The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1631–1639

AN OPERATOR THEOREM ON L_1 CONVERGENCE TO ZERO WITH APPLICATIONS TO MARKOV KERNELS

By Donald Ornstein[1] and Louis Sucheston[2]

Stanford University and Ohio State University

0. Summary. A recent theorem of Orey [12] (see also [1], [6], [7], [13]) asserts that if T is an L_1 operator induced on a discrete measure space by an irreducible recurrent aperiodic Markov matrix, then the condition (C) holds: $f \in L_1$, $\int f = 0$ implies that $T^n f$ converges to zero in L_1. In an attempt to determine when (C) holds for more general operators, we at first prove the following (Theorem 1.1): Let T be a positive linear contraction operator on L_1; if $T^n f$ and $T^{n+1}f$ intersect slightly, but uniformly in f in the unit sphere of L_1, then $T^n f - T^{n+1}f$ converges to zero in norm. (C) follows if T is conservative and ergodic (Corollary 1.3). In Section 2 we derive from this a simple proof of Orey's theorem. The main result of the paper is in Section 3 and could be called a "zero-two" theorem: Let $P(x, A)$ be a Markov kernel, and assume that there is a σ-finite measure m such that for each A, $m(A) = 0$ implies $P(x, A) = 0$ a.e. and $m(A) > 0$ implies $\sum_{n=0}^{\infty} P^{(n)}(x, A) = \infty$ a.e. Then the total variation of the measure $P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot)$ is either a.e. 2 for all n or it converges a.e. to 0 as $n \to \infty$. In Section 4 it is shown that a version of the zero-two theorem essentially contains the Jamison–Orey generalization of Orey's theorem to Harris processes.

Section 1 and Section 2 of this paper do not assume any knowledge of either operator ergodic theory or probability. Some known results in ergodic theory are applied in Section 3, but the proof of the main theorem does not depend on them.

1. Operator theorems. Let $(X, \mathscr{A}, m)$ be a measure space, and let L_1 be the class of integrable functions from X to the real line. L_1^+ is the class of nonnegative elements of L_1; an operator T on L_1 is called *positive* iff $TL_1^+ \subset L_1^+$. The relations below are frequently understood to hold modulo sets of measure zero. The L_1 norm of a function or an operator is denoted by $\|\ \|$. T is a *contraction* iff $\|T\| \leqq 1$.

If T is a contraction, $\|T^n(f - Tf)\|$ is a non-increasing sequence of numbers converging to a limit. We investigate when this limit is zero.

Theorem. 1.1. *Let T be a positive linear operator on L_1 with $\|T\| \leqq 1$. Assume that for every $f \in L_1^+$ with $\|f\| = 1$,*

$$\lim_n \|T^n f - T^{n+1} f\| < 2(1-\varepsilon) \tag{1.1}$$

where ε is a positive constant independent of f. Then for each $f \in L_1$

$$\lim_n \|T^n f - T^{n+1} f\| = 0. \tag{1.2}$$

Received July 8, 1969; revised April 1, 1970.

[1] The research of this author was supported by NSF Grant GP 8781.

[2] The research of this author was supported by NSF Grant GP 7693.

PROOF. It suffices to prove (1.2) for $f \in L_1^+$ with $||f|| = 1$. Given such an f, choose a fixed N so large that $||T^N f - T^{N+1} f|| < 2(1-\varepsilon)$. Let $h = \inf(T^N f, T^{N+1} f)$. Then for some $g, g' \in L_1^+$,

$$T^N f = h + g, \tag{1.3}$$

$$T^{N+1} f = h + g', \tag{1.4}$$

$$||g+g'|| = ||T^N f + T^{N+1} f - 2h|| = ||T^N f - T^{N+1} f||. \tag{1.5}$$

Applying T to (1.3) and adding the outcome to (1.4) gives

$$2T^{N+1} f = h + Th + Tg + g'.$$

Hence, on setting $n_1 = N+1$, $h_1 = h/2$, $g_1 = (Tg+g')/2$, we have

$$T^{n_1} f = h_1 + Th_1 + g_1 \tag{1.6}$$

and $||g_1|| < 1-\varepsilon$ since

$$2||g_1|| = ||Tg+g'|| \leqq ||g|| + ||g'|| = ||g+g'|| = ||T^N f - T^{N+1} f||.$$

Now apply the same argument to $g_1/||g_1||$, obtaining

$$T^{\bar{n}_1} g_1 = \bar{h}_1 + T\bar{h}_1 + \bar{g}_1 \tag{1.7}$$

with $||\bar{g}_1|| < (1-\varepsilon)^2$. On letting $n_2 = n_1 + \bar{n}_1$, one has from (1.6) and (1.7):

$$T^{n_2} f = h_2 + Th_2 + g_2$$

with $h_2 = T^{\bar{n}_1} h_1 + \bar{h}_1$ and $||g_2|| < (1-\varepsilon)^2$. Continuing in this way, one obtains a sequence of integers $n_1 < n_2 < \cdots$ such that for every positive integer k,

$$T^{n_k} f = h_k + Th_k + g_k$$

and $||g_k|| < (1-\varepsilon)^k$. Therefore, for N large, $T^N f$ may be approximated in norm by functions of the form $h + Th$, and $||f|| = 1$ implies that we can have $||h|| \leqq \frac{1}{2}$ or $||Th|| \leqq \frac{1}{2}$. We may and do assume $||h|| \leqq \frac{1}{2}$: if $||Th|| \leqq \frac{1}{2}$, replace h by Th and N by $N+1$. Again, for N' large, $T^{N'} h$ may be approximated by functions of the form $h' + Th'$ with $||h'|| \leqq \frac{1}{4}$; hence, for N'' large ($N'' = N + N'$), $T^{N''} f$ may be approximated by functions of the form $(I+T)^2 h''$ with $||h''|| \leqq \frac{1}{4}$, etc. Continuing in this way, we obtain that for N large $T^N f$ may be approximated by functions of the form $(I+T)^n h$, where n is arbitrary, h is nonnegative and $||h|| \leqq \frac{1}{2}^n$; hence $||T^N f - T^{N+1} f||$ may be approximated by $(I+T)^n (I-T)h$. To prove the theorem, it now suffices to show that

$$\lim_n \sum_{k=1}^{n} \left| \binom{n}{k} - \binom{n}{k-1} \right| 2^{-n} = 0. \tag{1.8}$$

It suffices to consider n even. Since $\binom{n}{k}$ is an increasing (decreasing) function of k for $k \leqq \frac{1}{2}n$ ($k > \frac{1}{2}n$), the sum in (1.8) may be written as a difference of sums $\sum_{k=1}^{\frac{1}{2}n}$ and $\sum_{k=\frac{1}{2}n+1}^{n}$, with the absolute value signs removed in each sum and cancellations occurring. (1.8) is now seen to follow from the convergence of $\binom{n}{\frac{1}{2}n} 2^{-n}$ to zero, which is an immediate consequence of Stirling's formula.

We now require a theorem about uniqueness of invariant functions. We state the result with greater generality than needed; the main particular cases of Theorem 1.2 are well known.

If L is a class of real-valued functions, we call an operator V on L *positive* iff $VL^+ \subset L^+$ where $L^+ = \{f \text{ in } L \text{ with } f \geqq 0\}$. We write V_∞ for the operator $I + V + V^2 + \cdots$.

THEOREM 1.2. *Let X be an abstract set, and let L be a set of real-valued functions on X, which is a linear space and a lattice under pointwise operations. Assume that V is a positive linear operator on L and*

$$\text{(1.9)} \qquad \textit{for each } f \not\equiv 0 \textit{ in } L^+, \qquad V_\infty f = \infty \textit{ on } X.$$

Let $f_1 \not\equiv 0$ *and* f_2 *in* L^+ *be such that* $Vf_1 = f_1$ *and* $Vf_2 = f_2$. *Then for some constant c,* $cf_1 = f_2$.

PROOF. Let g be such that $Vg = g$. Write g^+ for $\sup(g, 0)$. Then $Vg^+ \geqq Vg = g = g^+$ on $A = \{g > 0\}$ and $Vg^+ \geqq g^+ = 0$ on A^c; thus, $Vg^+ \geqq g^+$. Now set $g = f_1 - f_2$ and $f = Vg^+ - g^+$. $f + Vf + \cdots + V^{n-1}f = V^n g^+ - g^+ \leqq V^n g^+ \leqq V^n f_1 = f_1$; hence, by (1.9), $f \equiv 0$ and $Vg^+ = g^+$. Therefore, $A = \{V_\infty g^+ = \infty\}$, and, again by (1.9), $A = X$ or $A = \varnothing$ (empty set). Replace f_1 by cf_1 where c is a constant; we showed that for each c, $cf_1 > f_2$ or $cf_1 \leqq f_2$. It easily follows that there exists c_0 with $c_0 f_1 = f_2$.

REMARK. In ergodic theory the assumption (1.9) is usually expressed by calling V "conservative and ergodic".

Theorem 1.2 will be applied in the case when L is L_∞ of a σ-finite measure space $(X, \mathscr{A}, m)$ and V is T^*, the adjoint of a positive linear operator T on L_1. We note that: (i) (1.9) holds with $L = L_1$ and $V = T$, if and only if: (ii) (1.9) holds with $L = L_\infty$, $V = T^*$. We only show that (i) implies (ii); the converse implication is proved similarly. Assume (i). If $T^*_\infty f < \infty$ on a set of positive measure, then there exists a set A with $0 < m(A) < \infty$ and a function g in ${L_1}^+$ such that on A, $T^*_\infty f \leqq g$. By the duality relation, one has $\int f \cdot T_\infty 1_A \leqq \int g$, which by (i) implies $f \equiv 0$.

COROLLARY 1.3. *Assume that T is a positive linear operator on* L_1 *and* $\|T\| \leqq 1$. *Assume that* (1.9) *holds with* $L = L_1$ *and* $V = T$ (*equivalently*: *with* $L = L_\infty$ *and* $V = T^*$). *If for every* $f \in {L_1}^+$ *with* $\|f\| = 1$ (1.1) *holds with an* $\varepsilon > 0$ *independent of f, then* (C) *holds*: *for every* $f \in L_1$ *with* $\int f = 0$, *one has*

$$\text{(1.10)} \qquad \lim_n \|T^n f\| = 0.$$

PROOF. Let $L_0 = \{f : f \in L_1, \int f = 0\}$. To prove the corollary, it suffices to show that $(I - T)L_1$ is dense in L_0, and then to apply Theorem 1.1. A linear functional δ on L_0 may be extended to L_1; hence each δ corresponds to a bounded function h_δ by the relation: $\delta(f) = \int f h_\delta$, $f \in L_0$. Assume that δ vanishes on $(I - T)L_1$; then $T^* h_\delta = h_\delta$. On the other hand, applying (1.9) with $L = L_\infty$, $V = T^*$, $f = 1 - T^*1$, we obtain: $T^*1 = 1$. Therefore, by Theorem 1.2, h_δ is a constant, and hence δ vanishes on L_0. It follows (Hahn–Banach) that $(I - T)L_1$ is dense in L_0. (This argument is standard.)

2. Application to Markovian matrices. Let $(X, \mathscr{A}, m)$ be the space of non-negative integers with counting measure. A Markovian matrix (p_{ij}) acting on the right (left) is an $L_1(L_\infty)$ operator denoted by $T(T^*)$, and $T(T^*)$ is positive, linear, of norm one. If $f = 1_{\{i\}}$, then $(T^n f)_j = p_{ij}^{(n)}$ and $(T^{*n} f)_j = p_{ji}^{(n)}$. We assume (1.9) for L_1 and T (equivalently, for L_∞ and T^*); i.e., we assume that $\sum_n p_{ij}^{(n)} = \infty$ for all i, j; indeed, it suffices to consider functions f of the form $f = 1_{\{i\}}$. In the terminology of Markov chains, (1.9) is the assumption that the matrix (p_{ij}) is *irreducible* and *recurrent* (= *persistent*). Such a matrix is called *aperiodic* iff there exist integers i_0, j_0, n_0 with $\delta = \inf(p_{i_0 j_0}^{(n_0)}, p_{i_0 j_0}^{(n_0+1)}) > 0$. Since linear combinations of functions of the form $1_{\{i\}} - 1_{\{k\}}$ are dense in L_0, Orey's theorem may be stated as follows.

THEOREM 2.1. *Let (p_{ij}) be an irreducible recurrent aperiodic matrix. Then for any integers i, k,*

$$\lim_n \sum_j |p_{ij}^{(n)} - p_{kj}^{(n)}| = 0. \tag{2.1}$$

PROOF. The theorem will follow from Corollary 1.3 if we show that for some $\varepsilon > 0$, (1.1) holds for all f of the form $f = 1_{\{i\}}$, $i = 0, 1, \cdots$. Set

$$h_i = \lim_n \sum_j |p_{ij}^{(n)} - p_{ij}^{(n+1)}| = \lim_n \|T^n 1_{\{i\}} - T^{n+1} 1_{\{i\}}\|.$$

Note that $h_{i_0} \leqq 2(1-\varepsilon_0)$ if $\varepsilon_0 = \frac{1}{2}\delta$ because T is a contraction of L_1; therefore our proof will be completed if we show that h_i is independent of i. Let $h = (h_i)_{i=0}^{\infty}$; then

$$\begin{aligned}(T^*h)_k = \textstyle\sum_i p_{ki} h_i &= \lim_n \textstyle\sum_j \sum_i p_{ki} |p_{ij}^{(n)} - p_{ij}^{(n+1)}| \\ &\geqq \lim_n \textstyle\sum_j |p_{kj}^{(n+1)} - p_{kj}^{(n+2)}| = h_k.\end{aligned} \tag{2.2}$$

Thus $T^*h \geqq h$, and applying (1.9) with $L = L_\infty$, $V = T^*$, $f = T^*h - h$, we obtain: $T^*h = h$. Since $T^*1 = 1$, Theorem 1.2 implies that h is a constant.

3. Application to Markov kernels. From now on we assume that the σ-field $\mathscr{A}$ is *separable*, i.e. generated by a countable collection of sets and $\{x\} \in \mathscr{A}$ for each $x \in X$. These assumptions are satisfied in nearly all cases of interest. A function of two variables $P(x, A)$, $x \in X$, $A \in \mathscr{A}$ is called a *Markov kernel* iff the following conditions are satisfied: For each fixed $x \in X$, $P(x, \cdot)$ is a probability measure on $\mathscr{A}$; for each fixed $A \in \mathscr{A}$, $P(\cdot, A)$ is a measurable function of x. A Markov kernel is called *m-measurable* (in terminology of E. Hopf [3]) iff $m(A) = 0$ implies $P(x, A) = 0$ a.e. on X; the exceptional null set of x's for which $P(x, A) \neq 0$ depends on A. An m-measurable Markov kernel acting on the right (left) defines a positive linear contraction operator on $L_1(L_\infty)$, denoted by $T(T^*)$. Identifying under the Radon–Nikodym isomorphism L_1 with the space of m-continuous finite signed measures φ on $\mathscr{A}$, we write:

$$T\varphi(A) = \int \varphi(dx) P(x, A) \qquad \varphi \in L_1; \tag{3.1}$$

$$T^*h(x) = \int P(x, dy) h(y) \qquad h \in L_\infty. \tag{3.2}$$

We set $P^{(1)}(x, A) = P(x, A)$ and

$$P^{(n+1)}(x, A) = \int P(x, dy) P^{(n)}(y, A). \tag{3.3}$$

Then P^n acting on the right (left) corresponds to $T^n(T^{*n})$ by relations analogous to (3.1) and (3.2). An operator T on L_1 and the kernel P which induces it are called *conservative* iff for each $f \in L_1$, $T_\infty f = 0$ or ∞ a.e. on X. A set A is called $(T-)$ *closed* iff $f \in L_1^+$, support $f \subset A \in \mathscr{A}$ implies support $Tf \subset A$. T and P are called *ergodic* iff the only closed sets are $\varnothing$ and X. The assumption that T is both conservative and ergodic is equivalent with the assumption that T^* is (see Section 1); hence it may be stated as follows: For each non-null set $A \in \mathscr{A}$, $T^*_\infty 1_A = \sum_n P^{(n)}(x, A) = \infty$ a.e. on X, the exceptional null set depending on A.

Denote the set of all finite signed measures on $\mathscr{A}$ by Φ. Remark that (3.1) defines the action of T on Φ, not only on L_1. In particular, if δ_x is a probability measure concentrated at a point x, then $T^n\delta_x = P^{(n)}(x, \cdot)$. Total variation of measures in Φ is denoted by $\|\ \|$; for measures in L_1 the total variation coincides of course with the L_1 norm. It is easy to see that T is a positive linear contraction operator on Φ endowed with the norm $\|\ \|$.

THEOREM 3.1. (zero-two theorem). *Let P be an m-measurable conservative ergodic Markov kernel, and let for $x \in X$*

$$h(x) = \lim_n \|P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot)\|. \tag{3.4}$$

Either $h(x) = 0$ a.e. *on X or $h(x) = 2$* a.e. *on X. In the first case the condition* (C) *holds: For every measure $\varphi \in L_1$ with $\varphi(X) = 0$ one has*

$$\lim_n \|T^n\varphi\| = 0. \tag{3.5}$$

PROOF. We assume that the σ-field $\mathscr{A}$ is generated by a countable collection of sets, say $A_1, A_2, \cdots$. For n fixed let $\mathscr{E}_x(\cdot)$ be the signed measure $P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot)$ and let $\mathscr{E}_x{}^k$ be the restriction of $\mathscr{E}_x$ to the σ-field $\mathscr{A}_k$ generated by the sets $A_1, A_2, \cdots, A_k$; $k = 1, 2, \cdots$. By a version of the martingale convergence theorem ([11], page 144, IV.5.3)

$$\frac{d\mathscr{E}_x{}^k}{d|\mathscr{E}_x|} \to_k \frac{d\mathscr{E}_x}{d|\mathscr{E}_x|}$$

a.e. $(|\mathscr{E}_x|)$ and in $L_1(|\mathscr{E}_x|)$, hence the sequence of measurable functions of x

$$\|\mathscr{E}_x{}^k\| = \int \left|\frac{d\mathscr{E}_x{}^k}{d|\mathscr{E}_x|}\right| d|\mathscr{E}_x|$$

converges as $k \to \infty$ to

$$\int \left|\frac{d\mathscr{E}_x}{d|\mathscr{E}_x|}\right| d|\mathscr{E}_x| = \|\mathscr{E}_x\|,$$

which is therefore measurable. We also sketch an alternate argument, making no appeal to the martingale theorem. Let H^k and H be the Hahn sets of measures $\mathscr{E}_x{}^k$ and $\mathscr{E}_x$. For k large H and H^c may be approximated in the $|\mathscr{E}_x|$-metric of symmetric differences by $\mathscr{A}_k$ measurable sets, say B_k and $B_k{}^c$. The relation

$$\mathscr{E}_x{}^k(B_k) - \mathscr{E}_x{}^k(B_k{}^c) \leqq \mathscr{E}_x{}^k(H_k) - \mathscr{E}_x{}^k(H_k{}^c) \leqq \mathscr{E}_x(H) - \mathscr{E}_x(H^c)$$

now implies that

$$||\mathscr{E}_x^k|| = \mathscr{E}_x^k(H_k) - \mathscr{E}_x^k(H_k^c) \to_k \mathscr{E}_x(H) - \mathscr{E}_x(H^c) = ||\mathscr{E}_x||.$$

The measurability of $||\mathscr{E}_x||$ is needed in the following lemma.

LEMMA 3.2. *Let for* $x \in X$ *and each* n

$$h_n(x) = ||P^{(n)}(x, \cdot) - P^{(n+1)}(x, \cdot)||. \tag{3.6}$$

Then $h(x) = \lim_n \downarrow h_n(x)$ *is constant* a.e. *on* X.

PROOF OF THE LEMMA. For each $x \in X$, $A \in \mathscr{A}$, and each n let

$$\gamma_n(x, A) = P^{(n)}(x, A) - P^{(n+1)}(x, A); \tag{3.7}$$

then (3.3) implies

$$\int P(x, dy)\gamma_n(y, A) = \gamma_{n+1}(x, A). \tag{3.8}$$

Let $H_n(x)$ be the Hahn set of the measure $\gamma_n(x, \cdot)$; we have for every $y \in X$, every $A \in \mathscr{A}$

$$h_n(y) = \gamma_n(y, H_n(y)) - \gamma_n(y, H_n^c(y)) \geqq \gamma_n(y, A) - \gamma_n(y, A^c). \tag{3.9}$$

Applying this, (3.2) and (3.8), we obtain

$$T^*h_n(x) \geqq \gamma_{n+1}(x, A) - \gamma_{n+1}(x, A^c) \qquad A \in \mathscr{A};$$

hence on substituting $A = H_{(n+1)}(x)$ we conclude that $T^*h_n \geqq h_{n+1}$. Therefore $T^*h = T^* \lim\downarrow h_n = \lim T^*h_n \geqq \lim h_{n+1} = h$. (The monotone continuity property of the operator T^* is established e.g. in [11].) $T^*h \geqq h$ implies that $T^*h = h$ because the operator $T(T^*)$ is conservative and ergodic and (1.9) may be applied to the function $T^*h - h$. From Theorem 1.2 we now conclude that h is a constant, which proves the lemma.

To prove the theorem we may and do assume that the a.e. constant function h is less than $2(1-\varepsilon)$ on a set $X' = X - N_0$, where $m(N_0) = 0$ and $\varepsilon > 0$ is a constant. Let $\varphi \in \Phi$, $||\varphi|| = 1$ and let R_n be the Hahn set of the measure $\int \gamma_n(x, \cdot)\varphi(dx)$. We have

$$\begin{aligned} ||T^n\varphi - T^{n+1}\varphi|| &= ||\textstyle\int \gamma_n(x, \cdot)\varphi(dx)|| \\ &= \textstyle\int [\gamma_n(x, R_n) - \gamma_n(x, R_n^c)]\varphi(dx) \\ &\leqq \textstyle\int [\gamma_n(x, H_n(x)) - \gamma_n(x, H_n^c(x))]\varphi(dx) \\ &= \textstyle\int ||\gamma_n(x)||\, \varphi(dx) \downarrow \int h(x)\varphi(dx) \\ &\leqq 2(1-\varepsilon) + \textstyle\int_{N_0} h(x)\varphi(dx). \end{aligned} \tag{3.10}$$

If $\varphi \in L_1$, then the last integral is zero and the assumption (1.1) of Corollary 1.3 holds. We conclude that the condition (C) holds.

To prove that $h(x)$ is zero a.e., a somewhat more involved argument is needed. Let $N_1 = \{x : P^{(n)}(x, N_0) > 0 \text{ for some } n\}$ and, assuming N_k defined, let $N_{k+1} =$

$\{x: P^{(n)}(x, N_k) > 0$ for some $n\}$; $k = 1, 2, \cdots$. Let $N = N_0 \cup N_1 \cup \cdots$ and set $X' = X - N$. Define a "sub-Markov kernel" $P'(x, A)$ on $(X \times \mathscr{A})$ by setting $P'(x, A) = P(x, A \cap X')$ for $x \in X'$; $P'(x, A) = 0$ for $x \in N$. P' has the properties of a Markov kernel except that $P'(x, X)$ is less or equal, rather than equal, to one. The m-measurability of P implies that $m(N) = 0$, and it is easy to see that $P^{(n)}(x \cdot) = P'^{(n)}(x, \cdot)$, hence $h'(x) =_{\text{def}} \lim_n \|P'^{(n)}(x, \cdot) - P'^{(n+1)}(x, \cdot)\| = h(x)$ for $x \in X'$, while for $x \in N$, $h'(x) = 0$. Let Φ' be the class of measures in Φ with support in X', and define an operator T' on Φ' by $T'\varphi(A) = \int P'(x, A)\varphi(dx)$, $\varphi \in \Phi'$. Theorem 1.1 together with its proof remains valid if L_1, T are replaced by Φ', T', provided that the infimum of two measures δ, π at a set A is defined to be $\pi(A \cap H) + \delta(A \cap H^c)$, where H is the positive Hahn set of the measure $\delta - \pi$. We conclude that $\|T'^n\varphi - T'^{(n+1)}\varphi\| < 2(1-\varepsilon)$ for $\varphi \in \Phi'$; hence by the analogue of Theorem 1.1 we obtain that $\lim_n \|T'^n\varphi - T'^{(n+1)}\varphi\| = 0$ for $\varphi \in \Phi'$, and in particular $\lim_n \|T'^n\delta_x - T'^{(n+1)}\delta_x\| = h'(x) = h(x) = 0$ for $x \in X'$. This completes the proof of the theorem.

In keeping with the "ergodic" approach of the present paper, the assumption that all powers of T are ergodic replaces below the assumption of aperiodicity. We thus avoid the discussion of the existence of a period (for which see e.g. [9]), but the two treatments are essentially equivalent.

PROPOSITION 3.3. *Let T be a positive linear conservative contraction on L_1. If* (C) *holds, then every power of T is ergodic.*

PROOF. The proposition follows easily from results about positive contractions first proved by E. Hopf ([3], see also [11], Chapter V; it does not matter that in the present paper we allow m to be σ-finite rather than finite). The relation

$$T_\infty f = (T^k)_\infty (f + Tf + \cdots + T^{k-1}f) \qquad f \in L_1{}^+$$

and the Hopf decomposition theorem applied to T^k ([11], page 196) imply that T^k is conservative for each k. For a fixed integer k, set $S = T^k$; then S preserves the integral of positive functions, and S-closed (equivalently, S^*-invariant) sets form a σ-field ([11], page 196). If S is not ergodic, then there exists a non-trivial set $A \in \mathscr{A}$ such that both A and A^c are S-closed. Let φ in L_1 be such that support $\varphi^+ \subset A$, support $\varphi^- \subset A^c$, $\varphi(A) = \varphi(A^c) = 1$. Then $\|S^n\varphi\| = 2$ for all n, which contradicts (C).

There exist numerous examples of point-transformations to show that the ergodicity of every power of a conservative operator T does *not* imply the condition (C). However, the ergodicity of every T^k renders more striking the dichotomy 0–2 in Theorem 3.1. We have indeed the following version of Theorem 3.1.

THEOREM 3.4. *Let P be an m-measurable conservative Markov kernel, and assume that every power of P is ergodic. Then either the condition* (0) *or the condition* (2) *holds*:

(0) *For every k and* a.e. $x \in X$

$$\lim_n \|P^{(n)}(x, \cdot) - P^{(n+k)}(x, \cdot)\| = 0; \qquad (3.11)$$

(2) *for every* n, k *and* a.e. $x \in X$

$$\|P^{(n)}(x, \cdot)-P^{(n+k)}(x, \cdot)\| = 2. \tag{3.12}$$

Under the alternative (0) *the condition* (C) *holds.*

PROOF. Apply Theorem 3.1 to $S = T^k$, and apply monotonicity of the sequence h_n.

It may be worthwhile to give a probabilistic interpretation of the dichotomy 0–2 in the above theorem. Let $X_0, X_1, \cdots$ be a Markov process with stationary transition probabilities $P(x, A)$ on the state space $(X, \mathscr{A})$. If the alternative (2) holds, then for a.e. $x \in X$ there exists a decomposition of the state space X into countably many disjoint sets $A_0(x), A_1(x), \cdots$, such that $X_0 = x$ implies that with probability one each X_n belongs to exactly one of the sets $A_k(x)$, say $A_{k_n}(x)$, and $n \neq n'$ implies $k_n \neq k_{n'}$.

The process (X_n) is *time-determining* in the following sense: knowing the starting state, i.e., the value of X_0, and the state after a time n, i.e., the value of X_n, we can determine the elapsed time n with probability one. The probabilistic meaning of the alternative (0) is, in a sense, opposite. Assume again that we know the value of X_0 and the value of X_n, and n is unknown but equal to one of the numbers: $N, N+1, \cdots, N+k$; k fixed. Then for N large the probabilities of events $\{n = N\}, \{n = N+1\}, \cdots, \{n = N+k\}$ are close to each other. Thus for N large the elapsed time cannot be determined.

Harris [2] introduced the following recurrence condition for a Markov process: For each $x \in X$, each event A of positive measure, the probability is one that the process starting at x will visit A infinitely often. Notice that this condition is clearly incompatible with alternative (2). Indeed, assume that (2) holds and the Harris condition is satisfied. For each x the set $\bigcup_k A_k(x)$ is of positive measure since its complement is never visited by the process; hence for some k the set $A_k(x)$ is of positive measure and is visited infinitely many times by the process, which is a contradiction. Thus the alternative (0) holds for an aperiodic Harris process. The converse is not true, as is shown by the following simple example due to the referee (the authors knew of a more complicated one). Let X be the unit circle and suppose that a particle at x remains at x with probability $\frac{1}{2}$ and goes to $x+\alpha$ (α irrational) with probability $\frac{1}{2}$. Then the process is not time-determining, hence (0) must hold, but the process is not Harris because the complement of the (countable) orbit of x will never be entered.

4. The essential Harris condition. This condition may be stated as follows (see [4], [8], and especially [5], condition (C_3)):

(H) The Markov kernel $P(x, A)$ is m-measurable, conservative and ergodic. Furthermore, for each m-null set N there exists a point $x_0 \in X-N$ and an integer $n_0 > 0$ such that the measure $P^{(n_0)}(x_0, \cdot)$ is not singular with respect to m.

The following theorem is essentially due to Jamison and Orey [6].

THEOREM 4.1. *The condition* (H) *and ergodicity of every power of P imply the condition* (C).

PROOF. Assume that for a kernel $P(x, A)$, P^k is ergodic for all $k > 0$, (H) holds and (C) fails. Then the alternative (2) in Theorem 3.4 holds, and there is an m-null set N such that for $x \in X - N$ the measures $T^n \delta_x$, $n = 0, 1, \cdots$ have mutually disjoint supports. Let x_0 and n_0 be like in (H), let π be the non-vanishing m-continuous component of the measure $P^{(n_0)}(x_0, \cdot) = T^{n_0}\delta_{x_0}$, and set $p = d\pi/dm$. Since T is conservative and ergodic, $T_\infty p = \infty$ a.e. on X, and there exists an integer $k > 0$ such that the supports of p and $T^k p$ intersect. This is a contradiction, because these supports are contained in the disjoint supports of $T^{n_0}\delta_{x_0}$ and $T^{n_0+k}\delta_{x_0}$.

In Theorem 4.1 the condition (C) may be strengthened to (C′): $\|T^n \varphi\| \to 0$ for $\varphi \in \Phi$ with $\varphi(X) = 0$. It suffices to show that under (H) the m-singular part of $T^n\varphi$ is small in norm for large n (apply e.g. Theorem 2 of [10])[3], and then to apply Theorem 4.1 to $T^n\varphi$ instead of φ. Finally, in Theorem 4.1 the assumption that the σ-field $\mathscr{A}$ is separable may be removed by consideration of "admissible" σ-fields introduced by Doob (see [6]), in view of the fact that the σ-field generated by $T^n f$, $n = 1, 2, \cdots, f$ fixed, is separable.

REFERENCES

[1] BLACKWELL, D. and FREEDMAN, D. (1964). The tail σ-field of a Markov chain and a theorem of Orey. *Ann. Math. Statist.* **35** 1291–1295.

[2] HARRIS, T. E. (1956). The existence of stationary measures for certain Markov processes. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **2** 113–124. Univ. of California Press.

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

[4] ISAAC, R. (1964). Non-singular Markov processes have stationary measures. *Ann. Math. Statist.* **35** 869–871.

[5] JAIN, N. C. (1966). A note on invariant measures. *Ann. Math. Statist.* **37** 729–732.

[6] JAMISON B. and OREY, S. (1967). Markov chains recurrent in the sense of Harris. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 41–48.

[7] KRENGEL U. and SUCHESTON, L. (1969). On mixing in infinite measure spaces. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13** 150–164.

[8] METIVIER, M. (1969). Existence of an invariant measure and an Ornstein's ergodic theorem. *Ann. Math. Statist.* **40** 79–96.

[9] MOY, S. T. C. (1967). Period of an irreducible positive operator. *Illinois J. Math.* **11** 24–39.

[10] MOY, S. T. C. (1968). The continuous part of a Markov operator. *J. Math. Mech.* **18** 137–142.

[11] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability*. Holden–Day, San Francisco.

[12] OREY, S. (1962). An ergodic theorem for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **1** 174–176.

[13] ORNSTEIN, D. S. (1969). On a theorem of Orey. *Proc. Amer. Math. Soc.* **22** 549–551.

[3] That under the Harris condition the singular part of the kernel converges to zero belongs to the folklore of the subject, and can be proved by simple probabilistic arguments. The referee points out that a purely measure-theoretic proof simpler than Mrs. Moy's is given in the recent book of S. R. Foguel, *The Ergodic Theory of Markov Processes*, Theorem E, page 59.