_{i=1}^{n} P^{k_i} f \right\|_p = 0. \tag{8}$$

Under the same direction, we have the following

THEOREM 4. a) *If m is a finite invariant measure then $(\Omega, \mathcal{B}, m, P)$ is weakly mixing if and only if for any p with $1 \leqq p < \infty$, any $f \in L^p(\Omega)$, and any strictly increasing sequence $k_1, k_2, \cdots$ of non-negative integers such that the k_n/n*

are bounded, (7) *holds.*

b) *If* m *is an infinite subinvariant measure then* $(\Omega, \mathcal{B}, m, P)$ *is weakly mixing if and only if for any* p *with* $1<p<\infty$, *any* $f \in L^p(\Omega)$, *and any strictly increasing sequence* $k_1, k_2, \cdots$ *of non-negative integers such that the* k_n/n *are bounded,* (8) *holds.*

PROOF. Arguments analogous to those given in the above for the proofs of Theorems 1 and 2 are sufficient, and hence we omit the details.

REMARK 1. If m is a finite invariant measure and if $(\Omega, \mathcal{B}, m, P)$ is ergodic, then the following statements are equivalent:

a) $(\Omega, \mathcal{B}, m, P)$ is weakly mixing.

b) For any (1, 2)-sequence $k_1, k_2, \cdots$ (for the definition, see [**1**]) with lower density greater than 1/2 and $A \in \mathcal{B}$ with $m(A)>0$,

$$m(\{\omega \in \Omega ; \sum_{n=1}^{\infty} P^{k_n} 1_A(\omega)>0\})=m(\Omega).$$

c) For any $A \in \mathcal{B}$ with $m(A)>0$, the upper density of the set $\{n \geqq 1 ; \langle P^n 1_A, 1_A \rangle \neq 0\}$ is greater than 1/2.

This follows from arguments analogous to those given in [**1**] and [**3**], and hence we omit the details.

REMARK 2. If m is a finite invariant measure, then the following statements are equivalent:

a) $(\Omega, \mathcal{B}, m, P)$ is strongly mixing.

b) For any $f \in L^1(\Omega)$ and any strictly increasing sequence $k_1, k_2, \cdots$ of positive integers, there exists a decreasing sequence $c_1, c_2, \cdots$ of positive reals such that $\sum_{n=1}^{\infty} c_n$ diverges and

$$\lim_n \Big(\sum_{i=1}^{n} c_i P^{k_i} f(\omega) / \sum_{i=1}^{n} c_i \Big) = (1/m(\Omega)) \int f \, dm$$

almost everywhere.

This follows from a slightly modified argument of [**2**, pp. 238-239], and hence the proof is also omitted.

Department of Mathematics
Josai University
Sakado, Saitama 350-02, Japan

References

[1] J.R. Blum, A note on mixing transformations, Israel J. Math., **9** (1971), 464-465.

[2] A. Brunel and M. Keane, Ergodic theorems for operator sequences, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **12** (1969), 231-240.

[3] J.W. England and N.F.G. Martin, On weak mixing metric automorphisms, Bull. Amer. Math. Soc., **74** (1968), 505-507.

[4] S.R. Foguel, The ergodic theory of Markov processes, Van Nostrand, New York, 1969.
[5] D.L. Hanson and G. Pledger, On the mean ergodic theorem for weighted averages, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 13 (1969), 141–149.
[6] U. Krengel and L. Sucheston, On mixing in infinite measure spaces, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 13 (1969), 150–164.
[7] M. Lin, Mixing for Markov operators, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 19 (1971), 231–242.
[8] R. Sato, On the individual ergodic theorem for subsequences, Studia Math., (to appear).
[9] R. Sato, On a decomposition of transformations in infinite measure spaces, Pacific J. Math., (to appear).
[10] R. Sato, Abel-ergodic theorems for subsequences, Pacific J. Math., (to appear).
[11] K. Yosida, Functional analysis, Springer, Berlin-Göttingen-Heidelberg, 1965.