RECURRENT MARKOV CHAINS

S. OREY

In this paper Markov chains $\{X_i\}$, $i = 1, 2, \cdots$, which stationary transition probabilities are considered which take values in some measurable space $(S, \mathscr{B})$ and satisfy

(*) The Borel field $\mathscr{B}$ is separable and there exists a sigma finite measure m on $(S, \mathscr{B})$ such that P[entering E at some time $[X_0 = x] = 1$ for all $x \in S$ and all $E \in \mathscr{B}$ with $m(E) > 0$, where P is the underlying probability measure.

Such chains were introduced by Harris in [6], [7]. Let $P^n(x, E)$ be the n-step transition probability, $P^1(x, E) = P(x, E)$. In [7] it is proved that there exists a unique (up to constant factor) sigma finite measure Q which is *stationary* in the sense that $Q(E) = \int_S P(x, E)Q(dx)$.

Section 1 establishes some preliminary results. The relationship between (*) and Doeblin's condition is investigated. The results of Harris [6], [7] are summarized and extended. Note that many notational conventions used throughout the paper are introduced in § 1.

In § 2 it is shown that after the deletion of an inessential Q-null set the process splits up into a finite number, d, of disjoint cyclically moving classes.

Section 3 studies the asymptotic behavior of $P^n(x, \cdot)$ in case the stationary measure Q happens to be a probability measure. The approach is the "direct" approach of Markov and Doeblin and Doob [4]. It is shown that if $d = 1$, the total variation of $(P^n(x, \cdot) - Q)$ approaches 0 as n approaches ∞; for $d > 1$ the convergence statement must be modified in an obvious way. For the relationship of these results to those of [3] see the beginning of § 3.

Section 4 considers the asymptotic behavior of

$$U(n) = \sum_{N=1}^{\infty} P\left[\sum_{i=1}^{N} f(X_i) = n \right],$$

where f is a measurable function from S into the positive integers. If $\int_S f(x)Q(dx) < \infty$, $U(n)$ is for large n approximately a periodic function. The period depends both on the $\{X_i\}$ process and on f; this period may be greater than 1 even though the d associated with the $\{X_i\}$ process is 1 and $f(x) = 1$ for a set of x of positive Q-measure.

Section 5 is concerned with the behavior of normed sums,

$$\frac{1}{B_n}\sum_{i=1}^{n}f(X_i)\,,$$

where f is a real-valued measurable function on S. Neither the method of Doeblin (exploited and developed in [2]) nor that of Bernstein (used in [1]) is directly applicable to processes merely satisfying (*) and $Q(S) = 1$. Nevertheless, ideas from both those methods combined with the results of the previous sections make it possible to give conditions under which such normed sums obey the central limit theorem and even obey the Erdös-Kac-Donsker invariance principle. Results from [1] are made use of. As indicated at the relevant places, ideas of [2] and [8] are also used. The work of this section naturally leads to some problems related to recurrence times, and these are discussed in § 6.

1. **Preliminary results.** Let m be a measure satisfying (*). If ν is any measure on $(S, \mathscr{B})$ $p_\nu^n(x,\cdot)$, $P_{0\nu}^n(x,\cdot)$ are to denote, respectively, the absolutely continuous and singular component of P^n with respect to ν. Superscripts will be omitted when $n = 1$. It will be assumed that p_ν^n is measurable in (x, y) in the product space $S \times S$; this assumption is justified by Doob [4], p. 616. It will also be assumed that for all positive integers s and t

$$p_\nu^{s+t}(x, y) \geqq \int_S p_\nu^s(z, y)p_\nu^t(x, z)\nu(dz)$$

holds for all x and y; that such a choice of densities is always possible was shown in Doob [4], p. 146.

When the subscript ν is omitted it will be assumed in this section that ν is m; in all the other sections the omission of ν will mean $\nu = Q$.

If A is a Borel set, $m(A) > 0$, " the process on A " will have the same significance as in [7], i.e., if $X_{A0}, X_{A1}, \cdots$ are the successive members of the sequence $X_0, X_1, \cdots$ with values in A, $\{X_{Ai}\}$, $i = 0, 1, \cdots$ is the process on A. This is a Markov process with transition probability

$$(1.1) \qquad P_A(x, E) = P(x, E) + \int_{S-A} P(x, dy)P(y, E)$$
$$+ \int_{S-A}\int_{S-A} P(x, dy)P(y, dz)P(z, E) + \cdots$$

for every Borel subset E of A and $x \in A$.

The process on A will also satisfy (*). Notions defined for the original process can thus be relativized to A; notationally this is indicated by a subscript A, e.g., $p_A^n(x, y)$ is defined like $p^n(x, y)$ but using $P_A(x, E)$ in place of $P(x, E)$.

If $0 < m(A) < \infty$, v belongs to the open interval $(0, 1)$, and j is a positive integer, $K(A, v, j)$ is to be the set of all $x \in A$ such that

$$m\left\{y \in A : \sum_{i=1}^{j} p^i(x, y) > \frac{1}{j}\right\} > vm(A) \ .$$

By the convention explained above $K_C(A, v, j)$ is meaningful whenever C is a Borel set including A (the definition now involves p_C^i in place of p^i). Harris observed that if $m(A) < \infty$,

$$A = \bigcup_{k=1}^{\infty} K(A, r, k)$$

easily follows from (*), for every $r \in (0, 1)$. Lemma 1.1 is a modification of a lemma in [7].

LEMMA 1.1. *Let* A, B *be Borel sets such that* $m(B) > 0$, $m(A) < \infty$, $B \subseteq K(A, v, k)$ *for some* $v \in (0, 1)$ *and some positive integer* k. *Then* $K(B, r, k) = B$ *provided* $(1 - r) \geq (1 - v)m(A)/m(B)$, *where* $r \in (0, 1)$.

Proof. For every $x \in B$,

$$m\left\{y \in B : \sum_{i=1}^{k} p^i(x, y) < \frac{1}{k}\right\} < (1 - v)m(A)$$

$$= \left[(1 - v)\frac{m(A)}{m(B)}\right]m(B) < (1 - r)m(B)$$

if the proviso of the lemma holds.

LEMMA 1.2. *Let* $A = K(A, r, k)$, $r \in (0, 1)$, k *a positive integer,* A *a Borel set of finite m-measure. Then there exists a probability measure* φ *on* A *and numbers* $\alpha, \eta \in (0, 1)$ *such that* $P^{k+1}(x, E) \leq 1 - \eta$ *whenever* $E \subseteq A$, $E \in \mathscr{B}$, $\varphi(E) \leq \alpha$ *and* $x \in A$.

Proof. Let η be any number such that $0 < \eta < rm(A)k^{-3}$. Let x be some point in A. Define

$$C = \left\{z \in A : \sum_{i=1}^{k} p^i(x, z) > \frac{1}{k}\right\}, \ C_i = \left\{z \in C : p^i(x, z) > \frac{1}{k^2}\right\},$$

Then

$$C \subseteq \bigcup_{i=1}^{k} C_i \text{ and } m(C) > rm(A) \ .$$

Let E be a Borel subset of A, and suppose $P^{k+1}(x, E) > 1 - \eta$. Then

$$\eta > 1 - P^{k+1}(x, E) = P^{k+1}(x, S - E) \geq \frac{1}{k}\sum_{i=1}^{k}\int_{C_i} p^i(x, y)P^{k+1-i}(y, S - E)m(dy)$$

$$\geqq \frac{1}{k^3} \sum_{i=1}^{k} \int_{c_i} P^{k+1-i}(y, S-E) m(dy) = \frac{1}{k^3} \sum_{i=1}^{k} \int_{c_i} (1 - P^{k+1-i}(y, E)) m(dy)$$

$$\geqq \frac{1}{k^3} \left[r \cdot m(A) - \sum_{i=1}^{k} \int_{c_i} P^{k+1-i}(y, E) m(dy) \right]$$

$$\geqq \frac{1}{k^3} \left[r \cdot m(A) - k \int_A \sum_{i=1}^{k} P^i(y, E) m(dy) \right].$$

Let

$$\varphi(E) = \frac{1}{k m(A)} \int_A \sum_{i=1}^{k} P^i(y, E) m(dy) \text{ and } \alpha = \frac{r m(A) - k^3 \eta}{k^2 m(A)}$$

The inequality above yields $\varphi(E) > \alpha$, proving the lemma.

COROLLARY 1.2.A. *Let A, C be Borel sets, $m(C) < \infty$, $A \subseteqq C$ and $A = K_C(A, r, k)$, where $r \in (0, 1)$ and k is a positive integer. Then there exists a probability measure φ on A and numbers $\alpha, \eta \in (0, 1)$ such that $P_C^{k+1}(x, E) \leqq 1 - \eta$ whenever $x \in A$, and E is a Borel subset of A with $\varphi(E) \leqq \alpha$.*

Proof. Since the process on C also satisfies (*) the lemma applies to it, yielding the corollary.

COROLLARY 1.2B. *If A, C, r, k are as in Corollary 1.2A the process on A satisfies Doeblin's condition.*[1]

Proof. $A = K_A(A, r, k)$ since $K_C(A, r, k) \subseteqq K_A(A, r, k)$. So the conclusion of Corollary 1.2A applies with P_A for P_C; but this gives Doeblin's condition for the process on A.

Let $\mathscr{D}$ be the collection of $A \in \mathscr{B}$ such that the process on A satisfies Doeblin's condition.

It will be seen that $\mathscr{D}$ is an important collection. In [6] Harris announced a result which, slightly extended, asserts that when (*) holds one has for all $x, y \in S$

(1.2) $$\frac{\sum_{n=0}^{N} P^n(x, E)}{\sum_{n=0}^{N} P^n(y, F)} \longrightarrow \frac{Q(E)}{Q(F)} \text{ and } N \longrightarrow \infty$$

for all Borel sets E, F with $E \subseteqq F$, $F \in \mathscr{D}$. In this connection see also Theorem 1. In [7] the question of more general validity of (1.2) was raised. If merely $E, F \in \mathscr{B}$, $0 < m(F) < \infty$ is assumed, (*) does not

---

[1] It is actually Doob's generalization of Doeblin's condition that is referred to. See [4], p. 192, Hypothesis (D).

imply (1.2) for all $x, y \in S$. It is easy to give examples of chain satisfying (*) and with denumerable state space where (1.2) is violated even in the special case where $x = y$ or where $E = F$. On the other hand, when S is denumerable one shows easily that, regardless of whether or not the process satisfies (*), one has for $x \neq y$,

$$P[X_i = y \text{ for some positive } i \,|\, X_0 = x] \leqq \liminf_{N \to \infty} \frac{\sum\limits_{n=0}^{N} P^n(x, E)}{\sum\limits_{n=0}^{N} P^n(y, E)}$$

$$\leqq \limsup_{N \to \infty} \frac{\sum\limits_{n=0}^{N} P^n(x, E)}{\sum\limits_{n=0}^{N} P^n(y, E)} \leqq \frac{1}{P[X_i = x \text{ for some positive } i \,|\, X_0 = y]} \,.$$

It follows that (1.2) will hold for every pair of x and y belonging to the same recurrence class. In particular in case (*) holds and S is denumerable (1.2) will hold for all x and y outside a fixed Q-null set. It would be interesting if this could be shown to hold even when S is not denumerable.

Harris showed in [7] that if $A = K_A(A, r, k)$ the process on A has a stationary probability measure; this also follows from Corollary 1.2B. Whenever the process on some Borel set B has a stationary probability measure it will be denoted by Q_B.

LEMMA 1.3.   If $A, B \in \mathscr{B}$, $B \subseteq A$, $A \in \mathscr{D}$ then $B \in \mathscr{D}$.

Proof. Assume the hypotheses of the lemm. Then there is a positive function $\varepsilon(n)$ such that

$$\sum_{n=1}^{\infty} \varepsilon(n) < \infty$$

and $P_A^n(x, E) < Q_A(E) + \varepsilon(n)$ for every Borel subset E of A. Let N be an integer such that

$$\sum_{n=N}^{\infty} \varepsilon(n) < \frac{1}{4} \,.$$

Then there exists an integer M such that for all $x \in B$ and Borel subsets E of B,

$$P_B^N(x, E) = \sum_{i=N}^{\infty} P[X_{BN} \in E, X_{BN} = X_{Ai} \,|\, X_{A0} = x]$$

$$< \sum_{i=N}^{M} P[X_{BN} \in E, X_{BN} = X_{Ai}] + \frac{1}{4} < \sum_{i=N}^{M} [Q_A(E) + \varepsilon(i)] + \frac{1}{4} \,.$$

So when $Q_A(E) < 1/(4(M - N))$, $P_B^n(x, E) < 3/4$ proving the lemma.

LEMMA 1.4.   *Let* $0 < r < 1$, $A \in \mathscr{D}$ *and* $m(A) < \infty$. *Then* $A \subseteqq (A, r, k)$ *for some* k.

Proof.   If the lemma is false there is for every k an $x_k \in A$ such that if

$$E_k = \left\{ y \in A : \sum_{i=1}^{k} p^i(x_k, y) \leqq \frac{1}{k} \right\},$$

$m(E_k) > (1 - r)m(A)$.   Let E_k' satisfy $E_k' \subseteqq E_k$, $m(E - E_k') = 0$, and $P_0^j(x_k, E_k') = 0$, $j = 1, \cdots, k$.   So

$$\sum_{j=1}^{k} P^j(x_k, E_k') \leqq \frac{1}{k} m(E_k') \leqq \frac{1}{k} m(A) ,$$

and the last term approaches zero as k approaches ∞.   Therefore, $Q_A(E_k)$ approaches zero.   Since Q_A and m are finite on A and m is absolutely continuous with respect to Q_A, $m(E_k)$ must tend to zero, which results in a contradiction.

The restriction $m(A) > 0$ or $m(A) < \infty$ appeared frequently above. Note that there always exists finite measures q having the same null sets as Q and therefore satisfying (*).   If such a q is chosen for m in the preceding lemma the hypothesis $m(A) < \infty$ may be dropped and the conclusion may be weakened to $A \subseteqq K(S, r, k)$ for some k.   Letting $S_k = K(S, r, k)$, where r is fixed, $0 < r < 1$, the preceding sentence can be restated thus: $A \in \mathscr{D}$ implies $A \subseteqq S_k$ for some positive integer k. Clearly $S_k \subseteqq S_{k+1}$, $k = 1, 2, \cdots$.   By the remark preceding Lemma 1.1 $S = \bigcup_{i=1}^{\infty} S_k$.   Lemma 1.1 asserts that

$$K(S_k, r', k) = S_k \text{ if } (1 - r') \geqq (1 - r)q(S)/q(S_k).$$

For k sufficiently big such a choice of r' will be possible, since S_k approaches S.   So then by Corollary 1.2B such S_k belong to $\mathscr{D}$.   Now by Lemma 1.3 all S_k and all their Borel subsets belong to $\mathscr{D}$.   This proves the following theorem.

THEOREM 1.   *If* (*) *holds* S *can be represented as a union of Borel sets* S_i, $i = 1, 2, \cdots$ *such that* $S_i \subseteqq S_{i+1}$ *and a Borel set* A *belongs to* $\mathscr{D}$ *if and only if* $A \subseteqq S_k$ *for some* k.

Harris showed that if A is a Borel set such that the process on A has a stationary probability measure Q_A, Q_A can be extended to a stationary sigma finite measure on S, $\overline{Q}_A$. $\overline{Q}_A(E)$ is the expected number of visits to E up to and including the first return to A if the process starts with the initial distribution Q_A.   Analytically

(1.3) $$\overline{Q}_A(E) = \int_A Q_A(dx) P_A(x, E)$$

where $P_A(x, E)$ is defined by (1.1) (regardless of whether $E \subseteqq A$ or not).

The process $\{X_i\}$ is determined by the function $P(x, E)$ and the initial distribution. It will sometimes be convenient to indicate the initial distribution as a subscript on the expectation operator E, or on P.

Let $A \in \mathscr{B}$, and suppose $\{X_{Ai}\}$ possesses a stationary probability measure Q_A. Let V be the least positive integer such that $X_V \in A$. It is not hard to verify that if f is any measurable function from S into the positive integers one has

$$(1.4) \qquad E_{Q_A}\left\{\sum_{i=1}^{V} f(X_i)\right\} = \int_S f(x) \overline{Q}_A(dx)$$

where both sides are infinite if either one is.[2]

2. Cyclic decomposition.

In this section it is shown that the arguments applied by Doob in [4] to processes satisfying Doeblin's condition can be extended to the case where only (*) is assumed. In particular there exists a Borel set C with positive Q measure such that g.l.b. $\underset{x, y \in C}{p^a(x, y)} > 0$ for some positive integer α. This leads at once to the desired decomposition. The only place where it is necessary to deviate from the treatment of [4] is in the proof of Lemma 2.1 below (Lemma 5.3, p. 200 of [4]).

LEMMA 2.1. *If* (*) *holds there exist A, $B \in \mathscr{B}$ and a positive integer n such that $Q(A) > 0$, $Q(B) > 0$ and g.l.b. $\underset{\substack{x \in A \\ y \in B}}{p^n(x, y)} > 0$.*

Proof. Let $D \in \mathscr{B}$ satisfy

$$Q\left\{y \in D: \sum_{i=1}^{k} p^i(x, y) > \frac{1}{k}\right\} > r > 0 ,$$

for all $x \in D$. By section 1 such r, D, k exist. Then there exists a $D_1 \in B$ and a positive integer n_1 such that $D_1 \subseteqq D$, $Q(D_1) > 0$, and

$$Q\left\{y: p^{n_1}(x, y) > \frac{1}{k^2}\right\} > \frac{r}{k}$$

for all $x \in D_1$. Also there must exist $D_2 \in B$, and a positive integer n_2 such that $D_2 \subseteqq D_1$, $Q(D_2) > 0$ and for all $x \in D_2$

$$Q\left\{y: p^{n_1}(x, y) > \frac{1}{k^2} \text{ and } Q\left\{z \in D: p^{n_2}(y, z) > \frac{1}{k^2}\right\} > \frac{r}{k}\right\} > \frac{r}{k^2} .$$

Let

---

[2] This is part of the assertion of Lemma 6. The discrete analogue of this formula is formula (A) of the appendix to [**2**]. Cf. also footnote 7.

$$\hat{H}_1 = \Big\{ (x, y): x \in D_2,\ p^{n_1}(x, y) > \frac{1}{k^2} \text{ and } Q\Big\{ z \in D: p^{n_2}(y, z) > \frac{1}{k^2} \Big\} > \frac{r}{k} \Big\} ,$$

and

$$\hat{H}_2 = \Big\{ (y, z) \in D \times D: \text{ there exists an } x \text{ such that } (x, y) \in \hat{H}_1$$
$$\text{and } p^{n_2}(y, z) > \frac{1}{k^2} \Big\} .$$

Let $\hat{Q}$ be the product measure $Q \times Q$ in the product space $D \times D$. Given any $\hat{Q}$-null set N of $D \times D$ it is clear that it is always possible to choose two points $(x_0, y_0), (x_1, y_1)$ in $(D \times D) - N$ such that $(x_0, y_0) \in \hat{H}_1$, $(x_1, y_1) \in \hat{H}_2$ and $y_0 = x_1$. From here on the proof follows that of [4], middle of p. 201. Since the ν of Doob corresponds sometimes to n_1 and sometimes to n_2 in an obvious way, the conclusion here will be that $p^{n_1 + n_2}(x, y)$ is bounded away from zero for $x \in A$, $y \in B$. The lemma follows if $n = n_1 + n_2$.

THEOREM 2.1. *If* (*) *holds there is a* $C \in \mathscr{B}$, $Q(C) > 0$ *and a positive integer* α *such that* $\operatorname*{g.l.b.}_{x,y \in C} p^\alpha(x, y) > 0$.

Proof. This follows easily from the preceding lemma. For details see [4], Lemma 5.4.

Let C satisfy

$$\operatorname*{g.l.b.}_{x,y \in C} p^\alpha(x, y) > 0 .$$

It is known (cf. [4], p. 202) that if d is the greatest common divisor of

$$I(C) = \{ \alpha: \operatorname*{g.l.b.}_{x,y \in C} p^\alpha(x, y) > 0 \}$$

then all sufficiently large multiples of d belong to $I(C)$. With no more essential variations from the development as given in [4] one obtains:

THEOREM 2.2. *Suppose* (*) *holds. Then there exists a unique integer* d *such that whenever* $C \in \mathscr{B}$, $Q(C) > 0$ *and* $I(C)$ *non-void then* d *is the greatest common devisor of* $I(C)$. *There exists a partition of* S *into Borel sets* $C_0, C_1, \cdots, C_{d-1}, F$, *such that* $Q(F) = 0$ *and for* $x \in C_i$ $P(x, C_{i+1}) = 1$, *where the subscripts are integers modulo* d.

3. Convergence of $P^n(x, E)$ when Q is finite. In this section it is assumed that (*) holds and that Q is a probability measure. The basic method used here goes back to Markov and Doeblin; in detail, however, this presentation leans on Chapter V of Doob [4].

In [3] Doob investigated Markov chains possessing stationary probability measures. To see that under the assumptions of this section

Theorem 5 of [3] is applicable one needs only to check that $P_0^n(x, S) \to 0$ as $n \to \infty$ for all $x \in S$. For (cf. [3], p. 409) if for $m = 1, 2, \cdots, A_m$ is a Q-null set such that $P_0^m(x, S - A_m) = 0$,

$$P[X_n \in A_n \,|\, X_0 = x] = P[X_1 \in A_1, X_2 \in A_2, \cdots,$$
$$X_n \in A_n \,|\, X_0 = x] \;.$$

So if $P[X_n \in A_n \,|\, X_0 = x] > \varepsilon > 0$ for all n,

$$P\left[\text{ever entering } S - \bigcup_{i=1}^{\infty} A_i \,|\, X_0 = x \right] < 1 - \varepsilon \;,$$

contradicting (*). Theorem 3 below, however, shows somewhat more than would follow from an application of Theorem 5 of [3] since it proves, say when $d = 1$, that the variation of $(P^n(x, \cdot) - Q)$ approaches 0 when n approaches ∞ while [3] gives only $(P^n(x, E) - Q(E))$ approaches 0 as n approaches ∞ for every $E \in \mathscr{B}$.

If φ is a totally additive set function on S, $\|\varphi\|$ is to be the total variation of φ on S. $T\varphi$ $(T^n\varphi)$ will denote the measure

$$\int_S P(x, \cdot)\varphi(dx) \left(\int_S P^n(x, \cdot)\varphi(dx) \right) \;.$$

LEMMA 3.1. *Suppose there exists a Borel set C of positive Q-measure, an $\varepsilon > 0$ and for every $x, y \in S$ a positive integer $w(x, y)$ such that $p^w(x, z) > \varepsilon$, $p^w(y, z) > \varepsilon$ for all $z \in C$. Then for any two probability measures φ_1, φ_2 on S there is an n such that $\|T^n\varphi_1 - T^n\varphi_2\| \leq (1 - \varepsilon Q(C)/2)\|\varphi_1 - \varphi_2\|$.*

Proof. Note that $\|T^n\varphi_1 - T^n\varphi_2\|$ is nonincreasing in n. Let ρ_x be the measure such that $\rho_x(E) = 1$ (0) if $x \in E$ ($x \in S - E$), $E \in \mathscr{B}$. Consider first the case $\varphi_2 = \rho_x$, and write φ for φ_1. There exists a unique real number α and measure ν such that $\varphi = \nu + \alpha\rho_x$ and $\nu(\{x\}) = 0$. Let $\rho = (1 - \alpha)\rho_x$. Let $A_i = \{y : w(x, y) = i\}$. So the A_i are disjoint and $\bigcup_1^{\infty} A_i = S$. Let

$$\varepsilon' = \frac{\varepsilon Q(C)(1 - \alpha)}{4}$$

and choose N so large that

$$\nu\left(\bigcup_{i=N+1}^{\infty} A_i \right) < \varepsilon' \;.$$

Define the measures ν_i, ρ_i by $\nu_i(E) = \nu(E \cap A_i)$ and $\rho_i = \|\nu_i\|\rho_x$. Then

$$\left\| \nu - \sum_{i=1}^{N} \nu_i \right\| < \varepsilon', \left\| \rho - \sum_{i=1}^{N} \rho_i \right\| < \varepsilon' \;.$$

Now

$$\left\| T^N\!\Big(\sum_{i=1}^{N}\nu_i\Big) - T^N\!\Big(\sum_{i=1}^{N}\rho_i\Big)\right\| \leqq \sum_{i=1}^{N} \| T^N\nu_i - T^N\rho_i \| \leqq \sum_{i=1}^{N} \| T^i\nu_i - T^i\rho_i \| .$$

From the definition of ν_i and ρ_i it follows by an obvious argument that the densities of $T^i\nu_i$, $T^i\rho_i$ with respect to Q are equal at least to $\varepsilon\|\nu_i\|$ everywhere on C, from which it follows immediately that there is enough cancellation to insure $\| T^i\nu_i - T^i\rho_i \| \leqq 2\|\nu_i\|(1 - \varepsilon Q(C))$. So

$$T^N\nu - T^N\rho \| \leqq \left\|\sum_{i=1}^{N} T^N\nu_i - \sum_{i=1}^{N} T^N\rho_i\right\| + 2\varepsilon' \leqq 2\|\nu\|\Big(1 - \frac{\varepsilon Q(C)}{2}\Big),$$

proving the lemma for this special case.

Now let φ_1 and φ_2 be arbitrary probability measures. Let $\mu_1(\mu_2)$ be the positive (negative) variation of $\varphi_1 - \varphi_2$. So $\varphi_1 - \varphi_2 = \mu_1 - \mu_2$. Let α_i, ν_i be the unique real number and measure such that $\mu_i = \nu_i - \alpha_i\rho_x$, $\nu_i\{x\} = 0, i = 1, 2$. Let $\rho = \|\mu_1\|\rho_x$. For all big enough n and $i = 1, 2$ one has by the above:

$$\| T^n\mu_i - T^n\rho \| \leqq \|\mu_i - \rho\|\Big(1 - \frac{\varepsilon}{2}Q(C)\Big) \leqq 2\|\mu_i\|\Big(1 - \frac{\varepsilon Q(C)}{2}\Big).$$

So

$$\| T^n(\varphi_1 - \varphi_2) \| = \| T^n(\mu_1 - \mu_2) \| \leqq \| T^n\mu_1 - T^n\rho \|$$
$$+ \| T^n\mu_2 - T^n\rho \| \leqq 2(\|\mu_1\| + \|\mu_2\|)\Big(1 - \frac{\varepsilon Q(C)}{2}\Big)$$
$$= 2\|\varphi_1 - \varphi_2\|\Big(1 - \frac{\varepsilon Q(C)}{2}\Big) \ .$$

LEMMA 3.2.   *Assume* (*), *and Q a probability measure and the hypotheses of Lemma 3.1. Then $\| T^n\varphi - Q\|$ approaches 0 as n approaches ∞ for every probability measure φ.*

Proof. Assume the hypotheses of the lemma, and let φ be a probability measure. For every n one can find an m such that

$$\| T^m\varphi - T^mQ \| < \Big(1 - \frac{\varepsilon Q(C)}{2}\Big)^n,$$

by repeated applications of the previous lemma. Since $\| T^k\varphi - T^kQ \|$ is nonincreasing in k and $T^kQ = Q$ for all k the lemma follows.

LEMMA 3.3.   *Assume* (*) *and that $d = 1$ in the decomposition of Theorem 3.2. Then there is an $\varepsilon > 0$, a Borel set C of positive Q-measure,*

and for every $x, y \in S$ *a positive integer* w *such that* $p^w(x, z) > \varepsilon$, $p^w(yz) > \varepsilon$
for all $z \in C$.

Proof. Assume the hypotheses of the lemma. Let C be a set satis-
fying Theorem 2.1, so that $\underset{x,y \in C}{\text{g.l.b.}} \ p(x, y) > 0, Q(C) > 0$. Suppose first
that
(3.1) there exists a positive integer N, a $\delta > 0$, and for every $x, y \in S$
an integer γ such that

$$\sum_{i=\gamma+1}^{\gamma+N} P^i(x, C) > \delta, \ \sum_{i=\gamma+1}^{\gamma+N} P^i(y, C) > \delta \ .$$

Let M be a positive integer such that

$$\underset{x,y \in C}{\text{g.l.b.}} \ p^n(x, y) > 0$$

for all $n > M$. By (3.1) there exist positive integers $\beta(x), \beta(y)$ such that
$\beta(x) < N$ and $P^{\gamma+\beta(x)}(x, C) > \delta/N$ and $\beta(y) < N$ and $P^{\gamma+\beta(y)}(y, C) > \delta/N$.
Then for $z \in \{x, y\}$, $u \in C$ one has

$$p^{\gamma+N+M}(z, u) \geqq \int_C p^{N+M-\beta(z)}(v, u) P^{\gamma+\beta(z)}(z, dv)$$

$$\geqq \frac{\delta}{N} \ \underset{v,u \in C}{\text{g.l.b.}} \ p^{N+M-\beta(z)}(v, u) \geqq \frac{\delta}{N} \ \underset{M \leq n \leq N+M}{\min} \ \underset{\substack{v \in C \\ u \in C}}{\text{g.l.b.}} \ p^n(v, u) \ .$$

Thus to prove the lemma it suffices to prove (3.1).
Let δ' be a positive number less than 1. Let

$$A_n = \left\{ z : \sum_{i=1}^{n} P^i(z, C) > \delta' \right\} \ .$$

Clearly $A_n \subseteqq A_{n+1}$ and by (*) $Q(A_n) \to 1$ as $n \to \infty$. Let $Q(A_N) > 3/4$.
Under the present hypotheses the ergodic theorem shows that for *every*
x and y there is an n_0 such that for $n > n_0$ P[number of visits to A_N
in n steps $> 3n/4 | X_0 = z] > 3/4$ for $z \in \{x, y\}$. Let n be an integer
greater than n_0 divisible by four. Let $a_{ik} (b_{ik}) = P$[entering A_N for the
kth time at ith step $| X_0 = z], z = x \ (z = y)$. For $k = 1, 2, \cdots, 3n/4$,

$$\sum_{i=1}^{n} a_{ik} > \frac{3}{4}, \ \sum_{i=1}^{n} b_{ik} > \frac{3}{4} \ .$$

Let

$$a_i = \sum_{k=1}^{3/4n} a_{ik}, \ b_i = \sum_{k=1}^{3/4n} b_{ik} \ .$$

Then

$$\sum_{i=1}^{n} a_i > \frac{9}{16} n, \sum_{i=1}^{n} b_i > \frac{9}{16} n \ .$$

Now $a_i (b_i) \leqq P[\text{entering } A_N \text{ at step } i \,|\, X_0 = z] \leqq 1, \ z = x \ (z = y)$. This together with

$$\sum_{i=1}^{n} (a_i + b_i) > \frac{9}{8} n$$

proves that for some $\gamma \leqq n$ both $a_\gamma > 1/8, b_\gamma > 1/8$ holds. So $P^\gamma(x, A_N) > 1/8$, $P^\gamma(y, A_N) > 1/8$. Then

$$\sum_{i=\gamma+1}^{\gamma+N} P^i(z, C) > \int_{A_N} \sum_{i=1}^{N} P^\gamma(z, dv) P^i(v, C) > \frac{1}{8} \delta' \text{ for } z \in \{x, y\} \ .$$

This proves (3.1) with $\delta = \delta'/8$.

THEOREM 3.   *Assume* (*) *and that a stationary probability measure* Q *exists.*

If $d = 1$ *in Theorem 2.2* $\|P^n(x, \cdot) - Q\|$ *approaches* 0 *as* n *approaches* ∞.

More generally, if $d \geqq 1$, *and* $Q_k(A) = Q(A \cap C_k)$, $k = 0, 1, \cdots d - 1$, *for all* $A \in \mathscr{B}$, *and* d, C_k *as in theorem 2.2, and if for any initial distribution* φ *one defines* $\alpha_i(\varphi) = \lim_{n \to \infty} P_\varphi[X_{nd} \in C_i]$, $i = 0, \cdots, d-1$, *one has*

$$\left\| T^{nd+k} \varphi - d \sum_{i=0}^{d-1} (\alpha_i(\varphi) Q_{k+i}) \right\| \text{ approaches } 0 \text{ as } n \text{ approaches } \infty \ ,$$

where the subscripts are integers modulo d, $k = 0, 1, \cdots, d - 1$.

Proof.   The first assertion follows from Lemmas 3.2 and 3.3. The reduction of the second assertion to the first one is trivial.

Obviously $P^n(x, \cdot)$ cannot converge if Q is not finite. It may be conjectured that in this case $P^n(x, E) \to 0$ whenever $Q(E) < \infty$; such a result would be very useful. So far no proof of this conjecture has been found, not even under the additional hypothesis $E \in \mathscr{D}$.

4. A renewal theorem.   Let $\{X_i\}$ satisfy (*) and assume $\{X_i\}$ has a stationary probability measuae Q. Let f be a measurable function from S to the positive integers. Let

$$U_\varphi(n) = \sum_{N=1}^{\infty} P_\varphi \left[\sum_{i=1}^{N} f(X_i) = n \right] ,$$

where φ is the initial probability distribution.

Let $\tilde{S} = \{(x, k): x \in S, k \text{ a positive integer}\}$. Let $\tilde{\mathscr{B}}$ be the smallest Borel field such that for every $E \in \mathscr{B}$ and every positive integer k $\{(x, k): x \in E\} \in \tilde{\mathscr{B}}$. Let $\{\tilde{X}_i\}$ be a Markov process with values in $(\tilde{S}, \tilde{\mathscr{B}})$ having the following transition probabilities:

$$\tilde{P}((x, i), \{(x, i + 1)\}) = 1 \text{ for } i = 1, 2, \cdots, f(x) - 1 \text{ and } x \in S \text{ ,}$$
$$\tilde{P}((x, f(x)), E'\} = P(x, E) \text{ if } E' = \{(z, 1): z \in E\}, E \in \mathscr{B} \text{ ,}$$
$$\tilde{P}((x, i), \{(x, 1)\}) = 1 \text{ for } i > f(x) \text{ .}$$

For $E \in \mathscr{B}$ let $E^0 = \{(z, f(z)): z \in E\}$. Let $\tilde{\varphi}$ be the probability measure on $(\tilde{S}, \tilde{\mathscr{B}})$ defined by $\tilde{\varphi}(E^0) = \varphi(E), \tilde{\varphi}(\tilde{S} - S^0) = 0$.

Note that one has

(4.1) $$U_\varphi(n) = \tilde{P}_{\tilde{\varphi}}[X_n \in S^0] \text{ .}$$

Assume now that

$$\int_S f(x)Q(dx) < \infty \text{ .}$$

Since $P(x, E) = \tilde{P}_{S^0}((x, f(x)), E^0)$ for all $x \in S$ and $E \in \mathscr{B}$, $\{X_{S_i^0}\}$ has a stationary probability measure $\tilde{Q}_{S^0}$. $\{\tilde{X}_i\}$ satisfies (*), so $\tilde{Q}_{S^0}$ can be extended to a stationary measure $\tilde{Q}$ for $\{\tilde{X}_i\}$. From (1.1), (1.3), (1.4) it follows that

$$\tilde{Q}(\tilde{S}) = \int_S f(x)Q(dx) \text{ ,}$$

which was assumed finite. So

$$\overline{Q}(\cdot) = \frac{Q(\cdot)}{\displaystyle\int_S f(x)Q(dx)}$$

is a stationary probability measure for $\{\tilde{X}_i\}$. Apply Theorem 2.2 to $\{\tilde{X}_i\}$, obtaining an integer d and classes $C_1, \cdots, C_d, F$. Let

$$\alpha_i(\tilde{\varphi}) = \lim_{m \to \infty} P_{\tilde{\varphi}}[\tilde{X}_{m \cdot i} \in C_i] \text{ .}$$

Then

$$U(nd + j) \text{ approaches } d \sum_{i=0}^{d-1} \alpha_i(\overline{\varphi})\overline{Q}(S^0 \cap C_{j+1})$$

by (4.1) and Theorem 3. Let $C_i' = \{x \in S; (x, f(x)) \in C_i\}$ and let $\tilde{\alpha}_i(\varphi) = \alpha_i(\tilde{\varphi})$ for $i = 0, 1, \cdots, d - 1$. Then clearly

(4.2) $\tilde{\alpha}_i(\varphi) = \lim\limits_{n\to\infty} P_\varphi\Big(\bigcup\limits_{j=0}^{d-1}\Big[X_n \in C'_{i+j}, \sum\limits_{i=1}^{n} f(X_i) \equiv j(\mathrm{mod}\ d)\Big]\Big)$

where the subscript on the C' is an integer modulo d.

These arguments establish the following theorem.

THEOREM 4. *Let $\{X_i\}$ satisfy (*) and assume there is a stationary probability measure Q for X_i. Assume $\int_S f(x)Q(dx) < \infty$. Then S can be partitioned into Borel sets $C'_1, \cdots, C'_d, F'$ such that $P(x, F') = 0$ for all $x \in S - F'$, $Q(F') = 0$, and for any initial distribution φ of the $\{X_i\}$ process one has*

$$\lim_{n\to\infty} U_\varphi(nd + j) = \frac{d \sum\limits_{i=0}^{d-1} \tilde{\alpha}_i(\varphi)Q(C'_{j+1})}{\int_S f(x)Q(dx)}, \quad j = 0, 1, \cdots, d - 1\ ,$$

where $\tilde{\alpha}_i(\varphi)$ is defined by (4.2) and the subscripts are integers modulo d.

If the $\{X_i\}$ are independent (and then automatically identically distributed), $S = C'_0$ and the theorem yields

$$U_\varphi(nd) \to \frac{d}{\int_S f(x)Q(dx)} = \frac{d}{E\{f(X_i)\}}\ ;$$

in this case $U_\varphi(nd + j) \equiv 0$ for $j = 1, \cdots, d - 1$. This is a result of [5].

It seems plausible that $U_\varphi(n)$ approaches 0 as n approaches ∞ in case

$$\int_S f(x)Q(dx) = \infty\ .$$

This would follow from a proof of the conjecture made at the end of the previous section.

5. The invariance principle. Let f be a real-valued measurable function on the state space of the Markov chain $\{X_i\}$. Under certain conditions the sequence of sums

$$\sum_{i=1}^{n} f(X_i)$$

is known to behave like a sequence of partial sums of independent random variables, e.g., the central limit theorem holds for suitable norming constants. Two devices are available for proving results of this kind. The first method, which is due to S. Bernstein, uses the fact that in certain cases the dependence of X_{n+k} on $X_1, \cdots, X_n$ diminishes quickly as k increases; this method is applicable if $\{X_i\}$ satisfies Doeblin's condition. The second method is applicable to certain cases in which the

state space is denumerable; the idea in this case, due to Doeblin, is that if V_i is defined to be the ith nonnegative integer n such that $X_n = x$, where x is a fixed state, the sums $f(X_{V_i+1}) + \cdots + f(X_{V_{i+1}})$, $i = 1, 2, \cdots$, are independent and identically distributed. If $\{X_i\}$ is merely assumed to satisfy (*) and have a stationary probability measure Q neither method applies. However, it will be shown below that a combination of the two methods may be used in this case.

Assume now that $\{X_i\}$ satisfies (*). Let $A \in \mathscr{B}$ have positive Q-measure. Let V_i (or $V_i(A)$) be the ith nonnegative integer v such that $X_v \in A$, i.e., $X_{V_i} = X_{A(i-1)}$, $i = 1, 2, \cdots$. Let Y_i (or $Y_i(A)$) be the vector $(X_{V_i+1}, X_{V_i+2}, \cdots, X_{V_{i+1}})$, $i=1, 2, \cdots$. For $m=1, 2, \cdots$ an m-tuple $(x_1, \cdots, x_m)$ with components in S will be called a *path*. On the set of all paths impose the smallest Borel field containing for $m = 1, 2, \cdots$, all m-dimensional cylinder sets with one dimensional base set in $\mathscr{B}$. Then $\{Y_i\}$ is a Markov process; indeed one obviously has

$$P[Y_{n+k} \in W \mid Y_n, \cdots, Y_1] = P[Y_{n+k} \in W \mid Y_n] = P[Y_{n+k} \in W \mid X_{V_n}]$$

for every Borel set W. $\{Y_i(A)\}$ will be called the *A-path process*. The property of path processes that will be exploited is the following: *if* $A \in \mathscr{D}$, $\{Y_i(A)\}$ *satisfies Doeblin's condition*; the proof is obvious.

If f is a real-valued function defined on S, f^* is defined on paths by the relation $f^*((x_1, \cdots, x_m)) = f(x_1) + \cdots + f(x_m)$, $m = 1, 2, \cdots$. U is to be the function identically equal to 1 on S. For $n = 1, 2, \cdots$, define L_n (or $L_n(A)$) to be the random variable such that L_n is the biggest w such that $V_w \leqq n$.

The reference in the hypotheses of the following two theorems to *some A* may seem unsatisfactory. This point will be discussed at the end of this section and the results of the next section are also relevant.

Before proceeding to the theorems it will be useful to state a lemma. This lemma will serve in the present context in place of Lemma 7.2, p. 224 of [4]; since the lemma follows easily from Theorem 3 and the argument is similar to the corresponding one in [4] no proof needs to be given. If $\{X_i\}$ satisfies (*) the process is *acyclic* (*cyclic*) if $d = 1$ ($d > 1$) in Theorem 3.

LEMMA 5. *Assume $\{X_i\}$ satisfies (*) and has a stationary probability measure Q. Let $w(k)$, $k = 1, 2, \cdots$, be a sequence of positive integers diverging to infinity. Let M be a positive number and for $k = 1, 2, \cdots$, let F_k be a real-valued random variable measurable on $X_{w(k)}, X_{w(k)+1}, \cdots$, such that $|F_k|$ is bounded by M. Let T be the shift operator, i.e., $X_1(\omega) = X_0(T\omega)$, and let $TF_k(\omega) = F_k(T\omega)$, $k = 1, 2, \cdots$.*

If either $\{X_i\}$ is acyclic or for $k = 1, 2, \cdots$, and every $x \in S$

$$\lim_{k \to \infty} E\{F_k - TF_k \,|\, X_0 = x\} = 0 \ .$$

then

$$\lim_{k \to \infty} (E_\varphi\{F_k\} - E_Q\{F_k\}) = 0$$

for every initial probability distribution φ.

Since the process $\{Y_i\}$ is determined by the distribution of X_{V_1} it is natural to indicate the distribution of X_{V_1} by a subscript on the P or E when these operate on sets or random variables measurable on the $\{Y_i\}$ process.

THEOREM 5.1.   *Let $\{X_i\}$ satisfy (*) and have a stationary probability measure Q.   Let $A \in \mathscr{D}, Q(A) > 0, Y_i = Y_i(A), i = 1, 2, \cdots$.   Let $\delta > 0$, f a real-valued measurable function on S.   Let $E_{Q(A)}\{(f^*(Y_1))^{2+\delta}\} < \infty$, $m = E_{Q_A}\{f^*(Y_1)\}$.   Let $\overline{f} = f - mQ(A)U$ and $\sigma^2 = E_{Q_A}\{(\overline{f}(Y_1))^2\}$.   Let*

$$S_n = \sum_{i=1}^{n} \overline{f}(X_i), \ B_n = \sigma \sqrt{n \cdot Q(A)} \ .$$

Then the distribution of S_n/B_n approaches the normal with mean 0, variance 1.

Proof.   Let $V_i = V_i(A), L_n = L_n(A), Z_i = \overline{f}^*(Y_i), 1, 2, \cdots$.   Let $[\alpha]$ denote the largest integer in α for $\alpha > 0$.   The argument follows [2] and [8].   In particular the following decomposition is used:

$$(5.1) \qquad \frac{S_n}{B_n} = \frac{1}{B_n} \sum_{i=1}^{V_1} f(X_i) + \frac{1}{B_n} \sum_{i=V_{L_n}+1}^{n} f(X_i) + \frac{1}{B_n} \sum_{i=1}^{[nQ(A)]} Z_i$$

$$+ \frac{1}{B_n} \sum_{i=[nQ(A)]+1}^{L_n-1} Z_i + \frac{1}{B_n} mQ(A)(n - V_{L_n} + V_1) \ .$$

The distribution of the third term on the right tends to the desired normal by the central limit theorem for Markov chains satisfying Doeblin's condition (see [4], p. 228 or [1]).[3]   So it suffices to show that each of the other terms approaches zero in probability.   The corresponding facts were shown in [2] and [8], but some new arguments are needed in the present case; on the other hand, much of the following argument is due to [2] and [8].

The first term on the right causes no difficulty.   That the second term approaches zero in probability follows from (5.2), which will be proved.

---

[3] These references consider only the acyclic case, but an easy modification works in the general case.

(5.2) $$\lim_{w \to \infty} P[(n - V_{L_n}) > w] = 0 \text{ uniformly in } n.$$

Suppose first that $\{X_i\}$ is acyclic. By Theorem 3 there exists a function $\delta(w)$ tending to zero as w approaches infinity and such that one has for $h = 0, 1, \cdots$:

(5.3) $$\sum_{k=0}^{h} P[n - V_{L_n} = k] = \sum_{k=0}^{h} \int_A \sum_{j=k+1}^{\infty} P[V_2 = j \,|\, X_0 = x] P[X_{n-k} \in dx]$$

$$= \sum_{k=0}^{h} \int_A \sum_{j=k+1}^{\infty} P[V_2 = j \,|\, X_0 = x] Q(dx) + \sum_{k=0}^{h} \delta(n = k)$$

$$= Q(A) \sum_{k=0}^{h} \int_A \sum_{j=k+1}^{\infty} P[V_2 = j \,|\, X_0 = x] Q_A(dx) + \sum_{k=0}^{h} \delta(n - k) .$$

Now

$$\frac{1}{Q(A)} = Q_A(S) = \sum_{k=0}^{\infty} \int_A \sum_{j=k+1}^{\infty} P[V_2 = j \,|\, X_0 = x] Q_A(dx) .$$

Let $\varepsilon > 0$; there then exist h_ε such that the first term of the last member of (5.3) exceeds $1 - \varepsilon/2$ for $h = h_\varepsilon$. Choose n_ε so that

$$\sum_{k=0}^{h_\varepsilon} \delta(n - k) < \frac{\varepsilon}{2}$$

for $n \geqq n_\varepsilon$. Then

$$\sum_{k=0}^{h_\varepsilon} P[n - V_{L_n} = k] > 1 - \varepsilon$$

for $n > n_\varepsilon$. Clearly one can find h_ε' such that

$$\sum_{k=0}^{h_\varepsilon'} P[n - V_{L_n} = k] > 1 - \varepsilon$$

for all n. So for $h \geqq h_\varepsilon'$ $P[n - V_{L_n} > h] < \varepsilon$ for all n, as had to be shown. If $\{X_i\}$ is cyclic a simple variation of the above argument can be used to prove (5.2) provided A is included in one of the cyclic classes. Clearly (5.2) for arbitrary A with positive A-measure follows.

Obviously (5.2) also shows that the last term in (5.1) approaches zero in probability.

Note that

$$\frac{n}{L_n - 1} = \frac{V_{L_n} - V_1}{L_n - 1} + \frac{n - V_{L_n} + V_1}{L_n - 1}$$

$$= \frac{\sum\limits_{i=1}^{L_n-1} U^*(Y_i)}{L_n - 1} + \frac{n - V_{L_n} + V_1}{L_n - 1} + \frac{V_1}{L_n - 1} \ .$$

In the last member the middle term approaches zero in probability by (5.2), and the last term obviously tends to zero in probability. So the law of large numbers, valid for processes satisfying Doeblin's condition (see [1] or [4]),[3] applied to the first term leads to

(5.4) $\dfrac{n}{L_n - 1}$ approaches $\dfrac{1}{Q(A)}$ in probability as $n \to \infty$.

Write $[\alpha]$ for the largest integer in α for $\alpha > 0$ and define

$$\nu(n) = [Q(A)n(1 + \varepsilon)] \text{ and } \lambda(n) = [Q(A)n(1 - \varepsilon)] \ .$$

Let $\varepsilon > 0$. (5.4) shows that there must be an n_0 such that $P[\nu(n) \leq L_n - 1 \leq \lambda(n)] > 1 - \varepsilon$ for $n > n_0$. Thus to show the fourth term on the right in (5.1) approaches zero it needs only to be shown that

(5.5) $\dfrac{1}{B_n} \max\limits_{\nu(n) \leq s \leq \lambda(n)} \left| \sum\limits_{i=\nu}^{s} Z_i \right|$ approaches zero in probability.

To prove (5.5) assume temporarily that $\{Y_i\}$ is stationary, which will make $\{Z_i\}$ stationary. Then (5.5) is equivalent to

(5.6) $\dfrac{1}{B_n} \max\limits_{0 < s < \lambda(n) - \nu(n)} \left| \sum\limits_{i=1}^{s} Z_i \right|$ approaches zero in probability.

The expression in (5.6) equals

$$\frac{\sqrt{\lambda(n) - \nu(n)}}{B_n} \left\{ \max\limits_{0 < s < \lambda(n) - \nu(n)} \frac{1}{\sqrt{\lambda(n) - \nu(n)}} \left| \sum\limits_{i=1}^{s} Z_i \right| \right\} \ .$$

The distribution in the expression in braces approaches a limiting distribution by the Erdös-Kac-Donsker invariance principle, which is applicable here by [1],[3] and the corresponding fact for independent identically distributed random variables with normal distributions of mean 0. Since the quantity preceding the braces approaches zero (5.6), and hence (5.5), holds in this case. That (5.5) holds for any initial distribution follows from Lemma 5. So the theorem is proved.

In [1] Billingsley showed that the invariance principle of Erdös, Kac, and Donsker is applicable to certain sequences of dependent random variables. The following theorem extends these results to processes satisfying (*). The terminology is that of [1].

---

[3] These references consider only the acyclic case, but an easy modification works in the general case.

THEOREM 5.2.[4] *Under the conditions of Theorem 5.1 the invariance principle holds for the sequence $\{S_n\}$ with norming factors $\{B_n\}$.*

Proof. As in the proof of Theorem 4.1 of [1] it suffices to verify two conditions, (i) and (ii).

The verification of (i) in the present case is reduced to verifying the corresponding condition in the case where Doeblin's condition is satisfied in the same manner that the central limit theorem, Theorem 5.1, was reduced to the central limit theorem for processes satisfying Doeblin's condition. When Doeblin's condition is satisfied the argument of [1] applies.[3]

Verification of (ii) is carried out as in [1], except that Lemma 5 is used in place of Lemma 7.2, p. 224 of [4].[3] The fact that

$$\lim_{n \to \infty} E\left\{ \frac{S_n^2}{B_n^2} \right\} \text{ exists and is finite}$$

is also needed; this is easily reduced to Lemma 7.3, p. 224 of [4] by using the decomposition (5.1) and the fact, proved above, that in the right member of (5.1) all terms other than the third one approach zero in probability.

As remarked above the hypotheses of Theorems 5.1, 5.2, have the unsatisfactory feature that they refer to some $A \in \mathscr{B}$. In [2] there are analogous hypotheses referring to some state of the denumerable state space; there, however, the hypotheses are proved invariant in the sense that if they hold for some state they hold for each state. In the present situation there exists no similar invariance. Indeed, it is very simple to give examples of a Markov process satisfying (*) and Doeblin's condition and of a function f such that the conditions of Theorem 5.1 are true for some $A \in \mathscr{B}$ but not for $A = S$.[5] Such examples show also that even when dealing with processes satisfying Doeblin's condition the theorems above may be applicable when the result of [1] is not.

Though the existence of moments of random variables of the form $f^*(Y_1(A))$ or $g(Y_1(A))$ does depend on the choice of A certain facts can be established. This is the subject of the following section.

6. **Relations between path processes.** In this section $\{X_i\}$ will satisfy (*), Q will be the stationary measure, $A, D \in \mathscr{B}, D \subseteq A, 0 < Q(D) < \infty$, and g will be a positive, real-valued, measurable function on state space.[6]

---

[3] These references consider only the acyclic case, but an easy modification works in the general case.

[4] It is clear that in the special case where S is denumerable and A has only one point as member the conditions $\delta > 0$ may be dropped, i.e., δ may be zero.

[5] Example 3 of [2] illustrates this.

[6] The condition that g be positive can be relaxed. See however footnote 7.

Let

$$G_1 = g(Y_1(A))$$

and for $i = 2, 3, \cdots,$

$$G_i = g(Y_i(A)) \text{ if } X_{V_j} \in A - D \text{ for } j = 2, 3, \cdots, i = 0 \text{ otherwise.}$$

If $\bar{x} = (x_1, \cdots, x_v), \bar{y} = (y_1, \cdots, y_w)$ are two paths $\bar{x} + '\bar{y}$ is to stand for the path $(x_1, \cdots, x_v, y_1, \cdots, y_w)$.

The conventions concerning measures appearing as subscripts made § 5 will be used here. For example, $E_{Q_D}\{G_i\}$ is the expected value of G_i when X_{V_1} has the measure Q_D associated with it. In case a probability measure concentrates all its weight on some point it will be convenient to use this point as a subscript; e.g., $E_x(G_i)$ is meaningful when $x \in A$.

LEMMA 6.[7]

$$E_{Q_D}\left\{\sum_{i=1}^{\infty} G_i\right\} = \int_A E_y\{G_1\}\overline{Q}_D(dy) \ (= E_{Q_A}\{G_1\}\overline{Q}_D(A) \text{ if } Q(A) < \infty) \ .$$

If $A = S$ and $g = f^*$ (1.4) results.

Proof.

$$E_{Q_D}\left\{\sum_{i=1}^{\infty} G_i\right\} = \sum_{i=1}^{\infty} E_{Q_D}\{E\{G_i \mid X_0\}\} = \sum_{i=1}^{\infty} \int_D E_x\{G_i\} Q_D(dx)$$

$$= \int_D E_x\{G_i\} Q_D(dx) + \int_D \left(\int_{A-D} E_y\{G_1\} P_A(x, dy)\right.$$

$$+ \int_{A-D} \sum_{i=2}^{\infty} \int_{A-D} P_A(x, dy_1) \int_{A-D} P_A(y_1, dy_2) \cdots$$

$$\cdot \int_{A-D} P(y_{i-2}, dy_{i-1}) E_y\{G_1\} P_A(y_{i-1}, dy)\right) = \int_A E_y\{G_1\}\overline{Q}_D(dy) \ ,$$

the last equality following from (1.1), (1.3) with A for S, therefore P_A for P, and D for A. By [7] $\overline{Q}_A$ and $\overline{Q}_D$ differ only by a constant factor, if $Q(A) < \infty$. Then $\overline{Q}_A(A) = 1$ and $\overline{Q}_A = \overline{Q}_D \cdot Q_D(A)$. The equation in parenthesis follows.

If $g = f^*$ and $A = S$ one has

$$\int_S E_y\{G_1\}\overline{Q}_D(dy) = \int_S \left\{\int_S f(x)P(y, dx)\right\}\overline{Q}_D(dy)$$

---

[7] This lemma may be considered a generalization of (A) of the appendix to [2]. As in [2] the condition that g be positive can be weakened. Chung showed in Example 3 of [2] that even the special case (A) is false if no condition on g is assumed.

$$= \int_{x \in S} \left\{ f(x) \int_{y \in S} P(y, dx) \overline{Q}_D(dy) \right\} = \int_S f(x) \overline{Q}_D(dx) \ .$$

THEOREM 6.1. (a) *If* $g(\bar{x}) + g(\bar{y}) \leqq g(\bar{x} + '\bar{y})$ *for all paths* $\bar{x}$ *and* $\bar{y}$, *then*

$$\sum_{i=1}^{\infty} G_i < g(Y_1(D)) \quad and \quad E_{Q_D}\{f(Y_1(D))\} \geqq \int_A E_y\{f(Y_1(A))\} \overline{Q}_D(dy)$$

$$(= E_{Q_A}\{f(Y_1(A))\} \overline{Q}_D(A) \ if \ Q(A) < \infty) \ .$$

(b) *If* $g(\bar{x}) + g(\bar{y}) \geqq g(\bar{x} + '\bar{y})$ *for all paths* $\bar{x}$ *and* $\bar{y}$, *then*

$$\sum_{i=1}^{\infty} G_i \geqq f(Y_1(D)) \quad and \quad E_{Q_D}\{f(Y_1(D))\} \leqq \int_A E_y\{f(Y_1(A))\} \overline{Q}_D(dy)$$

$$(= E_{Q_D}\{f(Y_1(A))\} \overline{Q}_D(A) \ if \ Q(A) < \infty) \ .$$

Proof. The theorem follows immediately from Lemma 6.1.

Note that for $p \geqq 1$ $(0 < p \leqq 1)$ the function $f(\bar{x}) = (U^*(\bar{x}))^p$ satisfies condition (a) (condition (b)) of the theorem. Call $U^*(Y_1(B))$ the *recurrence time* to B when $B \in \mathscr{B}$; and if furthermore $Q(B) < \infty$, $p \geqq 0$ call $E_{Q_D}\{(U^*(Y_1(B)))^p\}$ the pth *stationary moment of the recurrence time*. It follows that when $Q(S) = 1$ $(Q(S) = \infty)$, $p > 0$, and $B \in \mathscr{B}$ $(B \in \mathscr{B}$ and $(B)) < \infty)$, the pth stationary moment of the recurrence time to B is finite only if the same is true for every Borel superset (subset of positive Q-measure).

The hypotheses of parts (a) and (b) of Theorem 6.1 cannot both be satisfied by the same g unless g is a constant multiple of U^*. The theorem below, on the other hand, is such that for a wide class of functions both the hypotheses of (a) and (b) may be satisfied.

With reference to the hypotheses in Theorem 6.2 observe that if $B \in \mathscr{B}$ the three statements

(a) $E_x\{g(Y_1(B))\}$ is uniformly bounded for all $x \in B$,

(b) $E_\varphi\{g(Y_1(B))\}$ is uniformly bounded for all probability measures φ on B,

(c) $E_\varphi\{g(Y_1(B))\} < \infty$ for all probability measures φ on B,

are all equivalent.

THEOREM 6.2. *Let* $A \in \mathscr{D}$.

(a) *Suppose* (i) $c > 0$ *and* $f(\bar{x} + '\bar{y}) \leqq c(f(\bar{x}) + f(\bar{y}))$ *for all paths* $\bar{x}$ *and* $\bar{y}$, *and* (ii) $M > 0$ *and* $E_x\{f(Y_1(A))\} < M$ *for all* $x \in A$. *Then* $E_x\{f(Y_1(D))\}$ *is uniformly bounded for all* $x \in D$.

(b) *Suppose* (i) $c > 0$ *and* $f(\bar{x}) + f(\bar{y}) \leqq cf(\bar{x} + '\bar{y})$ *for all paths* $\bar{x}$ *and* $\bar{y}$, *and* (ii) $M > 0$ *and* $E_x\{f(Y_1(D))\} \leqq M$ *for all* $x \in D$. *Then* $E_{Q_A}\{f(Y_1(A))\} < \infty$.

Proof. Assume (i) and (ii) of (a) and let $x \in D$. Let ν be the first n such that $n \geq 2$ and $X_{V_n(A)} \in D$, so that $G_i = 0$ for $i > \nu$. Using (i) of (a) repeatedly, and (ii) of (a), one has,

$$E_x\{f(Y_1(D))\} = \sum_{k=1}^{\infty} E_x\{f(Y_1(A) + 'Y_2(A) + \cdots + 'Y_k(A)|\nu = k\}\, P[\nu = k]$$

$$\leq \sum_{k=1}^{\infty} c^{[\log_2 k]+1} E_x\left\{\sum_{i=0}^{k} G_i | \nu = k\right\} P[\nu = k] \leq \sum_{k=1}^{\infty} c^{[\log_2 k]+1} kM P[\nu = k] .$$

Since $A \in \mathscr{D}$, $P[\nu = k]$ decreases exponentially and (a) follows.

Assume (i) and (ii) of (b). Let ν have the same significance as above. One has then, by Lemma 6 and repeated applications of (i) of (b),

$$\sum_{k=1}^{\infty} E_{Q_D}\{f(Y_1(A) +' \cdots + 'Y_k(A))|\nu = k\} P[\nu = k] c^{[\log_2 k]+1}$$

$$\geq E_{Q_D}\left\{\sum_{i=1}^{k} G_i | \nu = k\right\} P[\nu = k] = E_{Q_D}\left\{\sum_{i=1}^{\infty} G_i\right\} = E_{Q_D}\{f(Y_1(A))\} \ddot{Q}_D(A) .$$

Note that when $\nu = k$ $Y_1(D) = Y_1(A) + ' \cdots + 'Y_k(A)$; so assumption (ii) of (b) ensures that in the inequality above each of the expectations in the first member is at most M. Since $A \in \mathscr{D}$ $P[\nu = k]$ decreases exponentially. This proves (b).

As an application consider the following situation: there exist Borel sets D and C each of finite positive Q-measure and each containing only one point. Let A be the union of D and C. If g satisfies (i) of (a) in Theorem 6.2 and (i) of (b) in the same theorem, one has $E_{Q_D}\{g(Y_1(D))\} < \infty$ implies $E_{Q_A}\{g(Y_1(A))\} < \infty$, $E_{Q_A}\{g(Y_1(A))\} < \infty$ implies $E_{Q_C}\{g(Y_1(C))\} < \infty$, since (ii) of (a) and (b) are now automatically true. In particular, g will always satisfy (i) of (a) and (i) of (b) if $g(\bar{x}) = ((U^*(\bar{x}))^p$, $p > 0$. This gives again the result of Chung [2] that for two points, each of finite positive Q-measure, the pth moment of the recurrence time exists for both or neither.

References

1. P. Billingsley, *The invariance principle for dependent random variables*, Trans. Amer. Math. Soc., **83** (1956), 250-268.
2. K. L. Chung, *Contributions to the theory of Markov chains* Trans. Amer. Math. Soc., **76** (1954), 397-419.
3. J. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc., **63** (1948), 393-438.
4. ———, *Stochastic processes*, New York, 1953.
5. P. Erdös, W. Feller, and H. Pollard, *A theorem on power series*, Bull. Amer. Math. Soc., **55** (1949), 201-204.
6. T. E. Harris *Recurrent Markov processes*, II (abstract). Ann. Math. Stat., **26** (1955), 152-153.
7. ———, *The existence of stationary measures for certain Markov processes*, Third

Berkeley Symposium on Mathematical Statistics and Probability, vol. II, Berkeley, 1956, 113-124.

8. D. G. Kendall, *A note on Doeblin's central limit theorem*, Proc. Amer. Math. Soc., **8** (1957), 1037-1039.

UNIVERSITY OF MINNESOTA AND
UNIVERSITY OF CALIFORNIA