JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 7, 127-145 (1963)

Criteria for Stochastic Processes II: Passage-Time Moments*

JOHN LAMPERTI

The Rockefeller Institute, New York, New York

Submitted by Samuel Karlin

I. INTRODUCTION

In a previous paper [1], an attempt was made to find effective criteria to distinguish between recurrent and transient behavior for stochastic processes of a certain type. Here we shall give some similar criteria for certain other aspects of the behavior of these processes, most notably for the existence or nonexistence of passage-time moments.

Let us recall the general situation of [1]. Suppose $\{X_n\}$ is a Markov process with stationary transitions on the nonnegative reals; assume that

$$\limsup X_n = +\infty \quad \text{a.s.}$$

Define

$$\mu_k(x) = E\{(X_{n+1} - X_n)^k \mid X_n = x\}, \qquad k = 1, 2, \tag{1.1}$$

and assume that $\mu_2(x)$ is of the order of a constant for large x. Under an additional condition (existence and boundedness of the conditional $2 + \delta$ moment for some $\delta > 0$), it was shown that

$$\mu_1(x) \leq \frac{\mu_2(x)}{2x} + O(x^{-1-\epsilon}), \qquad \epsilon > 0, \tag{1.2}$$

as $x \to \infty$ implies that $\{X_n\}$ is *recurrent* in the sense that there is a finite interval which is (a.s.) visited infinitely often for any choice of X_0. If, on the other hand,

$$\mu_1(x) \geq (1 + \eta) \frac{\mu_2(x)}{2x}, \qquad \eta > 0, \tag{1.3}$$

for all large x, then $X_n \to \infty$ a.s. ($\{X_n\}$ is *transient*). A useful example of this theorem is provided by *random walks* on the nonnegative integers with transition probabilities of the form

$$p_{n,n+1} = 1 - p_{n,n-1} = \tfrac{1}{2}\left[1 + \frac{\beta}{n} + O(n^{-2})\right], \qquad n > 0, \tag{1.4}$$

* This work was supported by the U.S. National Science Foundation via grants to Dartmouth College and Stanford University.

and $p_{01} = 1 - p_{00} > 0$. It was shown (*inter alia*) by T. E. Harris [2] that such a walk is recurrent if and only if $\beta \leq \frac{1}{2}$; here (for $n > 0$) $\mu_2(n) = 1$ and $\mu_1(n) = \beta n^{-1} + O(n^{-2})$.

Walks satisfying (1.4) can also suggest some criteria for the existence or nonexistence of passage-time moments. Thus for $n > 0$, $E(T_{n0})$ (mean first-passage time from n to 0) is finite if $\beta < -\frac{1}{2}$, infinite if $\beta \geq -\frac{1}{2}$ [2]. We shall show below that

$$2x\,\mu_1(x) + \mu_2(x) \leq -\epsilon < 0 \tag{1.5}$$

for large x implies the existence of a moment analogous to $E(T_{n0})$, while (under some additional assumptions), if the left side of (1.5) is nonnegative for large x the moment is infinite. These results are proved in Sections II and III respectively, along with similar criteria for higher moments.

In Section IV we obtain some corollaries of these results, under the additional hypothesis that the transition function

$$F(x, y) = \Pr\,(X_{n+1} \leq y \mid X_n = x) \tag{1.6}$$

is continuous in x in the weak* topology for measures on $[0, \infty)$. Attention is devoted to the existence of finite stationary measures, the Cesaro convergence of the n-step transition probabilities, and especially to whether these converge to 0. The reason for our extra concern with the latter is that a condition of this kind is essential for the theory developed in [3], but a sufficiently general criterion was not given there. We shall say (as in [3]) that $\{X_n\}$ is *null* if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Pr\,(X_i \in I \mid X_0 = x) = 0 \tag{1.7}$$

for every x and every compact interval I; $\{X_n\}$ is *uniformly null* if (1.7) holds uniformly in x. Here, using a result from Section III, we prove quite easily a sufficient condition for $\{X_n\}$ to be uniformly null which is much more general than that found in [3], and consequently the applicability of the results of [3] is considerably extended.

Finally in the concluding Section V we extend the basic criteria of Sections II and III to certain processes with states in Euclidean spaces of several dimensions. The idea is to apply the one-dimentional results to the radial component, as was done with recurrence criteria in [1]. The radial component, of course, is usually not a Markov process, but it may have an "approximately Markov" character which is sufficient for the proofs of the theorems. Application is made to certain multidimensional random walks with transition probabilities somewhat analogous to (1.4). This section is quite similar to Section IV of [1] and so some of the details are omitted.

The chief tool used in our proofs is the concept of a semimartingale [4, Chap. 7], although only elementary parts of the corresponding theory are needed. The main tricks are to find modifications of the process to be studied which have suitable martingale properties, and to perform some calculations with the moments $\mu_k(x)$ resembling those in [3]. The first of these devices has frequently been used before, most notably by J. L. Doob, but the results that can be obtained with it have evidently not been exhausted.

II. Criteria for Finite Moments

Let $\{X_n\}$ be a discrete-time stochastic process on a Borel subset S of the nonnegative reals; for the time being we assume it is Markov with stationary transitions, though later we shall comment on how this can be relaxed. Let us write for $x \in S$

$$\mu_k(x) = E[(X_{n+1} - X_n)^k \mid X_n = x] = \int_0^\infty (y - x)^k F(x, dy) \tag{2.1}$$

when the corresponding absolute moments are finite.

THEOREM 2.1. *Suppose there exists $\epsilon > 0$ and $A < \infty$ such that, for $x \geq A$, $\mu_2(x)$ exists and*

$$2x\mu_1(x) + \mu_2(x) \leq -\epsilon.\text{[1]} \tag{2.2}$$

Let $T \geq 0$ be the time at which the process first enters the interval $[0, A]$. *Then*

$$E(T) \leq \frac{E(X_0^2)}{\epsilon}. \tag{2.3}$$

PROOF. Let us define a process $\{Z_n\}$ by

$$Z_n = \begin{cases} X_n^2 + \epsilon n & \text{for} \quad n < T, \\ X_T^2 + \epsilon T & \text{for} \quad n \geq T. \end{cases} \tag{2.4}$$

Let $\mathscr{F}_n$ be the Borel field of sets determined by $X_0, X_1, \cdots, X_n$. Clearly Z_n is $\mathscr{F}_n$ measurable Then for $n < T$ and hence for $X_n > A$,

$$\begin{aligned} E(Z_{n+1} - Z_n \mid \mathscr{F}_n) &= E([X_n + (X_{n+1} - X_n)]^2 - X_n^2 + \epsilon \mid X_n) \\ &= 2X_n\mu_1(X_n) + \mu_2(X_n) + \epsilon \leq 0. \end{aligned} \tag{2.5}$$

[1] Of course (2.2) is to hold for $x \in S$. If there are no points of S greater than A, both hypothesis and conclusion are vacuously true. This remark applies as well to the theorems which follow this one.

For $n \geq T$, obviously $E(Z_{n+1} - Z_n \mid \mathscr{F}_n) = 0$. Therefore $\{Z_n\}$ is a supermartingale relative to $\{\mathscr{F}_n\}$ and hence $E(Z_n)$ decreases. By Fatou's lemma, consequently,

$$E(\lim_{n\to\infty} Z_n) \leq \lim_{n\to\infty} E(Z_n) \leq E(Z_0) = E(X_0^2).$$

But from (2.2), $\mu_1(x) < 0$ for $x \geq A$. It follows from our previous work [1, p. 317] that $\{X_n\}$ a.s. enters $[0, A]$, so that $\lim Z_n = X_T^2 + \epsilon T$. As a result we have

$$E(X_T^2) + \epsilon E(T) \leq E(X_0^2),$$

which obviously implies (2.3).

In much the same way a result on the pth moment of T can be obtained:

THEOREM 2.2. *Suppose that for all sufficiently large x*

$$2x\,\mu_1(x) + (2p-1)\,\mu_2(x) \leq -\epsilon \tag{2.6}$$

for some $\epsilon > 0$. Suppose also that

$$\mu_2(x) = O(1), \qquad \mu_k(x) = o(x^{k-2}) \qquad \text{for} \qquad 2 < k \leq 2p.^2 \tag{2.7}$$

Then for any sufficiently large A we have

$$E(T^p) = O(E(X_0^{2p})), \tag{2.8}$$

where T again is the smallest time n at which $X_n \leq A$.

PROOF. Define $\{Z_n\}$ by letting

$$Z_n = \begin{cases} (X_n^2 + cn)^p & \text{for} \quad n < T, \\ (X_T^2 + cT)^p & \text{for} \quad n \geq T, \end{cases} \tag{2.9}$$

where c is a constant such that $0 < c < \epsilon$. We again wish to show that $\{Z_n\}$ is a supermartingale relative to $\{\mathscr{F}_n\}$; again clearly $E(Z_{n+1} - Z_n \mid \mathscr{F}_n) = 0$ when $n \geq T$. Suppose $n < T$ so that $X_n > A$ (where A is still to be chosen large enough). Then from (2.9),

$$E(Z_{n+1} - Z_n \mid \mathscr{F}_n) = E\{[X_{n+1}^2 + c(n+1)]^p - [X_n^2 + cn]^p \mid X_n\}$$

$$= \sum_{l=0}^{p} \binom{p}{l} E[X_{n+1}^{2l} \mid X_n]\, c^{p-l}(n+1)^{p-l} - (X_n^2 + cn)^p. \tag{2.10}$$

[2] Actually $\mu_{2p}(x) = o(x^{2p-2})$ and $\mu_2(x) = O(1)$ imply $\mu_k(x) = o(x^{2k-2})$ for $2 < k < 2p$ by the Schwartz inequality, so this assumption is slightly redundant.

However, using (2.7) we find that

$$E(X_{n+1}^{2l} \mid X_n) = E[(X_n + \Delta X_n)^{2l} \mid X_n] = \sum_{j=0}^{2l} \binom{2l}{j} X_n^{2l-j} \mu_j(X_n)$$

$$= X_n^{2l} + \left[2l\, X_n\, \mu_1(X_n) + \binom{2l}{2} \mu_2(X_n)\right] X_n^{2l-2} + o(X_n^{2l-2}).$$

Substituting this in (2.10), expanding $(X_n^2 + cn)^p$ in binomial series, and rearranging slightly we can write

$$E(Z_{n+1} - Z_n \mid \mathscr{F}_n) = \sum_{l=0}^{p} \binom{p}{l} c^{p-l} X_n^{2l}[(n+1)^{p-l} - n^{p-l}]$$

$$+ \sum_{l=1}^{p} \binom{p}{l} c^{p-l}(n+1)^{p-l} X_n^{2l-2}\, l[2X_n\, \mu_1(X_n) + (2l-1)\, \mu_2(X_n) + o(1)]$$

Combining like powers of X_n this expression becomes

$$\sum_{l=0}^{p-1} X_n^{2l} \frac{p!}{l!(p-l-1)!} c^{p-l-1} \left\{ \frac{c}{p-l} \Delta(n^{p-l}) \right.$$
$$\left. + (n+1)^{p-l-1} [2X_n\, \mu_1(X_n) + (2l+1)\, \mu_2(X_n) + o(1)] \right\}. \qquad (2.11)$$

We shall see that for large X_n, (2.11) is negative for all n. Indeed, since the coefficients are positive it is only necessary to examine the terms in the braces. The term $\Delta(n^{p-l})$ is overestimated as $(p-l)(n+1)^{p-l-1}$; using this, the choice of $c < \epsilon$ and (2.6) we have

$$E(Z_{n+1} - Z_n \mid \mathscr{F}_n) \leq \sum_{l=0}^{p-1} X_n^{2l} \frac{p!}{l!(p-l-1)!} [(n+1)\, c]^{p-l-1} \{c - \epsilon + o(1)\},$$

which is negative for all large enough values of X_n. Thus if A is chosen suitably we will have $\{Z_n\}$ a supermartingale.

Finally, just as in the proof of the previous theorem, Fatou's lemma yields

$$E[(X_T^2 + cT)^p] = E[\lim_{n\to\infty} Z_n] \leq E(Z_0) = E(X_0^{2p}).$$

When the left side is expanded by the binomial theorem all terms are positive; therefore we have

$$c^p E(T^p) \leq E(X_0^{2p}),$$

which proves (2.8).

Remark on the proof. As suggested in the introduction, we shall sometimes wish to apply Theorems 2.1 and 2.2 when $\{X_n\}$ is not a Markov process. Suppose that $\{Y_n\}$ is a Markov process, not necessarily on the real line, and that $X_n = f(Y_n)$ for some real-valued function f which need not be $1 - 1$. We replace (2.1) by

$$\mu_k'(y) = E[(X_{n+1} - X_n)^k \mid Y_n = y]. \tag{2.1'}$$

Suppose, to take first Theorem 2.1, that

$$2f(y)\,\mu_1'(y) + \mu_2'(y) \leq -\epsilon \tag{2.2'}$$

for all y such that $f(y) \geq A$. If the fields $\mathscr{F}_n$ are now defined in terms of Y_0, $\cdots$, Y_n, the proof of the theorem is valid without further change, and (2.3) still holds. In the same way, provided the analogues of (2.6) and (2.7) are known to hold when $x = f(y)$, the proof and conclusion of Theorem 2.2 remain valid. This remark will be applied in Section V.

III. Criteria for Infinite Moments

In this section we consider the slightly more difficult converse questions. Our approach depends on two facts which hold under the assumptions below. These are that the time of first passage from a large x to a fixed interval $[0, A]$ is at least of the order of x^2 with high probability, and that an appropriate function of $\{X_n\}$ is a submartingale, making it possible to estimate from below the probability of reaching large states before entering $[0, A]$. We begin with the first moment:

THEOREM 3.1. *Suppose that the conditional moments defined in (2.1) satisfy*

$$2x\,\mu_1(x) + \mu_2(x) \geq \epsilon > 0 \tag{3.1}$$

for all $x \geq A$, and in addition that

$$\mu_1(x) = O(x^{-1}), \qquad \mu_2(x) = O(1), \qquad \mu_4(x) = o(x^2). \tag{3.2}$$

Then the time T_{x_0} of first passage from $x_0 > A$ to $[0, A]$ has infinite expectation.

For the proof of both the theorems of this section we need the following

LEMMA 3.1. *If (3.2) holds there is a constant $\epsilon > 0$ such that*

$$\Pr(T_x > \epsilon x^2) \geq \tfrac{1}{2} \tag{3.3}$$

for all $x \geq A + \delta$.

Proof. Let us consider a process $\{\tilde{X}_n\}$ with $\tilde{X}_0 = x > A$, and which has the same transition law as $\{X_n\}$ except that it is *stopped* upon first entering the interval $[0, A]$, maintaining the same value from that time on. We see by writing $X_{n+1} = X_n + \Delta X_n$ that

$$E[\tilde{X}_{n+1}^2 - \tilde{X}_n^2 \mid \tilde{X}_n] = 2\tilde{X}_n \, \mu_1(\tilde{X}_n) + \mu_2(\tilde{X}_n) \tag{3.4}$$

for $\tilde{X}_n > A$; for $\tilde{X}_n \le A$, $\tilde{X}_{n+1} = \tilde{X}_n$ so the result is 0. Thus for all values of $\tilde{X}_n$ we have

$$-C_1 \le E[\tilde{X}_{n+1}^2 - \tilde{X}_n^2] = E\{2\tilde{X}_n \, \mu_1(\tilde{X}_n) + \mu_2(\tilde{X}_n)\} \le C_1$$

because of (3.2). It follows that

$$x^2 - nC_1 \le E(\tilde{X}_n^2) \le x^2 + nC_1. \tag{3.5}$$

In the same way we can estimate $E(\tilde{X}_n^4)$; thus for $\tilde{X}_n > A$

$$E[\tilde{X}_{n+1}^4 - \tilde{X}_n^4 \mid \tilde{X}_n] = 4\tilde{X}_n^3 \, \mu_1(\tilde{X}_n) + 6\tilde{X}_n^2 \, \mu_2(\tilde{X}_n) + 4\tilde{X}_n \, \mu_3(\tilde{X}_n) + \mu_4(\tilde{X}_n).$$

The left side is then seen to be $O(\tilde{X}_n^2)$ because of (3.2) when $\tilde{X}_n > A$; it is 0 for $\tilde{X}_n \le A$. (We actually need here only $\mu_4(x) = O(x^2)$.) Using (3.5) we have

$$E[\tilde{X}_{n+1}^4 - \tilde{X}_n^4] = E\{E[\tilde{X}_{n+1}^4 - \tilde{X}_n^4 \mid \tilde{X}_n]\} \le C_2(x^2 + nC_1),$$

which implies that

$$E(\tilde{X}_n^4) \le C_2 x^2 n + C_3 n^2 + x^4. \tag{3.6}$$

From (3.5) and (3.6) we obtain

$$E[(\tilde{X}_n^2 - x^2)^2] \le C_4 x^2 n + C_3 n^2. \tag{3.7}$$

To complete the proof we apply Chebychef's inequality to (3.7) and obtain

$$\Pr(|\tilde{X}_n^2 - x^2| \ge x^2 - A^2) \le \frac{C_4 x^2 n + C_3 n^2}{(x^2 - A^2)^2}. \tag{3.8}$$

But $|\tilde{X}_n^2 - x^2| < x^2 - A^2$ implies $\tilde{X}_n > A$ so that the process $\{\tilde{X}_n\}$ has not yet been stopped; therefore $T_x > n$. Putting $n = \epsilon x^2$ (more precisely, the integer part of ϵx^2) we have

$$\Pr(T_x > \epsilon x^2) \ge \Pr(|\tilde{X}_n^2 - x^2| < x^2 - A^2) \ge 1 - \frac{C_4 \epsilon + C_3 \epsilon^2}{[1 - (A^2/x^2)]^2}.$$

As long as $x \ge A + \delta$, $\delta > 0$, the denominator on the right-hand side is bounded from 0; choosing a small enough ϵ we have (3.3).

PROOF OF THEOREM 3.1. Let us choose $B > x_0$, and form the process $\{\tilde{X}_n\}$ with initial state $x_0 > A$, transition probabilities $F(x, y)$ for $A < x < B$, and *stopped* when first either $\tilde{X}_n \leq A$ or $\tilde{X}_n \geq B$. The process $\{\tilde{X}_n^2\}$ is a submartingale; the inequality follows from (3.4) and (3.1) for $A < \tilde{X}_n < B$, and is trivial for other $\tilde{X}_n$ since the process has stopped in that case.

Let us suppose for the moment that from no state x in the interval (A, B) is it possible to make in one step a transition to a state beyond $2B$; i.e., that

$$F(x, 2B) = 1 \qquad \text{for} \qquad A < x < B. \tag{3.9}$$

Then $\{\tilde{X}_n^2\}$ is bounded, and a martingale system theorem [4, p. 302] applies and yields $E(\tilde{X}_\tau^2) \geq E(\tilde{X}_0^2) = x_0^2$, where τ is the stopping time. But, again using (3.9),

$$E(\tilde{X}_\tau^2) \leq A^2 + (2B)^2 \Pr(\tilde{X}_\tau \geq B),$$

so that if we denote the probability on the right by $q(x_0, B)$,

$$q(x_0, B) \geq \frac{x_0^2 - A^2}{4B^2}. \tag{3.10}$$

Combining (3.10) with (3.3) (i.e., Lemma 3.1), we obtain

$$\Pr(T_{x_0} > \epsilon B^2) \geq \frac{q(x_0, B)}{2} \geq \frac{x_0^2 - A^2}{8B^2}. \tag{3.11}$$

The constants in (3.11) are independent of B; thus for large u,

$$\Pr(T_{x_0} > u) \geq \frac{c}{u}$$

for a fixed $c > 0$, which proves that $E(T_{x_0}) = \infty$.

The difficulty with this is the unpleasant assumption (3.9) which we shall now remove by a truncation argument. Suppose the original transition probability F satisfies (3.1) and (3.2) but not (3.9); choose B and form, for $A < x < B$,

$$F^*(x, y) = \begin{cases} F(x, y) & \text{for} \quad y < 2B \\ 1 & \text{for} \quad y \geq 2B. \end{cases} \tag{3.12}$$

Thus the transitions to points beyond $2B$ are moved back to $2B$. The process $\{\tilde{X}_n^*\}$, formed from F^* just as $\{\tilde{X}_n\}$ was formed from F above, has exactly the same $q(x_0, B)$ as does $\{\tilde{X}_n\}$. Thus if $\{(\tilde{X}_n^*)^2\}$ is a submartingale, (3.10) will hold and the conclusion of the theorem follows as before. It is therefore necessary to examine the functions $\mu_1^*(x)$ and $\mu_2^*(x)$ defined by using F^* in place of F in (2.1).

This last step is easily accomplished. It is clear from (3.12) and (2.1) that, for $A < x < B$,

$$| \mu_2(x) - \mu_2^*(x) | \leq \int_{2B}^{\infty} (y - x)^2 F(x, dy).$$

But the right-hand side is less than

$$\int_{2B}^{\infty} \frac{(y-x)^4}{B^2} F(x, dy) < \frac{1}{B^2} \int_0^{\infty} (y - x)^4 F(x, dy) = \frac{\mu_4(x)}{B^2}.$$

Since $x < B$ and $\mu_4(x) = o(x^2)$, we have

$$\mu_2^*(x) = \mu_2(x) + o(1), \tag{3.13}$$

where $o(1)$ is uniform in $x < B$ as $B \to \infty$. Similarly it is shown that

$$\mu_1^*(x) = \mu_1(x) + o(B^{-1}), \tag{3.14}$$

with the error term interpreted as above. Thus if B is large enough, from (3.1), (3.13), and (3.14) we have

$$2x\, \mu_1^*(x) + \mu_2^*(x) \geq \frac{\epsilon}{2}$$

for $A < x < B$ and so $\{(\tilde{X}_n^*)^2\}$ is a submartingale, by the same argument used above for $\{\tilde{X}_n^2\}$. This fact, with the observation of the previous paragraph, completes the proof.

Remark (a). The reader may have noticed that except for the need to remove condition (3.9), the weaker hypothesis

$$2x\, \mu_1(x) + \mu_2(x) \geq 0 \tag{3.1'}$$

for $x \geq A$ would have been sufficient since it makes $\{\tilde{X}_n^2\}$ a submartingale. Thus if (3.9) is actually satisfied by $F(x, y)$ (for instance, if $| X_{n+1} - X_n |$ is uniformly bounded a.s.) (3.1') may replace (3.1). However, a sharper theorem can be obtained by a modification of the argument: we define $Y_n = \tilde{X}_n^2 \log \tilde{X}_n$ and try to show $\{Y_n\}$ to be a submartingale. If it is (more precisely, if $\{Y_n^*\}$ is a submartingale, where the transition function F^* has replaced F) we have

$$q(x_0, B) \geq \frac{x_0^2 \log x_0 - A^2 \log A}{4B^2 \log 2B},$$

analogous to (3.10). From this as before we obtain

$$\Pr (T_{x_0} > u) \geq \frac{c}{u \log u}$$

for large u, and again $E(T_{x_0}) = \infty$. A fairly straightforward argument, involving estimates slightly more tedious than those carried out hitherto, shows that if

$$2x\,\mu_1(x) + \mu_2(x) \geq O(x^{-\epsilon}), \qquad \epsilon > 0, \tag{3.1*}$$

holds together with

$$\mu_1(x) = O(x^{-1}), \qquad \eta \leq \mu_2(x) \leq M, \qquad \mu_4(x) = O(x^{2-\delta}) \tag{3.2*}$$

for some η and $\delta > 0$, then $\{Y_n^*\}$ is indeed a submartingale when A is large enough. Thus if (3.1*) replaces (3.1) and (3.2*) replaces (3.2), Theorem 3.1 is still valid. Doubtless other improvements are possible; however, the "dividing line" between cases of finite and infinite mean passage time is drawn now fairly sharply by Theorems 2.1 and 3.1.

Remark (b). The result we have just proved can be compared with one by Doob [4, p. 308] which also asserts $E(T) = \infty$. In our notation, Doob's proof works when $\mu_1(x) \geq 0$ for large x and the corresponding absolute moment is bounded. While our theorem allows $\mu_1(x)$ to be "slightly" negative, and so is in a sense sharper, Doob's does not need higher moments. It does not seem likely that our result can be obtained by his method.

Now we turn to higher moments of T, and use much the same approach to prove

THEOREM 3.2. *Suppose for some integer $p > 1$ that*

$$2x\,\mu_1(x) + (2p-1)\,\mu_2(x) \geq \epsilon > 0 \tag{3.15}$$

for all large x; suppose also that $\mu_{2p}(x)$ exists and

$$\mu_1(x) = O(x^{-1}), \qquad \mu_2(x) = O(1), \qquad \mu_{2p}(x) = o(x^{2p-2}). \tag{3.16}$$

Then for all sufficiently large A we have $E(T_{x_0}^p) = \infty$ for every $x_0 > A$.

PROOF. Define $\{\tilde{X}_n\}$ as in the proof of Theorem 3.1 with $A < x_0 < B$, where A must be chosen large enough so that $\{\tilde{X}_n^{2p}\}$ is a submartingale. To see that this is possible, write for $n < \tau$

$$E[\tilde{X}_{n+1}^{2p} - \tilde{X}_n^{2p} \mid \mathscr{F}_n] = E[(\tilde{X}_n + \Delta\tilde{X}_n)^{2p} - \tilde{X}_n^{2p} \mid \tilde{X}_n]$$
$$= \sum_{l=1}^{2p} \binom{2p}{l} \tilde{X}_n^{2p-l}\,\mu_l(\tilde{X}_n). \tag{3.17}$$

Because of (3.16), $\mu_l(x) = o(x^{l-2})$ for $2 < l < 2p$; using this and (3.15) we see that the quantity in (3.17) is positive for all large enough values of $\tilde{X}_n$.

We now choose A so that it is positive for $\tilde{X}_n > A$ and have established our assertion.

Just as in Theorem 3.1, we temporarily assume (3.9). Proceeding as before we then obtain

$$q(x_0, B) \geq \frac{x_0^{2p} - A^{2p}}{(2B)^{2p}}; \tag{3.18}$$

using (3.3) there results

$$\Pr(T_{x_0} > u) \geq \frac{c}{u^p}, \qquad c > 0,$$

for large u, and it follows that $E(T_{x_0}^p) = \infty$. The assumption (3.9) is removed by again introducing F^* and showing that $\{(\tilde{X}_n^*)^{2p}\}$ is still a submartingale. We shall omit the details since the procedure is entirely analogous to our earlier one.

Theorem 3.2 can doubtless be somewhat improved in the manner indicated for the first moment in Remark (a) above, but we shall not attempt this. Finally, we comment that a generalization of the kind indicated at the end of Section II is also valid for the results of this section, and will be applied in dealing with multidimensional processes in Section V. The proof requires no additional arguments; under assumptions bearing the same relation to those of the theorems of this section which (2.2′) bore to (2.2), the proofs we have given apply with only notational changes.

IV. Applications

Throughout this section we shall suppose that the state-space S is a closed subset of $[0, \infty)$, and that the transition probability $F(x, y)$ defined in (1.6) is weak*-continuous as a function of x. Specifically, by this we mean that the transformation

$$Tf(x) = \int_0^\infty f(y)\, F(x, dy) \tag{4.1}$$

maps the class C_∞ of continuous functions on S vanishing at ∞ into itself. We will refer to this as "continuity," or condition (c). Note that if S is the positive integers, or any set without finite limit points, condition (c) reduces to the requirement that a function vanishing at ∞ is carried into another such. This must be true if $\mu_2(x)$ is bounded, or even $o(x^2)$, by Chebychef's inequality.

Lemma 4.1. *Under* (c), *a process* $\{X_n\}$ *which is not uniformly null has at least one finite stationary measure.*

PROOF. Let us write

$$G_n(x, y) = \frac{1}{n} \sum_{i=0}^{n-1} \Pr(X_i \leq y \mid X_0 = x). \tag{4.2}$$

Then $\{X_n\}$ uniformly null means (as in the introduction) that $G_n(x, y) \to 0$ as $n \to \infty$ uniformly in x for each y. If this is not the case, there is a sequence x_n such that $G_n(x_n, y)$ does not go to 0 for some, and hence all large enough, values of y. It is then possible to select a subsequence $\{n'\}$ of the integers such that

$$\lim_{n' \to \infty} G_{n'}(x_{n'}, \cdot) = G(\cdot) \tag{4.3}$$

in the weak* sense;[3] G is nondecreasing and $0 = G(0-) < G(\infty) \leq 1$.

We shall show that G provides an invariant measure for $\{X_n\}$; that is, for all y which are continuity points of G,

$$G(y) = \int_0^\infty F(x, y)\, dG(x). \tag{4.4}$$

On the one hand, we have

$$\int_0^\infty F(x, y)\, G_{n'}(x_{n'}, dx)$$

$$= \frac{1}{n'} \sum_{i=0}^{n'-1} \int_0^\infty F(x, y)\, d \Pr(X_i \leq x \mid X_0 = x_{n'})$$

$$= \frac{1}{n'} \sum_{i=0}^{n'-1} \Pr(X_{i+1} \leq y \mid X_0 = x_{n'})$$

$$= G_{n'}(x_{n'}, y) + \frac{1}{n'} \{\Pr(X_{n'} \leq y \mid X_0 = x_{n'}) - \Pr(X_0 \leq y \mid X_0 = x_{n'})\} \tag{4.5}$$

which obviously tends to $G(y)$ at continuity points, hence in the weak* sense. On the other hand, if $f \in C_\infty$,

$$\int_0^\infty f(y)\, d \int_0^\infty F(x, y)\, G_{n'}(x_{n'}, dx) = \int_0^\infty Tf(x)\, G_{n'}(x_{n'}, dx) \to \int_0^\infty Tf(x)\, dG(x)$$

by (4.3) and (c). Thus the top expression in (4.5) converges in the weak* sense to $\int_0^\infty F(x, y)\, dG(x)$; combining this and the above yields (4.4) and proves the lemma.

[3] This is equivalent to ordinary convergence at all continuity points of G.

In order that $\{X_n\}$ be null, the following condition is clearly necessary:

$$\Pr(\limsup X_n = \infty \mid X_0 = x) = 1 \qquad \text{for all } x \in S. \tag{4.6}$$

This is usually not hard to verify; it is, for instance, true if

$$F(x, x + \epsilon) < 1 - \epsilon, \qquad \epsilon > 0, \qquad x \in S, \tag{4.7}$$

and this can be considerably generalized. We now prove

THEOREM 4.1. *If a Markov process satisfies* (c), *(4.6), and the hypotheses of Theorem 3.1, then it is uniformly null.*

PROOF. If it is not, by the lemma there is a finite invariant measure which can be normalized and used to define a strictly stationary process $\{X_n\}$ with the transition probability $F(x, y)$. The stationary distribution function, say G, is less than 1 for all finite x since the contrary would contradict (4.6). Choose an interval $[0, A]$ with A large enough that $G(A) > 0$, and satisfying the conditions of Theorem 3.1 so that the mean passage time from x into $[0, A]$ is infinite for every $x \in S \cap (A, \infty)$. Let

$$R = \{x \in S : x \le A, F(x, A) < 1\};$$

because of (4.6) it is easy to see that R has positive G-measure. But for all $x \in R$, the mean time to return to $[0, A]$ is infinite. This contradicts a theorem of Kac [5], which asserts (in particular) that the mean passage time from $x \in [0, A]$ into $[0, A]$ is finite a.e. (G) in $[0, A]$ for such a stationary process.

Remarks. It would be natural to investigate whether the Cesaro convergence in Theorem 4.1 (implicit in the definition of "null process") can be replaced by an ordinary limit. It appears likely that this can usually be done, but since it is somewhat far from the theme of this paper and not relevant to the needs of [3] we shall not pursue the question here. Theorem 4.1 as it stands much generalizes the results in the appendix of [3], where the restriction that S be discrete was needed.

We shall now take a brief look at the "positive" (i.e., not null) case. The results are summarized as

THEOREM 4.2. *Assume that condition* (c) *and condition (2.2) hold; assume also that*

$$E\left[(X_{n+1} - X_n)^2 \left(\log \frac{X_{n+1} + 1}{X_n + 1}\right)^+ \,\middle|\, X_n = x\right] = o(\log x). \tag{4.8}$$

Then for any fixed x every subsequence $\{n'\}$ of integers contains a sub-subse-

quence $\{n''\}$ such that $G_{n''}(x, \cdot)$ (defined in (4.2)) converges weak to an invariant probability distribution. If the invariant distribution is unique, say G, then for all x*

$$w^* \lim_{n \to \infty} G_n(x, \cdot) = G(\cdot). \tag{4.9}$$

Remark. Under a certain type of recurrence condition an invariant distribution must be unique, as T. E. Harris has shown in [6]. Our conditions imply (by results in [1]) a type of recurrence, but it is much weaker than Harris' and in fact it is easy to see by examples that (c) and (2.2) do not imply the uniqueness of a stationary distribution.

PROOF. Under condition (c), as we saw in the proof of Lemma 4.1, when a subsequence of $G_n(x, \cdot)$ is w^* convergent the limit function is an invariant "distribution" which may have total mass less than 1. If this possibility can be ruled out, the conclusions of the theorem will follow.

In [3], Theorem 2.3, it was essentially shown[4] that (2.2) implies that $\{X_n\}$ is not null. The proof consists simply of calculating $E(X_n^2)$, and observing that (2.2) plus the assumption that $G_n(x, y) \to 0$ for fixed x and (large enough) y implies $E(X_n^2)$ eventually negative. We will use a similar idea here, but to get our stronger conclusion it is necessary to calculate

$$B_n = E[(X_n + 1)^2 \log (X_n + 1)].$$

The point of departure is the expression

$$\Delta B_n = E[E\{(X_n + \Delta X_n + 1)^2 \log (X_n + \Delta X_n + 1) - (X_n + 1)^2 \log (X_n + 1) \mid X_n\}].$$

Now by Taylor's theorem for any $x > 0$, $x + h > 0$

$$(x + h)^2 \log (x + h) - x^2 \log x = h(2x \log x + x) + \frac{h^2}{2} (2 \log \xi + 3),$$

where ξ is between x and $x + h$. Combining the preceding and using (4.8), we obtain

$$\Delta B_n = E\{[2(X_n + 1)\, \mu_1(X_n + 1) + \mu_2(X_n + 1)] \log (X_n + 1) + o\,(\log X_n)\},$$

so that, using (2.2), for large A

$$\Delta B_n \leq -\frac{\epsilon}{2} \int_A^\infty \log (x + 1)\, d \Pr (X_n \leq x) + O\,[\Pr (X_n \leq A)].$$

[4] The result was stated slightly less generally.

Summing this we find that

$$\frac{B_n}{n} \leq \frac{(x_0+1)^2 \log(x_0+1)}{n} + O[G_n(x_0, A)] - \frac{\epsilon}{2}\int_A^\infty \log(x+1)\, G_n(x_0, dx). \tag{4.10}$$

Now suppose that a subsequence $\{n'\}$ exists such that $G_{n'}(x_0, \cdot) \xrightarrow{w^*} G(\cdot)$ where $G(\infty) < 1$. Since $G_n(x_0, \infty) = 1$ for all n, we have

$$\lim_{n' \to \infty} \{1 - G_{n'}(x_0, y)\} \geq 1 - G(\infty) > 0 \tag{4.11}$$

for all y. Because the integrand in the last term in (4.10) tends to $+\infty$, (4.11) means the integral must do the same as $n \to \infty$, and so $B_{n'} < 0$ for large n'. This is a contradiction, and so the only adverse possibility is ruled out and the theorem follows.

V. Several Dimensions

In this section we will impose hypotheses which are stronger than necessary in order to simplify the discussion. Let $\{\mathbf{X}_n\}$ be s-dimensional random vectors[5] forming a Markov process with stationary transition probability function

$$F(y_1, \cdots, y_s; \mathbf{x}) = \Pr(X_{n+1}^{(i)} - X_n^{(i)} \leq y_i,\, i = 1, \cdots, s \mid \mathbf{X}_n = \mathbf{x}). \tag{5.1}$$

We assume that the increments are bounded; i.e., that

$$\| \mathbf{X}_{n+1} - \mathbf{X}_n \| \leq B \text{ a.s.} \qquad \text{for} \qquad n = 0, 1, \cdots. \qquad (5.2$$

The vector and matrix valued functions

$$\begin{aligned} \mathbf{m}(\mathbf{x}) &= E[\mathbf{X}_{n+1} - \mathbf{X}_n \mid \mathbf{X}_n = \mathbf{x}], \\ \mathbf{v}(\mathbf{x}) &= [v_{ij}(\mathbf{x})] = E[(\mathbf{X}_{n+1} - \mathbf{X}_n)(\mathbf{X}_{n+1} - \mathbf{X}_n)^{\mathrm{T}} \mid \mathbf{X}_n = \mathbf{x}] \end{aligned} \tag{5.3}$$

are then defined and can be calculated from the function F of (5.1). The basic idea is to consider the radial component $R_n = \| \mathbf{X}_n \|$, but $\{R_n\}$ is not, in general, a Markov process. However, as pointed out in the remarks at the end of Section II and III, it is sufficient that the key inequalities such as

[5] Boldface letters denote s-dimensional vectors or s by s matrices in this section. Vectors are columns unless written with superscript T for transpose; "$\| \cdot \|$" is the Euclidean norm.

(2.2′) hold for each of the states of the underlying process $\{\mathbf{X}_n\}$. By proceeding in this way we will prove the following

THEOREM 5.1. *Suppose that*

$$\lim_{\|\mathbf{x}\|\to\infty} \mathbf{v}(\mathbf{x}) = \mathbf{v} \tag{5.4}$$

exists, that $\mathbf{v}$ *is (strictly) positive definite, and that*

$$\lim_{\|\mathbf{x}\|\to\infty} \mathbf{x}^T\mathbf{v}^{-1}\,\mathbf{m}(\mathbf{x}) = \gamma \tag{5.5}$$

exists also. Then if $2\gamma < 2 - 2p - s$, *there exists a sphere* $\|\mathbf{x}\| \leq A$ *such that the first-passage time* $T_\mathbf{x}$ *from* $\mathbf{x}$ *to the sphere satisfies*

$$E(T_\mathbf{x}^p) = O(\|\mathbf{x}\|^{2p}). \tag{5.6}$$

Conversely, if $2\gamma > 2 - 2p - s$ *and* A *is large,* $E(T_\mathbf{x}^p) = \infty$ *for all* $\mathbf{x}$ *with* $\|\mathbf{x}\| > A$.

PROOF. The calculations are essentially the same as those in the proof of Theorem 4.1 of [1]; we outline them for completeness. We begin with the case $\mathbf{v} = \mathbf{I}$, the identity matrix, and calculate

$$\mu_1'(\mathbf{x}) = E[R_{n+1} - R_n \mid \mathbf{X}_n = \mathbf{x}]$$
$$= \int_{-B}^{B} \cdots \int_{-B}^{B} \left\{ \left(\sum_{i=1}^{s} (x_i + y_i)^2 \right)^{1/2} - \left(\sum_{i=1}^{s} x_i^2 \right)^{1/2} \right\} dF(y_1, \cdots, y_s; \mathbf{x})$$

approximately for large $\|\mathbf{x}\|$. The result is

$$\mu_1'(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{m}(\mathbf{x})}{\|\mathbf{x}\|} + \frac{s-1}{2\|\mathbf{x}\|} + o(\|\mathbf{x}\|^{-1}). \tag{5.7}$$

In the same way, it is seen that

$$\mu_2'(\mathbf{x}) = E[(R_{n+1} - R_n)^2 \mid \mathbf{X}_n = \mathbf{x}] = 1 + o(1) \tag{5.8}$$

as $\|\mathbf{x}\| \to \infty$. The higher moments are bounded because of (5.2).

Now if (5.5) holds, the theorems of Sections II and III can be applied to the process $\{R_n\}$ to yield the conclusion of Theorem 5.1. For instance, if $2\gamma < -s$, we have

$$2\|\mathbf{x}\|\,\mu_1'(\mathbf{x}) + \mu_2'(\mathbf{x}) = 2\gamma + (s-1) + 1 + o(1) < -\epsilon$$

for large $\| \mathbf{x} \|$, so that by Theorem 2.1

$$E(T_{\mathbf{x}}) = O(\| \mathbf{x} \|^2).$$

The other cases are just the same, except that a different one of our theorems applies. We are, of course, appealing to the extended version of these theorems explained at the end of Section II.

To remove the restriction $\mathbf{v} = \mathbf{I}$, let $\mathbf{Q}$ be a nonsingular matrix such that $\mathbf{QvQ}^{\mathrm{T}} = \mathbf{I}$, and define $\mathbf{Y}_n = \mathbf{QX}_n$. Then $\{\mathbf{Y}_n\}$ is again a Markov process of the type we are considering and

$$E[(\mathbf{Y}_{n+1} - \mathbf{Y}_n)(\mathbf{Y}_{n+1} - \mathbf{Y}_n)^{\mathrm{T}} \mid \mathbf{Y}_n = \mathbf{Qx}] = \mathbf{Qv(x)}\,\mathbf{Q}^{\mathrm{T}}.$$

Since $\mathbf{v(x)} \to \mathbf{v}$ as $\| \mathbf{x} \|$ (and so $\| \mathbf{y} \|$) tends to ∞, $\{\mathbf{Y}_n\}$ has the limiting covariance matrix $\mathbf{I}$. Similarly

$$\boldsymbol{\mu}^*(\mathbf{Qx}) \equiv E[\mathbf{Y}_{n+1} - \mathbf{Y}_n \mid \mathbf{Y}_n = \mathbf{Qx}] = \mathbf{Q}\boldsymbol{\mu}(\mathbf{x}),$$

so that $\mathbf{y}^{\mathrm{T}}\boldsymbol{\mu}^*(\mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\boldsymbol{\mu}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{v}^{-1}\boldsymbol{\mu}(\mathbf{x})$. Thus the special case of the theorem which we have already proved applies to $\{\mathbf{Y}_n\}$, and the general version as stated is the result.

As in [1] we shall apply the result to certain s-dimensional *random walks*, by which we mean Markov chains on the lattice points of E_s with transition probability matrices of the form

$$P_{\mathbf{x},\mathbf{x}+\mathbf{u}_i} = p_i(\mathbf{x}), \qquad P_{\mathbf{x},\mathbf{x}-\mathbf{u}_i} = q_i(\mathbf{x}), \qquad P_{\mathbf{x},\mathbf{y}} = 0 \text{ otherwise},$$

where $\mathbf{u}_i$ is a unit vector in the direction of the positive ith coordinate axis. It is obvious that (5.2) holds and that

$$\mathbf{v(x)} = \operatorname{diag}[p_i(\mathbf{x}) + q_i(\mathbf{x})].$$

We assume that (5.4) holds with $\mathbf{v}$ definite, which means that

$$\lim_{\|\mathbf{x}\| \to \infty} [p_i(\mathbf{x}) + q_i(\mathbf{x})] = d_i > 0$$

exists. It is then easily verified that

$$\mathbf{x}^{\mathrm{T}}\mathbf{v}^{-1}\,\boldsymbol{\mu}(\mathbf{x}) = \sum_{i=1}^{s} x_i \frac{p_i(\mathbf{x}) - q_i(\mathbf{x})}{d_i}. \tag{5.9}$$

We consider two classes of walks. The first class is defined by supposing $p_i(\mathbf{x})$ and $q_i(\mathbf{x})$ to be of the form

$$p_i(\mathbf{x}) = p_i\left[1 + \frac{\alpha x_i}{\|\mathbf{x}\|^2} + o(\|\mathbf{x}\|^{-1})\right],$$

$$q_i(\mathbf{x}) = p_i\left[1 - \frac{\alpha x_i}{\|\mathbf{x}\|^2} + o(\|\mathbf{x}\|^{-1})\right], \qquad (5.10)$$

where $p_i > 0$. (It is also assumed in both classes of walks that all states communicate and that $\Sigma_{i=1}^{s}(p_i(\mathbf{x}) + q_i(\mathbf{x})) = 1$ for each $\mathbf{x}$.) Theorem 5.1 then has the following:

COROLLARY 5.1. *Let $\{\mathbf{X}_n\}$ be an s-dimensional random walk satisfying (5.10); let T_{00} be the recurrence time from state 0 to itself. Then if $2\alpha < 2 - 2p - s$, $E(T_{00}^p)$ is finite, while if $2\alpha > 2 - 2p - s$, $E(T_{00}^p) = \infty$.*

The second class of random walks to which we shall apply our theorem is defined by letting

$$p_i(\mathbf{x}) = p_i\left[1 + \frac{\alpha_i}{x_i - \xi_i} + o(\|\mathbf{x}\|^{-1})\right], \qquad p_i > 0$$

$$q_i(\mathbf{x}) = p_i\left[1 - \frac{\alpha_i}{x_i - \xi_i} + o(\|\mathbf{x}\|^{-1})\right]. \qquad (5.11)$$

It is necessary for our proof to have the constants ξ_i not integral, though no doubt if $\xi_i = 0$ and $p_i(\mathbf{x})$ is redefined when $x_i = 0$, the results are the same. We do not apply (5.9) as it stands, but translate the origin from $\mathbf{0}$ to $\boldsymbol{\xi}$ in order that the limit (5.5) should exist; this does not change $\mathbf{v}$. Applying the theorem we obtain

COROLLARY 5.2. *Let $\{\mathbf{X}_n\}$ be an s-dimensional walk satisfying (5.11). Then if $2(\alpha_1 + \cdots + \alpha_s) < 2 - 2p - s$, $E(T_{00}^p)$ exists, while if*

$$2(\alpha_1 + \cdots + \alpha_s) > 2 - 2p - s, \qquad E(T_{00}^p) = \infty.$$

REFERENCES

1. LAMPERTI, J. Criteria for the recurrence or transience of stochastic processes, I. *J. Math. Anal. and Appl.* **1**, 314-330 (1960).
2. HARRIS, T. E. First passage and recurrence distributions. *Trans. Am. Math. Soc.* **73**, 471-486 (1952).
3. LAMPERTI, J. A new class of probability limit theorems. *J. Math. and Mech.* **11**, 749-772 (1962). (An announcement of the main results appeared in *Bull. Am. Math. Soc.* **67**, 267-269 (1961).)

4. DOOB, J. L. "Stochastic Processes." Wiley, New York, 1953.
5. KAC, M. On the notion of recurrence in discrete stochastic processes. *Bull. Am. Math. Soc.* **53**, 1002-1010 (1947).
6. HARRIS, T. E. The existence of stationary measures for certain Markov processes. *Proc. Third Berkeley Symp., Berkeley, 1956*, Vol. II, pp. 113-124.