The Annals of Mathematical Statistics
1968, Vol. 39, No. 6, 2145–2148

A GENERALIZATION OF ITO'S THEOREM CONCERNING THE POINTWISE ERGODIC THEOREM

By Choo-Whan Kim

Simon Fraser University

1. Introduction. Let $(X, \mathfrak{F}, \lambda)$ be a σ-finite measure space. Let $L_1(\lambda) = L_1(X, \mathfrak{F}, \lambda)$ be the real Banach space of λ-integrable real functions and $L_\infty(\lambda)$ the dual of $L_1(\lambda)$. All subsets of X discussed in this paper are elements of $\mathfrak{F}$. For two sets A and B, $A \subset B$, $A = B$ mean that $\lambda(A - B) = 0$, $\lambda(A \triangle B) = 0$, respectively. All functions on X are $\mathfrak{F}$-measurable real functions and will always be considered up to λ-equivalence. For two functions f and g on X, $f = g$, $f \leqq g$ mean that the equality and the inequality, respectively, are satisfied in the almost everywhere (a.e.) sense with respect to λ. $\{f \geqq g\}$ denotes the set $\{x \mid f(x) \geqq g(x)\}$. For any set A, A' denotes its complement and 1_A designates the characteristic function of A.

Let $T: f \rightarrow fT$ be a positive linear contraction, (i.e., $\|T\|_1 \leqq 1$) on $L_1(\lambda)$ to $L_1(\lambda)$. We call T a *Markov operator* on $L_1(\lambda)$. The adjoint of T which acts on $L_\infty(\lambda)$ will be denoted by T, but we will write Tg for $g \in L_\infty(\lambda)$. The adjoint T is characterized by (1) T is a positive linear operator, (2) $T1 \leqq 1$, (3) $g_k \downarrow 0$ implies $Tg_k \downarrow 0$ ([8], p. 86). We have then $\int fT \cdot g\lambda(dx) = \int f \cdot Tg\lambda(dx)$ for $f \in L_1(\lambda)$, $g \in L_\infty(\lambda)$.

The purpose of this paper is to prove the following generalization of Ito's results ([6], Theorem 1, Lemma 2).

Theorem. *Let T be a Markov operator on $L_1(\lambda)$. Suppose that the sequence $\{(1/n) \sum_{k=0}^{n-1} wT^k \mid n = 1, 2, \cdots\}$ is weakly sequentially compact for some $w \in L_1(\lambda)$ such that $w > 0$. Then the following assertions hold:*

Assertion 1. (the pointwise ergodic theorem). For each $f \in L_1(\lambda)$,

$$\lim_{n\to\infty} (1/n) \sum_{k=0}^{n-1} fT^k \quad \textit{exists } (\lambda\text{-a.e.}).$$

Assertion 2. (the $L_1(\lambda)$-mean ergodic theorem). For each $f \in L_1(\lambda)$,

$$\lim_{n\to\infty}(1/n) \sum_{k=0}^{n-1} fT^k \quad \textit{exists in the } L_1(\lambda)\textit{-norm}.$$

We will prove Assertions 1 and 2 in Section 2 and 3, respectively. Certain relevant facts are stated in Section 2.

The author wishes to thank the referee for valuable suggestions and, in particular, for a simplified proof of Assertion 2.

2. Proof of Assertion 1. The following lemma follows readily from the mean ergodic theorem of Yosida and Kakutani ([7], p. 441; [11], p. 192).

Lemma 1. *If the sequence $\{(1/n) \sum_{k=0}^{n-1} wT^k\}$ is weakly sequentially compact, then the sequence converges in the $L_1(\lambda)$-norm to a function $u \in L_1^+(\lambda)$ which is invariant under T*, i.e., $uT = u$.

Received 5 January 1968.

Thus for the sequence $\{(1/n) \sum_{k=0}^{n-1} wT^k\}$, the three concepts: weak sequential compactness, weak convergence and strong convergence are equivalent. Henceforth we assume $u = s - \lim_{n\to\infty} (1/n) \sum_{k=0}^{n-1} wT^k$, where $s - \lim$ denotes the strong limit. Then $uT = u$.

We review basic concepts related to Markov operator ([2], [5]). It has been shown by Hopf ([5], Theorem 8.1) and Chacon ([2], Theorem 2) that for each Markov operator T on $L_1(\lambda)$, there exists a subset C of X, unique up to equivalence, such that for each $f \varepsilon L_1^+(\lambda)$, (1) $\sum_{k=0}^{\infty} fT^k = 0$ or ∞ on C, (2) $\sum_{k=0}^{\infty} fT^k < \infty$ on C'. The subsets C and $D = C'$ are called, respectively, *the conservative part* and *the dissipative part of* X relative to T. In fact, we have $C = \{\sum_{k=0}^{\infty} fT^k = \infty\}$, where f is an arbitrary but fixed positive element of $L_1(\lambda)$. Following Neveu [9] we write $C_f = \{\sum_{k=0}^{\infty} fT^k = \infty\}$ for $f \varepsilon L_1^+(\lambda)$. Then $C_f = C \cap \{\sum_{k=0}^{\infty} fT^k > 0\}$ for each $f \varepsilon L_1^+(\lambda)$. Hereafter C and D, respectively, denote the conservative and the dissipative parts of X relative to a given Markov operator T.

LEMMA 2. *If* $u = s - \lim_n (1/n) \sum_{k=0}^{n-1} wT^k$, *then* $C_u = \{u > 0\}$.

PROOF. Since $u \varepsilon L_1^+(\lambda)$ and $uT = u$, we have

$$\{u > 0\} \subset C_u = C \cap \{\sum_{k=0} uT^k > 0\} = C \cap \{u > 0\} \subset \{u > 0\}.$$

The following lemma of Hopf ([5], Lemma 9.4) proved for a finite measure λ, is also true for a σ-finite measure λ. For completeness of our argument we state and prove

LEMMA 3. *If* $Th \leqq h$ $(Th \geqq h)$ *on* C *for* $h \varepsilon L_\infty(\lambda)$, *then* $Th = h$ *on* C.

PROOF. It is enough to consider the case where $Th \leqq h$ on C. Let $A = \{Th < h\} \cap C$. Let $f \varepsilon L_1^+(\lambda)$ be such that $\{f > 0\} = C$. It follows from Theorem 8.2 of [5] that $fT^k = 0$ on D for each $k = 1, 2, \cdots$. Hence we have the inequality

$$\int_A (h - Th) \sum_{k=0}^{n} fT^k\lambda(dx) \leqq \int (h - Th) \sum_{k=0}^{n} fT^k\lambda(dx)$$
$$= \int (h - T^{n+1}h)f\lambda(dx) \leqq 2\|h\|_\infty \cdot \|f\|_1 < \infty, \qquad n = 1, 2, \cdots.$$

We have, from the monotone convergence theorem,

$$\int_A (h - Th) \sum_{k=0}^{\infty} fT^k\lambda(dx) < \infty.$$

However, since $\sum_{k=0}^{\infty} fT^k = \infty$ on each λ-non-null subset of C, we have $\lambda(A) = 0$.

COROLLARY. *The following equalities hold on* C.

$$T1 = 1, \quad T1_{C_f'} = 1_{C_f'}, \quad T1_{C_f} = 1_{C_f}, \quad T1_C = 1_C, \quad \text{where} \quad f \varepsilon L_1^+(\lambda).$$

A set A is called *closed* (*stochastically*) if $T1_A = 1_A$ on A. We prove the following.

LEMMA 4. *Let* B *be a subset of* C. *The following are equivalent:*

(1) B *is closed;* i.e., $T1_B = 1_B$ *on* B.

(2) *If for* $f \varepsilon L_1^+(\lambda)$, $\{f > 0\} \subset B$, *then* $\{fT > 0\} \subset B$.

(3) $B = C_g$ *for some* $g \varepsilon L_1^+(\lambda)$.

(4) $T1_B = 1_B$ *on* C.

PROOF. (1) $\Rightarrow$ (2): We know from (1) and the Corollary to Lemma 3, that

$T1_{C-B} = 0$ on B. Suppose that $\{f > 0\} \subset B$ for some $f \in L_1^+(\lambda)$. Since $fT = 0$ on D from Theorem 8.2 of [5], it remains to show $fT = 0$ on $C - B$. But

$$\int_{C-B} fT \cdot \lambda(dx) = \int f \cdot T1_{C-B}\lambda(dx) = \int_B f \cdot T1_{C-B}\lambda(dx) = 0,$$

so the assertion holds. (2) $\Rightarrow$ (3): Let $g \in L_1^+(\lambda)$ be such that $\{g > 0\} = B$. It follows from (2) that $gT^k = 0$ on $X - B$ for each k. Clearly, $B \subset C$. Then $C - B \subset X - B \subset \{\sum_{k=0}^{\infty} gT^k = 0\}$ and $C - B \subset C \cap \{\sum_{k=0}^{\infty} gT^k = 0\} = C - C_g$ from Lemma 8.4 of [5]. Hence $B \supset C_g$. On the other hand, $C_g = C \cap \{\sum_{k=0}^{\infty} gT^n > 0\} \supset C \cap \{g > 0\} = B$. The implications (3) $\Rightarrow$ (4) $\Rightarrow$ (1) are obvious.

A subset B of C is called an *invariant set* if it satisfies one of the four conditions of Lemma 4. We may readily show that the class of all invariant sets forms a σ-algebra of subsets of C.

LEMMA 5. *If* $u = s - \lim_n (1/n) \sum_{k=0}^{n-1} wT^k$, *then* $\lim_k T^k 1_{C_u'} = 0$.

PROOF. Since the set C_u is an invariant set and $T1_{C_u'} = 0$ on C_u, we readily have $T1_{C_u'} \leqq 1_{C_u'}$. However T being positive implies that $\{T^k 1_{C_u'} \mid k \geqq 1\}$ is a decreasing sequence. If we write $h = \lim_k T^k 1_{C_u'}$, then $h = \lim_n (1/n) \cdot \sum_{k=0}^{n-1} T^k 1_{C_u'}$. Let $\mu = \lambda_w$ be defined by $\mu(A) = \int_A w\lambda(dx)$. Then μ is a finite measure equivalent to λ. Now we have the following equality:

$$\int_{C_u'} (1/n) \sum_{k=0}^{n-1} wT^k \lambda(dx) = \int (1/n) \sum_{k=0}^{n-1} T^k 1_{C_u'} \mu(dx).$$

By using the Lebesgue dominated convergence theorem on the right hand side and the weak convergence of $\{(1/n) \sum_{k=0}^{n-1} wT^k\}$ on the other side, we have

$$0 = \int_{C_u'} u\lambda(dx) = \int h\mu(dx).$$

Since $h \geqq 0$, $h = 0$ μ-a.e.; equivalently, $h = 0$ λ-a.e.

LEMMA 6. *If* $u = s - \lim_n (1/n) \sum_{k=0}^{n-1} wT^k$, *then* $C = C_u = \{u > 0\}$.

PROOF. It is enough to show that $\lambda(A) = 0$, where $A = C \cap C_u'$. Since $T^k 1_A \leqq T^k 1_{C_u'}$, $k = 1, 2, \cdots$, it follows from Lemma 5 that $\lim_k T^k 1_A = 0$. On the other hand the set $A = C \cap C_u'$ is invariant; i.e., $T1_A = 1_A$ on C. By the usual argument we have

$$0 = \lim_k \int T^k 1_A \cdot w\lambda(dx) \geqq \lim_k \int_C T^k 1_A \cdot w\lambda(dx) = \int_C 1_A \cdot w\lambda(dx) \geqq 0,$$

so $\lambda(A) = 0$.

PROOF (Assertion 1). It is enough to show that for each $f \in L_1^+(\lambda)$, $\lim_n (1/n) \sum_{k=0}^{n-1} fT^k$ exists (λ-a.e.) on C. We assume from Lemma 1 that $u = s - \lim_n (1/n) \sum_{k=0}^{n-1} wT^k$. By the general ergodic theorem of Chacon-Ornstein [1] and Lemma 6, we have, for $f \in L_1^+(\lambda)$, $\lim_n \sum_{k=0}^{n-1} fT^k / \sum_{k=0}^{n-1} uT^k = u^{-1} \lim_n (1/n) \sum_{k=0}^{n-1} fT^k$ exists (λ-a.e.) on $C \cap \{\sum_{k=1}^{\infty} uT^k > 0\} = C \cap \{u > 0\} = C$. Hence the assertion holds.

3. Proof of Assertion 2. Our point of departure is the following lemma.

LEMMA 7. *Let* $\{g_n\}$ *and* $\{h_n\}$ *be sequences in* $L_1(\lambda)$ *such that* (1) $0 \leqq g_n \leqq h_n$, $n = 1, 2, \cdots$, (2) $h_n \to h$ *in the* $L_1(\lambda)$*-norm. Then* $\{g_n\}$ *is weakly sequentially compact.*

PROOF. It is well known ([4], Theorem C, p. 108) that the measures $\mu_n(E) = \int_E h_n \, d\lambda, E \varepsilon \mathfrak{F}, n = 1, 2, \cdots$, are uniformly absolutely continuous with respect to the measure λ and equicontinuous from above at the empty set $\varnothing$ under the condition (2). In particular, the conditions (1) and (2) imply that the measures $\nu_n(E) = \int_E g_n \, d\lambda, E \varepsilon \mathfrak{F}, n = 1, 2, \cdots$, are equicontinuous from above at $\varnothing$. It is easy to see that $\{g_n\}$ is bounded in the $L_1(\lambda)$-norm. From a theorem of Dunford and Pettis ([3], Theorem 9, p. 292), we establish the assertion.

PROOF (Assertion 2). In view of the mean ergodic theorem of Yosida and Kakutani ([7], [11]), it suffices to show that the sequence $\{(1/n) \sum_{k=0}^{n-1} fT^k\}$ is weakly sequentially compact for every f belonging to some fundamental subset of $L_1(\lambda)$. (A subset of a Banach space is called fundamental if the linear span of the set is dense in the space.) Let for $t = 1, 2, \cdots, B_t = \{x \mid w(x) > 1/t\}$, where $w(x)$ is the function appearing in the assumption of the theorem. Then, for each t, $B_t \varepsilon \mathfrak{F}$, $\lambda(B_t) < \infty$ and $\bigcup_{t=1}^{\infty} B_t = X$. If we denote by M the set of all functions f of the form $f = 1_{B \cap B_t}$, where B is an arbitrary set in $\mathfrak{F}$ and t a positive integer, then it is easy to see that M is a fundamental subset of $L_1(\lambda)$. We now show that for each f in M, the sequence $\{(1/n) \sum_{k=0}^{n-1} fT^k\}$ is weakly sequentially compact. So, let $f \varepsilon M$. Then, $f = 1_{B \cap B_t}$ for some t, and we have $f(x) \leqq tw(x)$ for all x. The positivity of T now implies $0 \leqq (1/n) \sum_{k=0}^{n-1} fT^k \leqq t(1/n) \sum_{k=0}^{n-1} wT^k, n = 1, 2, \cdots$. We complete the proof by using Lemmas 1 and 7 to the sequences $\{g_n = (1/n) \sum_{k=0}^{n-1} fT^k\}$ and $\{h_n = t(1/n) \sum_{k=0}^{n-1} wT^k\}$.

REFERENCES

[1] CHACON, R. V. AND ORNSTEIN, D. S. (1960). A general ergodic theorem. *Illinois J. Math.* **4** 153–160.

[2] CHACON, R. V. (1962). Identification of the limit of operator averages. *J. Math. Mech.* **11** 961–968.

[3] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators,* Part I. Interscience, New York.

[4] HALMOS, P. R. (1954). *Measure Theory.* Van Nostrand, New York.

[5] HOPF, E. (1954). The general temporally discrete Markoff process. *J. Math. Mech.* **3** 13–45.

[6] ITO, Y. (1965). Uniform integrability and the pointwise ergodic theorem. *Proc. Amer. Math. Soc.* **16** 222–227.

[7] LOÈVE, M. (1960). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

[8] MOY, S. C. (1965). λ-continuous Markov chains. *Trans. Amer. Math. Soc.* **117** 68–91.

[9] NEVEU, J. (1961). Sur le théorème ergodique ponctuel. *C. R. Acad. Sci. Paris,* **252** 1554–1556.

[10] NEVEU, J. (1964). *Bases Mathématiques du Calcul des Probabilités.* Masson, Paris.

[11] YOSIDA, K. and KAKUTANI, S. (1941). Operator-theoretical treatment of Markoff's process and mean ergodic theorem. *Ann. Math.* **42** 188–228.