Z. Wahrscheinlichkeitstheorie verw. Geb. 19, 231–242 (1971)

Mixing for Markov Operators

MICHAEL LIN*

1. Definitions and Notations

Let (X, Σ, m) be a finite measure space with $m(X)=1$. A positive linear contraction on $L_1(X, \Sigma, m)$ is called a Markov operator on L_1 and its adjoint, acting on $L_\infty(X, \Sigma, m)$ is a Markov operator on L_∞. P will be written to the right of its variable in L_1 and the adjoint in L_∞ will also be denoted by P and written to the left of its variable. Thus $\langle uP, f\rangle = \langle u, Pf\rangle$ for $u \in L_1$ and $f \in L_\infty$. Identifying L_1 with the space of finite signed measures absolutely continuous with respect to m (via the Radon-Nikodym theorem) P acts on that space: $\mu P(A) = \int P 1_A \, d\mu$ for $\mu \ll m$, $A \in \Sigma$. The same formula defines μP for a σ-finite positive measure $\mu \ll m$. (X, Σ, m, P) will be called a Markov process (and sometimes P alone will be written). If λ is σ-finite and $\lambda \sim m$, (X, Σ, λ, P) is defined by the action of P on the space of finite signed measures $\ll \lambda \sim m$. λ is called *subinvariant* if $\lambda P \leqq \lambda$ and invariant if $\lambda P = \lambda$. λ will always denote either a finite invariant measure $\sim m$ or a σ-finite subinvariant measure $\sim m$ (with $\lambda(X) = \infty$). The adjoint process P^* defined in $L_1(X, \Sigma, \lambda)$ and its properties are described in [4, Chapter VII]. The process is *ergodic* if $Pf = f$ a.e. and $f \in L_\infty$ imply that f is a constant. The process is *irreducible* if $L_1(A, \Sigma \cap A)$ is not invariant under P for $\emptyset \neq A \neq X$ [4, Example I.f], which is equivalent to $P 1_B \leqq 1_B \Rightarrow B$ is either X or $\emptyset$.

The decomposition of X into the *conservative* and *dissipative* parts is given in [4, Chapter II]. If the process is irreducible, it is either conservative or dissipative.

2. Mixing with a Finite Invariant Measure

Definition 2.1. Let P be a Markov process with a finite invariant measure $\lambda \sim m$. P is called *mixing* if for every $A \in \Sigma$ the sequence $\{P^n 1_A\}$ is weak-* convergent in L_∞ to $\lambda(A)/\lambda(X)$, i.e. $\langle \mu, P^n 1_A\rangle \to \mu(X)\, \lambda(A)/\lambda(X)$ for every $\mu \ll \lambda$ or, equivalently, $\langle u, P^n 1_A\rangle \to (\lambda(A)/\lambda(X)) \langle u, 1\rangle$ for every $u \in L_1(\lambda)$. (Integrations in this section are with respect to λ.) In this section we assume $\lambda(X) = 1$. In [4, Chapter VII] it is shown that P on L_∞ is also a contraction of $L_1(\lambda)$, and hence also defines a contraction on $L_2(\lambda)$, again denoted by P. Clearly P is mixing if and only if for every $A \in \Sigma$ $P^n 1_A \to \lambda(A)$ weakly in $L_2(\lambda)$.

Lemma 2.1. *P is mixing if and only if the adjoint process P^* is mixing.*

* This paper is a part of the author's Ph. D. thesis prepared at The Hebrew University of Jerusalem under the direction of Professor S. R. Foguel, to whom the author is grateful for his helpful advice and kind encouragement.

Proof. If $A, B \in \Sigma$ and P is mixing $\langle P^{*n} 1_A, 1_B\rangle = \langle 1_A, P^n 1_B\rangle \to \langle 1_A, \lambda(B)\rangle = \lambda(A)\,\lambda(B)$ and by standard approximation $P^{*n} 1_A \to \lambda(A)$ weakly in $L_2(\lambda)$. The converse – by symmetry.

Lemma 2.2. *Let* $\{a_{ij}\}$ *be a bounded sequence of real numbers satisfying* $\lim_{|i-j|\to\infty} a_{ij} = 0$. *Then* $\lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} = 0$.

The (very simple) *proof* is given in [1].

Theorem 2.1. *Let P be a Markov process with finite invariant probability measure $\lambda \sim m$. Then the following conditions are equivalent:*

(a) *P is mixing.*

(b) *All weak-* limit points in L_∞ of $\{P^n 1_A\}$ are constants.*

(c) *For every $u \in L_1(\lambda)$ with $\int u\, d\lambda = 0$ $u P^n \to 0$ weakly in $L_1(\lambda)$.*

(d) *For every $v \in L_1(\lambda)$ and any increasing subsequence $\{n_i\}$*

$$\left\| \frac{1}{N} \sum_{i=1}^{N} v P^{n_i} - \langle \lambda, v\rangle \right\|_1 \to 0.$$

Proof. (b) $\Rightarrow$ (c). It is enough to show that $\langle u P^n, 1_A\rangle \to 0$ for every $A \in \Sigma$. Suppose that (c) is false, i.e. for some $u \in L_1(\lambda)$ with $\int u\, d\lambda = 0$ and some $A \in \Sigma$ $\langle u P^n, 1_A\rangle \not\to 0$. Hence for a subsequence $\{n_i\}$ and some $\varepsilon > 0$ we have $\langle u, P^{n_i} 1_A\rangle \geqq \varepsilon$ (we may have to change u by $-u$). If $f \in L_\infty$ is a weak-* limit point of $\{P^{n_i} 1_A\}$ it is a constant by (b) and $\langle u, f\rangle = 0$. But $\{g \in L_\infty : |\langle u, g\rangle| < \varepsilon\}$ is a weak-* open neighborhood of f, and must contain infinitely many $\{P^{n_i} 1_A\}$, which contradicts the choice of $\{n_i\}$. Hence (c) is true.

(c) $\Rightarrow$ (a). If $A \in \Sigma$ we have, as $\lambda P = \lambda$, $\|P^n 1_A\|_1 = \int P^n 1_A\, d\lambda = \lambda(A)$. Let $v \in L_1(\lambda)$ and define $u = v - \langle \lambda, v\rangle$. Since constants are in $L_1(\lambda)$, $u \in L_1(\lambda)$ with $\int u\, d\lambda = 0$. By (c)

$$\langle v, P^n 1_A\rangle - \lambda(A)\,\langle \lambda, v\rangle = \langle u, P^n 1_A\rangle = \langle u P^n, 1_A\rangle \to 0$$

which shows that $P^n 1_A$ is weak-* convergent in L_∞ to $\lambda(A)$. Thus P is mixing.

(d) $\Rightarrow$ (b). (d) implies that if $u \in L_1(\lambda)$ with $\int u\, d\lambda = 0$ then $\left\| \frac{1}{N} \sum_{i=1}^{N} u P^{n_i} \right\|_1 \to 0$ for every subsequence $\{n_i\}$.

Let $A \in \Sigma$ and suppose that $\{P^n 1_A\}$ has a non-constant function $f \in L_\infty$ as a weak-* limit point. Clearly $0 \leqq f \leqq 1$. We can find a function $u \in L_1$ with $\int u\, d\lambda = 0$ and $\int u f\, d\lambda > 0$. But since f is a weak-* limit point of $\{P^n 1_A\}$, there exists a subsequence $\{n_i\}$ with $\langle u, P^{n_i} 1_A\rangle \to \langle u, f\rangle$. Hence

$$\left\langle \frac{1}{N} \sum_{i=1}^{N} u P^{n_i}, 1_A \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \langle u, P^{n_i} 1_A\rangle \to \langle u, f\rangle \neq 0$$

which contradicts (d). Thus (b) must hold.

(a) $\Rightarrow$ (d). Since λ is invariant, P on L_∞ is also a contraction of $L_2(\lambda)$, with adjoint P^*. We use an argument of [4, Theorem VIII.A]:

$$\begin{aligned} \|P^{*k} P^k P^n f - P^n f\|_2^2 &= \|P^{*k} P^{n+k} f\|_2^2 - 2\,\|P^{n+k} f\|_2^2 + \|P^n f\|_2^2 \\ &\leqq \|P^n f\|_2^2 - \|P^{n+k} f\|_2^2. \end{aligned}$$

P is a contraction, so $\|P^n f\|_2$ converges, hence $\|P^{*k} P^k P^n f - P^n f\|_2 \to 0$ as $n \to \infty$ *uniformly* in k. Take $A \in \Sigma$.

$$|\langle P^i 1_A, P^j 1_A\rangle - \lambda(A)^2| = |\langle P^{*j} P^j P^{i-j} 1_A, 1_A\rangle - \lambda(A)^2|$$
$$\leqq |\langle P^{*j} P^j P^{i-j} 1_A - P^{i-j} 1_A, 1_A\rangle| + |\langle P^{i-j} 1_A, 1_A\rangle - \lambda(A)^2|$$

as $i-j \to \infty$ the first term tends to zero by the Hilbert space argument above, and the last tends to zero since P is mixing. Hence

$$\lim_{|i-j| \to \infty} |\langle P^i 1_A, P^j 1_A\rangle - \lambda(A)^2| = 0.$$

$$\left\| \frac{1}{N} \sum_{i=1}^{N} P^{n_i} 1_A - \lambda(A) \right\|_2^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle P^{n_i} 1_A - \lambda(A), P^{n_j} 1_A - \lambda(A)\rangle$$
$$= \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \{\langle P^{n_i} 1_A, P^{n_j} 1_A\rangle - \lambda(A)^2\} \to 0.$$

The convergence to zero is by Lemma 2.2. $\lambda(X) = 1$ implies that $1 \in L_2$ and hence

$$\left\| \frac{1}{N} \sum_{i=1}^{N} P^{n_i} 1_A - \lambda(A) \right\|_1 \to 0.$$

By linearity and standard approximation

$$\left\| \frac{1}{N} \sum_{i=1}^{N} P^{n_i} u - \int u \, d\lambda \right\|_1 \to 0$$

for every $u \in L_1$.

By Lemma 2.1 P^* is also mixing, and applying the last result to P^* yields

$$\left\| \frac{1}{N} \sum_{i=1}^{N} v P^{n_i} - \int v \, d\lambda \right\|_1 = \left\| \frac{1}{N} \sum_{i=1}^{N} P^{*n_i} v - \int v \, d\lambda \right\|_1 \to 0.$$

Remark. The definition of mixing coincides with (strong) mixing when P is induced by a point transformation. Theorems 2.1 and 2.2 generalize the results of Blum and Hanson [1] to the case when P is not necessarily induced by a point transformation.

Theorem 2.2. *Let P and λ be as before. Then the following conditions are equivalent:*

(a) *P is mixing.*

(b) *For every $1 \leqq p < \infty$, $f \in L_p(\lambda)$ and sequence $\{n_i\}$*

$$\left\| \frac{1}{N} \sum_{i=1}^{N} P^{n_i} f - \int f \, d\lambda \right\|_p \xrightarrow[N \to \infty]{} 0.$$

(c) *For some $1 \leqq p < \infty$ we have:*

$$\left\| \frac{1}{N} \sum_{i=1}^{N} P^{n_i} f - \int f \, d\lambda \right\|_p \xrightarrow[N \to \infty]{} 0$$

for every $f \in L_p$ and any increasing subsequence $\{n_i\}$.

Proof. (a)$\Rightarrow$(b). For $p=1$ (b) is proved in Theorem 2.1 (a)$\Rightarrow$(d). If $A\in\Sigma$, $P1_A\leqq 1$, so

$$\left\|\frac{1}{N}\sum_{i=1}^{N}P^{n_i}1_A-\lambda(A)\right\|_p^p\leqq\left\|\frac{1}{N}\sum_{i=1}^{N}P^{n_i}1_A-\lambda(A)\right\|_1\to 0$$

and by standard approximation (b) follows.

(b)$\Rightarrow$(c) is obvious.

(c)$\Rightarrow$(a). Let $A\in\Sigma$ be fixed. The given convergence implies that for every $B\in\Sigma$ and any subsequence $\{n_i\}$

$$\frac{1}{N}\sum_{i=1}^{N}\langle P^{n_i}1_A, 1_B\rangle\to\lambda(A)\,\lambda(B).$$

Hence necessarily $\langle P^n 1_A, 1_B\rangle\to\lambda(A)\,\lambda(B)$ for every $B\in\Sigma$, and $P^n 1_A\to\lambda(A)$ weakly in $L_2(\lambda)$, so P is mixing. Q.E.D.

For unaveraged convergence we have:

Theorem 2.3. *Let P and λ be as before. Then the following conditions are equivalent:*

(a) *For every $1\leqq p<\infty$ and $f\in L_p(\lambda)$ $\|P^n f-\int f\,d\lambda\|_p\to 0$.*

(b) *For some $1\leqq p_0<\infty$ we have: $\|P^n f-\int f\,d\lambda\|_{p_0}\to 0$ for any $f\in L_{p_0}(\lambda)$.*

Proof. Clearly (a)$\Rightarrow$(b). To prove (b)$\Rightarrow$(a) we first show that we may assume $p_0=1$. Let $f\in L_1(\lambda)$ be bounded. Then $f\in L_{p_0}(\lambda)$. Let q satisfy $1/p_0+1/q=1$. By Hölder's inequality

$$\|P^n f-\textstyle\int f\,d\lambda\|_1=\int|P^n f-\int f\,d\lambda|\,d\lambda\leqq\|P^n f-\int f\,d\lambda\|_{p_0}\|1\|_q\to 0.$$

Standard approximations show that (b) holds with $p_0=1$. If $f\in L_p(\lambda)$, for $1\leqq p<\infty$, and f is bounded, we have

$$\begin{aligned}\|P^n f-\langle\lambda,f\rangle\|_p^p&\leqq\|P^n f-\langle\lambda,f\rangle\|_1\,\|P^n(f-\langle\lambda,f\rangle)\|_\infty^{p-1}\\&\leqq\|P^n f-\textstyle\int f\,d\lambda\|_1\,\|f-\int f\,d\lambda\|_\infty^{p-1}\to 0.\end{aligned}$$

Standard approximations complete the proof.

Example 2.1. Mixing does not imply unaveraged convergence in L_p-norm. If P is invertible (e.g. induced by an invertible measure preserving transformation which is mixing) $P^*=P^{-1}$ and $\|P^n(f-\langle\lambda,f\rangle)\|_p\to 0$ implies $f=\langle\lambda,f\rangle$.

Example 2.2. The results of Theorem 2.3 hold for P^* but not for P. See [5, p. 116].

3. Mixing with a σ-Finite Subinvariant Measure

Definition 3.1. Let P be a Markov process with a σ-finite subinvariant measure $\lambda\sim m$, and $\lambda(X)=\infty$. P is called *mixing* if for every $A\in\Sigma$ with $\lambda(A)<\infty$ the sequence $\{P^n 1_A\}$ is weak-* convergent in L_∞ to zero.

P on L_∞ defines a contraction on $L_2(\lambda)$ [4, Chapter VII], and clearly P is mixing if and only if $\lambda(A)<\infty$ implies $P^n 1_A \to 0$ weakly in $L_2(\lambda)$.

All integrations in this section are with respect to λ, and we assume $\lambda(X)=\infty$.

Lemma 3.1. *P is mixing if and only if the adjoint process P^* is mixing.*

Proof. If $A, B \in \Sigma$ with finite λ-measure and P is mixing, then

$$\langle P^{*n} 1_A, 1_B \rangle = \langle 1_A, P^n 1_B \rangle \to 0$$

and by standard approximation $P^{*n} 1_A \to 0$ weakly in $L_2(\lambda)$, so P^* is mixing. The converse – by symmetry.

The following well-known lemma gives a multitude of examples to which the following results can be applied.

Lemma 3.2. *If P is dissipative (with σ-finite subinvariant measure $\lambda \sim m$) it is mixing.*

Proof. Let $A \in \Sigma$ with $\lambda(A)<\infty$. Then $1_A \in L_1(\lambda)$ and since P^* is also dissipative [4, (7.2)] $\sum_{n=1}^{\infty} P^n 1_A = \sum_{n=1}^{\infty} 1_A P^{*n} < \infty$ a.e., so $P^n 1_A(x) \to 0$ a.e. $0 \leqq P^n 1_A \leqq 1$ a.e. Hence, if $\lambda(B)<\infty$ $\int 1_B P^n 1_A \, d\lambda \to 0$ by Lebesgue's dominated convergence theorem, whence $P^n 1_A \to 0$ weakly in $L_2(\lambda)$ and P is mixing. Q.E.D.

We note that if P is dissipative it always has a σ-finite subinvariant measure $\lambda \sim m$ see [4, Chapter II].

Theorem 3.1. *Let P be a Markov process with σ-finite subinvariant measure $\lambda \sim m$. Then the following conditions are equivalent:*

(a) *P is mixing.*

(b) *If $\lambda(A)<\infty$ all weak-* limit points in L_∞ of $\{P^n 1_A\}$ are constants.*

(c) *If $u \in L_1(\lambda)$ satisfies $\int u \, d\lambda = 0$, then $\langle u P^n, 1_A \rangle \to 0$ for $A \in \Sigma$ with $\lambda(A)<\infty$.*

Proof. (a) $\Rightarrow$ (b) is obvious, by definition.

(b) $\Rightarrow$ (c). The proof is the same as in Theorem 2.1 (b) $\Rightarrow$ (c), starting from the second sentence.

(c) $\Rightarrow$ (a). Remember that $\lambda(X)=\infty$. Let $A \in \Sigma$ with $\lambda(A)<\infty$. If $P^n 1_A$ is not weak-* convergent to zero, there exist a set B with $0<\lambda(B)<\infty$, an $\varepsilon>0$ and a sequence $\{n_i\}$ with $\langle 1_B, P^{n_i} 1_A \rangle \geqq \varepsilon$. $\lambda(X)=\infty$ implies $\lambda(X-B)=\infty$, and since λ is σ-finite, we may find a set $E \in \Sigma$ disjoint from B with $2\lambda(A)\,\lambda(B)/\varepsilon < \lambda(E) < \infty$. Define $u = 1_B - (\lambda(B)/\lambda(E))\, 1_E$. Then $u \in L_1(\lambda)$ and $\int u \, d\lambda = 0$. But $\langle 1_E, P^{n_i} 1_A \rangle \leqq \lambda(A)$ so $\langle u, P^{n_i} 1_A \rangle = \langle 1_B, P^{n_i} 1_A \rangle - (\lambda(B)/\lambda(E)) \langle 1_E, P^{n_i} 1_A \rangle \geqq \varepsilon - \lambda(A)\,\lambda(B)/\lambda(E) > \varepsilon/2$.

Hence $\langle u P^n, 1_A \rangle$ does not tend to zero, which contradicts (c). Hence $P^n 1_A \to 0$ in L_∞ weak-* topology. (This proof follows [8, Theorem 1.2].) Q.E.D.

Remark. Part (b) of the theorem shows that the notion of mixing in this section generalizes that of the preceding section (Theorem 2.1 (b)). For P induced by a point transformation our definition coincides with that of Krengel and Sucheston, by [8, Theorem 1.3]. The next theorem carries over to the general operator a result obtained in [8] for P induced by a point transformation. It is the generalization of Theorem 2.2 for the σ-finite subinvariant measure.

Theorem 3.2. *Let P and λ be as before. Then the following conditions are equivalent:*

(a) *P is mixing.*

(b) *For every $1<p<\infty$, $f\in L_p(\lambda)$ and sequence $\{n_i\}$*

$$\left\| N^{-1}\sum_{i=1}^{N} P^{n_i} f\right\|_p \xrightarrow[N\to\infty]{} 0.$$

(c) *For some $1<p_0<\infty$ we have: $\left\| N^{-1}\sum_{i=1}^{N} P^{n_i} f\right\|_{p_0} \to 0$ for every $f\in L_{p_0}(\lambda)$ and any increasing subsequence $\{n_i\}$.*

Proof. (a)$\Rightarrow$(b). It is enough to prove the assertion for f of the form 1_A with $\lambda(A)<\infty$, as the general result will follow for f simple function by linearity and for general $f\in L_p(\lambda)$ by standard approximation.

If $p>2$ then, as $P^n 1_A \leqq 1$,

$$\int \left| N^{-1}\sum_{i=1}^{N} P^{n_i} 1_A\right|^p d\lambda \leqq \int \left| N^{-1}\sum_{i=1}^{N} P^{n_i} 1_A\right|^2 d\lambda$$

and thus it is enough to prove the assertion for $1<p\leqq 2$. Define $\delta=p-1$. As P is mixing $P^n 1_A\to 0$ weakly in $L_2(\lambda)$, so $\langle P^n 1_A, 1_A\rangle\to 0$. Hence $\lim_{|i-j|\to\infty}\langle P^i 1_A, P^j 1_A\rangle=0$ by the same arguments as in the proof of Theorem 2.1 (a)$\Rightarrow$(d) (putting zero instead of $\lambda(A)$ there). We proceed as in [8]. Given $\varepsilon>0$, fix $\varepsilon_1>0$ such that $\varepsilon_1^\delta<\varepsilon/\lambda(A)$ and choose $\varepsilon_2>0$ to satisfy $0<\varepsilon_2<\varepsilon_1\varepsilon$. Let M be an integer such that $|i-j|>M\Rightarrow\langle P^i 1_A, P^j 1_A\rangle<\varepsilon_2$. $\{n_i\}$ is increasing, so $|i-j|>M\Rightarrow|n_i-n_j|>M$.

$$\begin{aligned}\int\left(\frac{1}{N}\sum_{i=1}^{N}P^{n_i}1_A\right)^p d\lambda &= \int\frac{1}{N}\sum_{i=1}^{N}P^{n_i}1_A\left(\frac{1}{N}\sum_{j=1}^{N}P^{n_j}1_A\right)^\delta d\lambda\\ &=\frac{1}{N}\sum_{i=1}^{N}\int P^{n_i}1_A\left(\frac{1}{N}\sum_{j=1}^{N}P^{n_j}1_A\right)^\delta d\lambda\\ &\leqq\frac{1}{N}\sum_{i=1}^{N}\int P^{n_i}1_A\left(\frac{1}{N}\sum_{|j-i|\leqq M}P^{n_j}1_A\right)^\delta d\lambda\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\int P^{n_i}1_A\left(\frac{1}{N}\sum_{\substack{|j-i|>M\\ j\leqq N}}P^{n_j}1_A\right)^\delta d\lambda.\end{aligned}\tag{$*$}$$

The last inequality follows from $0<\delta\leqq 1$, since for $0\leqq a, b$ and $0<\delta\leqq 1$ $(a+b)^\delta\leqq a^\delta+b^\delta$. The first term is bounded by

$$\frac{1}{N}\sum_{i=1}^{N}\int P^{n_i}1_A((2M+1)/N)^\delta\, d\lambda=(2M+1)^\delta\,\lambda(A)/N^\delta \xrightarrow[N\to\infty]{} 0$$

so the first term, as M is fixed, tends to zero. To bound the second term in $(*)$, we define $G_{Ni}=\{x:(1/N)\sum_{\substack{|j-i|>M\\ j\leqq N}} P^{n_j}1_A(x)>\varepsilon_1\}$.

$(1/N)\sum_{\substack{|j-i|>M\\ j\leqq N}} P^{n_j}1_A\geqq\varepsilon_1 1_{G_{Ni}}$, so for each i:

$$\begin{aligned}\int P^{n_i}1_A\Big(N^{-1}\sum_{|j-i|>M} P^{n_j}1_A\Big)^\delta d\lambda &= \int_{G_{Ni}} P^{n_i}1_A\Big(N^{-1}\sum_{|j-i|>M} P^{n_j}1_A\Big)^\delta d\lambda+\int_{X-G_{Ni}}\\ &\leqq \int_{G_{Ni}} P^{n_i}1_A\,d\lambda+\varepsilon_1^\delta\lambda(A)\\ &\leqq N^{-1}\varepsilon_1^{-1}\int P^{n_i}1_A\sum_{|j-i|>M}P^{n_j}1_A\,d\lambda+\varepsilon\leqq\varepsilon_2/\varepsilon_1+\varepsilon<2\varepsilon.\end{aligned}$$

Since this is for each i, by averaging the second term in $(*)$ is bounded by 2ε, hence

$$\limsup_{N\to\infty}\int\Big|N^{-1}\sum_{i=1}^N P^{n_i}1_A\Big|^p d\lambda\leqq 2\varepsilon$$

and as ε was arbitrary, (a) $\Rightarrow$ (b) is proved.

(b) $\Rightarrow$ (c) is obvious.

(c) $\Rightarrow$ (a). Let $A\in\Sigma$ with $\lambda(A)<\infty$. By (c) $\Big\langle N^{-1}\sum_{i=1}^N P^{n_i}1_A, 1_A\Big\rangle\to 0$. Since this is for any subsequence $\{n_i\}$ we have necessarily $\langle P^n 1_A, 1_A\rangle\to 0$. By Theorem 3.1 of Foguel's [3] we have $P^n 1_A\to 0$ weakly in $L_2(\lambda)$. Since this is for any $A\in\Sigma$ with $\lambda(A)<\infty$, P is mixing. Q.E.D.

Remarks. (1) For $p=1$ the theorem is not true. This situation is discussed in the next section.

(2) The results of Theorem 3.2 may be applied to P^*, which is mixing when P is.

(3) Mixing does not imply ergodicity when $\lambda(X)=\infty$. For an example let P be the symmetric random walk on the integers with $p(i,i\pm 1)=\frac{1}{2}$. P is ergodic and conservative, $\lambda\{i\}=1$ is invariant, P^2 is not ergodic. P is mixing and hence so is P^2.

Theorem 3.3. *Let P and λ be as before. Then in the following conditions* (a) $\Rightarrow$ (b) $\Rightarrow$ (c):

(a) *For every $f\in L_1(\lambda)$ with $\int f\,d\lambda=0$ $\|P^n f\|_1\to 0$.*

(b) *For every $f\in L_2\cap L_1$ with $\int f\,d\lambda=0$ $\|P^n f\|_2\to 0$.*

(c) *P is mixing.*

Proof. (a) $\Rightarrow$ (b). If $f\in L_2\cap L_1$ with $\int f\,d\lambda=0$ and f is bounded, then

$$\|P^n f\|_2^2=\int P^n f P^n f\,d\lambda\leqq\|P^n f\|_1\,\|P^n f\|_\infty\leqq\|P^n f\|_1\,\|f\|_\infty\to 0.$$

If f is not bounded, we write $f=f^+-f^-$ and find sequences of simple functions $\{f_j^+\}$ and $\{f_j^-\}$ with $0\leqq f_j^+\uparrow f^+$, $0\leqq f_j^-\uparrow f^-$. Hence $\int f_j\,d\lambda\uparrow\int f^+\,d\lambda$ and

$\int f_j^- \, d\lambda \uparrow \int f^- \, d\lambda$. We may assume $\int f_j^+ \, d\lambda \geqq \int f_j^- \, d\lambda$ (by changing f with $-f$ and taking only a subsequence, if necessary). We define $g_j^+ = (\|f_j^-\|_1 / \|f_j^+\|_1) f_j^+$. Hence f_j^-, g_j^+ are simple and tend in L_1 and in L_2 to f^- and f^+ respectively.

$$\|P^n f\|_2 \leqq \|P^n[f-(g_j^+ - f_j^-)]\|_2 + \|P^n(g_j^+ - f_j^-)\|_2.$$

The first term can be made arbitrarily small by fixing j large enough. Then, as $\int (g_j^+ - f_j^-) \, d\lambda = 0$ the second term tends to zero as $n \to \infty$ by the beginning of the proof. Hence (b) holds.

(b) $\Rightarrow$ (c). By Theorem 3.1 (c) we have to show that $\langle f, P^{*n} 1_A \rangle \to 0$ for every $A \in \Sigma$ with $\lambda(A) < \infty$ and any $f \in L_1$ with $\int f \, d\lambda = 0$, in order to conclude that P^* is mixing, and P will be mixing by Lemma 3.1. Fix $A \in \Sigma$ with $\lambda(A) < \infty$. If $f \in L_1$ satisfies $\int f \, d\lambda = 0$ and $f \in L_2$,

$$|\langle f, P^{*n} 1_A \rangle| = |\langle P^n f, 1_A \rangle| \leqq \|P^n f\|_2 \, \lambda(A)^{\frac{1}{2}} \to 0.$$

If $f \in L_1$ with $\int f \, d\lambda = 0$, but $f \notin L_2$, define f_j^- and g_j^+ as above, so that $f_j = g_j^+ - f_j^-$ satisfies $f_j \to f$ in L_1, $f_j \in L_1 \cap L_2$ and $\int f_j \, d\lambda = 0$.

$$|\langle f, P^{*n} 1_A \rangle| \leqq \|f - f_j\|_1 + |\langle f_j, P^{*n} 1_A \rangle|$$

the first term is made small enough by taking a fixed j large enough and the second then tends to zero by the beginning of the proof of (b) $\Rightarrow$ (c). Hence $\langle f, P^{*n} 1_A \rangle \to 0$. Thus P is mixing. Q.E.D.

Remark. The following examples show that we have no possibility of a complete extension of Theorem 2.3.

Example 3.1. (b) does not imply (a) in Theorem 3.3. Let P be the symmetric random walk of remark (3) above. It is easy to check that $P^* = P$. Since P^2 has an invariant set A (the even numbers), (a) cannot hold. Define

$$K = \{f : f \in L_2(\lambda) \colon \|P^n f\|_2 = \|P^{*n} f\|_2 = \|f\|_2 \, \forall n\}.$$

By [4, proof of Theorem VIII.D] $K = \{0\}$, so by a theorem of Horowitz [6] $\|P^n f\|_2 \to 0$ for every $f \in L_2$ and (b) holds.

Example 3.2. Mixing does not imply conditions (b) of Theorem 3.3.

Let T be a point transformation which is an infinite K-automorphism [10, p. 965]. T is invertible so the induced operator P is invertible with $P^* = P^{-1}$, so (b) does not hold. By [10] P has countable Lebesgue spectrum and is mixing. (T can be chosen conservative and ergodic.)

4. Complete Mixing and L_1-Convergence to Zero

Definition 4.1. A Markov process (X, Σ, m, P) is called *completely mixing* if for every $A \in \Sigma$ all weak-* limit points in L_∞ of $\{P^n 1_A\}$ are constants.

By Theorem 2.1 (b) this definition generalizes the notion of mixing when there is an equivalent finite invariant measure.

Theorem 4.1. *Let P be a Markov process on $L_1(X, \Sigma, m)$ having no finite invariant measure $\ll m$. If $0 \neq u \in L_1(m)$ satisfies $u P^n \to 0$ weakly in L_1, then $\|u P^n\|_1 \to 0$.*

The proof of [8, Theorem 5.1] can be adapted to our situation, as indicated at § 5 of [8].

Theorem 4.2. *Let P be a conservative (or an irreducible) Markov process. Then the following conditions are equivalent:*

(a) *P is completely mixing.*

(b) *For every $u \in L_1(m)$ with $\int u\,dm = 0$ $uP^n \to 0$ weakly in L_1.*

(c) *For every $u \in L_1(m)$ with $\int u\,dm = 0$ and any subsequence $\{n_i\}$ we have*

$$\lim_N \left\| N^{-1} \sum_{i=1}^{N} u P^{n_i} \right\|_1 = 0.$$

Proof. (a) $\Rightarrow$ (b). The same as (b) $\Rightarrow$ (c) in Theorem 2.1.

(b) $\Rightarrow$ (c). We first show that P is ergodic. If $Pf = f$ for some non-constant $f \in L_\infty$, we may find a function $u \in L_1(m)$ with $\int u\,dm = 0$ and $\langle u, f\rangle \neq 0$. By (b) we get a contradiction, as $0 \neq \langle u, f\rangle = \langle u, P^n f\rangle = \langle uP^n, f\rangle \to 0$. Thus P is ergodic. By [4, Theorem II.B], if P is ergodic and conservative it is irreducible. Hence every invariant measure is equivalent to m (since S the support of an invariant measure defines an invariant subspace $L_1(S, \Sigma \cap S, m)$ of $L_1(m)$). If there exists an invariant measure $\lambda \sim m$ with $\lambda(X) = 1$, then for $u \in L_1(m)$ with $\int u\,dm = 0$ we define $d\mu = u\,dm$, and $v = d\mu/d\lambda$. Thus $\int v\,d\lambda = \mu(X) = \int u\,dm = 0$ implies by Theorem 2.1(d) that $\left\| N^{-1} \sum_{i=1}^{N} \mu P^{n_i} \right\| \to 0$, so $\left\| N^{-1} \sum_{i=1}^{N} u P^{n_i} \right\|_1 \to 0$.

If there is no finite invariant measure equivalent to m, there is none at all, hence $uP^n \to 0$ weakly in L_1 implies $\|uP^n\|_1 \to 0$ by Theorem 4.1, hence (c) clearly holds.

(c) $\Rightarrow$ (a). Let f be a weak-* limit point of $\{P^n 1_A\}$. If f is not a constant, we may find a function $u \in L_1(m)$ with $\int u\,dm = 0$ and $\langle u, f\rangle \neq 0$. Since f is a weak-* limit point of $\{P^n 1_A\}$ there exists a sequence $\{n_i\}$ with $\langle u, P^{n_i} 1_A\rangle \to \langle u, f\rangle$, hence

$$\left\langle N^{-1} \sum_{i=1}^{N} u P^{n_i}, 1_A \right\rangle \to \langle u, f\rangle \neq 0$$

which contradicts (c). Hence f is constant.

Remarks. (1) Theorem 4.2 was conjectured in [8, § 5] for P induced by a point transformation. This limitation did not permit the use of P^*, which was needed to get part (d) of Theorem 2.1.

(2) (a) and (b) are equivalent without conservativity, nor irreducibility.

(3) We do not know if a conservative completely mixing process has always a σ-finite invariant measure. A finite one does not necessarily exist.

Theorem 4.3. *Let T be a contraction in a Banach space L and let S^* be the closed unit sphere of the conjugate space L^*. Define $A^* = \bigcap_{n=1}^{\infty} T^{*n} S^*$.*

(a) *$u \in L$ satisfies $\|T^n u\| \to 0$ if and only if $\langle f, u\rangle = 0$ for every $f \in A^*$.*

(b) *T^* maps A^* onto itself.*

Proof. (a) Suppose first that $\|T^n u\| \to 0$. If $f \in A^*$ there is a sequence $\{f_n\}$ in S^* with $T^{*n} f_n = f$.

$$|\langle f, u\rangle| = |\langle T^{*n} f_n, u\rangle| = |\langle f_n, T^n u\rangle| \leqq \|T^n u\| \to 0.$$

Suppose now that u satisfies $\langle f, u\rangle = 0$ for every $f \in A^*$, and define $A_n^* = T^{*n} S^*$. By Alaoglu's theorem S^* is compact in the weak-* topology, and as T^* is continuous in that topology, A_n^* is compact. $\{A_n^*\}$ decreases to A^*, which is thus compact.

By [2, p. 65] for every n there is an $f_n \in S^*$ with

$$\|T^n u\| \leqq |\langle f_n, T^n u\rangle| + \frac{1}{n} = |\langle T^{*n} f_n, u\rangle| + \frac{1}{n}.$$

Let f be a weak-* limit point of $\{T^{*n} f_n\}$. Clearly $f \in A^*$, and by hypothesis $\langle f, u\rangle = 0$. Hence there exists a sequence $\{n_i\}$ with $\langle T^{*n_i} f_{n_i}, u\rangle \to 0$, which implies $\|T^{n_i} u\| \to 0$. Since $\|T\| \leqq 1$, $\|T^n u\|$ converges, and thus the limit is zero.

(b) $A_{n+1}^* = T^* A_n^*$ implies $T^* A^* \subseteq A^*$. If $f \in A^*$, $f = T^{*n} f_n$ for some sequence $\{f_n\}$ in S^*. Let g be a weak-* limit point of $T^{*n-1} f_n$, which is clearly in A^*. Since T^* is continuous in the weak-* topology, $T^* g$ is a limit point of $T^{*n} f_n$, hence $T^* g = f$, and T^* is onto. Q.E.D.

For the next theorem we define

$$S = \{f \colon f \in L_\infty, \|f\|_\infty \leqq 1\}; \qquad S^+ = \{f \in L_\infty \colon 0 \leqq f \leqq 1\}.$$

Theorem 4.4. *The following conditions are equivalent for a Markov process P:*

(a) $\|u P^n\|_1 \to 0$ *for every* $u \in L_1$ *with* $\int u\, dm = 0$.

(b) $\bigcap_{n=1}^{\infty} P^n S$ *contains only constant functions.*

(c) *Every sequence* $\{f_n\}$ *in* S *satisfying* $P f_{n+1} = f_n$ $(n = 1, 2, \ldots)$ *contains only constants.*

(d) $\bigcap_{n=1}^{\infty} P^n S^+$ *contains only constant functions.*

(e) *Every sequence* $\{f_n\}$ *in* S^+ *satisfying* $P f_{n+1} = f_n$ *contains only constant functions.*

Proof. (b) $\Rightarrow$ (c) is immediate.

(c) $\Rightarrow$ (b). By Theorem 4.3(b) P maps $A = \bigcap_{n=1}^{\infty} P^n S$ onto itself. If $f \in A$, we put $f_1 = f$ and there is an $f_2 \in A$ with $P f_2 = f_1$, $f_3 \in A$ with $P f_3 = f_2$ and so on. By (c) f_1 is a constant.

(b) $\Leftrightarrow$ (a) by Theorem 4.3(a).

(b) $\Rightarrow$ (d) $\Rightarrow$ (e) is obvious.

(e) $\Rightarrow$ (d). The proof of Theorem 4.3(b) can easily be adapted to show that P maps $\bigcap P^n S^+$ onto itself (S^+ is compact in the weak-* topology), and (e) $\Rightarrow$ (d) by the proof of (c) $\Rightarrow$ (b).

(d) ⇒ (a). Let μ_n be defined by $d\mu_n = u P^n dm$. Let $X = A_n \cup B_n$ be the Hahn decomposition of μ_n. Hence $\|\mu_n\| = \mu_n(A_n) - \mu_n(B_n)$. It is enough to show $\mu_n(A_n) \to 0$, $\mu_n(B_n) \to 0$, which is done in the same way as in the proof of Theorem 4.3(a) by considering weak-* limit points of $\{P^n 1_{A_n}\}$, which are constants by (d).

Remark. For the case that P is given by a transition probability and is considered as an operator on the space of all finite measures on (X, Σ) this result is due to Jamison and Orey [7], who gave a probabilistic proof.

Corollary 4.1. *Let P be induced by the point transformation T (i.e. $Pf(x) = f(Tx)$ a.e.). A necessary and sufficient condition for the convergence $\|(\mu - \nu) P^n\| \to 0$ for any two probability measures $\ll m$ is that $\bigcap_{n=1}^{\infty} T^{-n} \Sigma = \{\emptyset, X\}$.*

Proof. Define $\Sigma_n = T^{-n}\Sigma$. Then Σ_n is a σ-algebra, which is the smallest with respect to which all the functions $P^n f$ are measurable. If $A \in \bigcap_{n=1}^{\infty} \Sigma_n$, then $A = T^{-n} A_n$, or $P^n 1_{A_n} = 1_A$. Thus by Theorem 4.4 the condition is clearly necessary. If $f \in \bigcap P^n S$, then f is Σ_n-measurable, so the condition implies that f is a constant. Hence by Theorem 4.4 $(\mu - \nu)(X) = 0$ implies $\|(\mu - \nu) P^n\| \to 0$.

Remark. A proof of Corollary 4.1 using the Martingale convergence theorem can be found in [6] (a slight modification is needed to dispose of the assumption of a finite invariant measure). The results of [8] imply the theorem only when there is *no* finite invariant measure $\ll m$.

Acknowledgment and Remarks (added March, 1971)

It is a pleasure to thank Professors Krengel and Sucheston for their comments on the manuscript and for sending me preprints of some of their work. These lead to the following remarks.

1. The proof of Theorem 2.1 can be used to show that if P is a contraction on a Hilbert space H, then $P^n f \to 0$ weakly iff $N^{-1} \Sigma P^{n_i} f \to 0$ strongly for every increasing $\{n_i\}$. This result was proved independently by Akcoglu and Sucheston ("On operator convergence in Hilbert space and in Lebesgue space", to appear in Periodica Math. Hungarica).

2. Theorem 4.2 was proved independently (in essentially the same way) by Mr. Winkler at the Ohio State University. Akcoglu and Sucheston (loc. cit.) have then proved it, without assuming conservativity (or irreducibility), as a corollary of a more general theorem. (It should be mentioned that an indication of a proof for point transformations, much less elementary when the space is not a Lebesgue space, *is* given in [8].)

3. A conservative completely mixing process does not necessarily have a σ-finite invariant measure. If P_1 is a conservative ergodic operator without invariant measure, then $P = \delta I + (1 - \delta) P_1$ is completely mixing for $0 < \delta < 1$ (Ornstein and Sucheston, "An operator theorem on L_1 convergence to zero", Ann. Math. Stat. **41**, 1631–1639 (1970)), and is conservative without invariant measure [4, p. 94]. (Example due to Sucheston.)

4. The author has proved that an irreducible process is completely mixing if and only if its cartesian square is completely mixing.

References

1. Blum, J. R., Hanson, D. L.: On the mean ergodic theorem for subsequences. Bull. Amer. math. Soc. **66**, 308–311 (1960).
2. Dunford, N., Schwartz, J.: Linear operators, Part I. New York: Interscience 1958.
3. Foguel, S. R.: Powers of a contraction in Hilbert space. Pacific J. Math. **13**, 551–562 (1963).
4. – The ergodic theory of Markov processes. New York: Van Nostrand 1969.

5. Horowitz, S.: Some limit theorems for Markov processes. Israel J. Math. **6**, 107–118 (1968).
6. – Strong ergodic theorems for Markov processes. Proc. Amer. math. Soc. **23**, 328–334 (1969).
7. Jamison, B., Orey, S.: Markov chains recurrent in the sense of Harris. Z. Wahrscheinlichkeitstheorie verw. Geb. **8**, 41–48 (1967).
8. Krengel, U., Sucheston, L.: On mixing in infinite measure spaces. Z. Wahrscheinlichkeitstheorie verw. Geb. **13**, 150–164 (1969).
9. Neveu, J.: Mathematical foundations of the calculus of probability. San Francisco: Holden-Day 1965.
10. Parry, W.: Ergodic and spectral analysis of certain infinite measure preserving transformations. Proc. Amer. math. Soc. **16**, 960–966 (1965).

M. Lin
Institute of Mathematics
Hebrew University
Jerusalem
Israel

(Received June 26, 1970)