

# FLUCTUATION THEORY OF RECURRENT EVENTS

BY  
WILLIAM FELLER<sup>(1)</sup>

**1. Introduction.** The problems treated in this paper belong to pure analysis, but they have an intuitive background which can be roughly described as follows. We are concerned with sequences of repeated trials which may be independent or dependent; for simplicity we shall assume that each trial can result only in denumerably many events. We shall study a repetitive pattern or event  $\mathcal{E}$  characterized by the property that, as far as  $\mathcal{E}$  is concerned the initial situation repeats itself every time when  $\mathcal{E}$  occurs: the trials following the occurrence of  $\mathcal{E}$  are a replica of the whole sequence. In other words, once  $\mathcal{E}$  has occurred we start from scratch. The following examples will provide an idea of the variety of problems to which the theory applies.

**EXAMPLES.** A classical example with revived modern interest is provided by *success-runs* of fixed length  $r$ . In a sequence of Bernoulli trials we shall say that  $\mathcal{E}$  has occurred at the  $n$ th trial if the last failure has occurred at the trial number  $n-r$ , or  $n-2r$ , or  $n-3r$ , and so on<sup>(2)</sup>. If this is the case any probability of a run or of combinations of runs in trials number  $n+k_1, n+k_2, \dots, n+k_r$  is the same as the corresponding probability for the trials number  $k_1, k_2, \dots, k_r$ . Instead of runs we can consider more complicated patterns to which older methods do not apply and we can also treat problems involving combinations of several patterns. It should be noted that the present theory applies not only to independent trials but to Markov chains as well.

An even more interesting application is to the various types of *return to equilibrium*. For example, in any random-walk problem  $\mathcal{E}$  may be the event of a return to the initial position. In particular, consider repeated tossings of a coin and let  $\mathcal{E}$  be the event that the number of heads equals the number of tails. This particular case has been investigated by P. Lévy [8]<sup>(3)</sup> but his results are unfortunately not sufficiently known. They show that the traditional emphasis (and misinterpretation) of the laws of large numbers caused a complete misconception of what is bound to occur in an actual experiment. If  $x_1, x_2, \dots, x_k$  are the actual observed durations between two successive equalizations (or, alternatively, the waiting times up to the *first* equalization in  $k$  experiments), then the mean  $(x_1 + \dots + x_k)/k$  is not stable but will be

---

Presented to the Society, September 23, 1948; received by the editors July 23, 1948.

(1) Research under an ONR contract at Cornell University.

(2) With this definition an uninterrupted sequence of 17 successes will contain 3 runs of length  $r=5$ . This is at variance with the several usual conventions of counting such runs but our definition simplifies the computations and reduces the theory of runs to a special case of a general theory.

(3) Numbers in brackets refer to the literature cited at the end of the paper.

of the order of magnitude of  $k$ . Even more curiously, let in  $n$  tossings  $f_n$  be the fraction of time where the first player is ahead of the second (so that  $nf_n$  is the number of trials in which the accumulated number of heads exceeds the accumulated number of tails). One expects naively  $f_n$  to be “in probability” near  $1/2$ . Actually the limiting density function is  $U$ -shaped with a minimum at  $f=1/2$  and infinite tails at the end  $f=0$  and  $f=1$ . In a large number of trials the chances are, respectively,  $2/3$ ,  $1/5$ , and  $1/20$  that the duration of the lead of one player will extend over the fraction  $.75$ ,  $.975$ , or  $.9985$  of the entire duration leaving the other player in comfort only during roughly  $1/4$ ,  $1/40$ , or  $1/65$  of the time (P. Lévy’s arcsin law). This unfamiliar behavior is due to the fact that here  $\mathcal{E}$  has an infinite mean recurrence time. On the other hand, the situation is typical of many recurrence problems and shows the interest in finding general limit theorems covering such applications.

Further examples of recurrent events are provided by the *zeros of sums*  $S_n = X_1 + \dots + X_n$  of integral-valued mutually independent random variables with a common distribution; the passage of the system through a specified state of a discrete *Markov chain*; coincidences in multiple trials (for example, in tossing several coins  $\mathcal{E}$  may denote the event that the accumulated numbers of heads are the same for all coins). A most familiar application is to *renewal theory*, where each renewal or birth marks a recurrent event. Finally, most of the waiting time and counter problems, as well as many problems in telephone traffic deal with recurrent events, although there we have a continuous time parameter instead of discrete trials.

*Summary of results.* Our point of departure is a new theorem on power series which shows directly that the probability  $u_n$  that  $\mathcal{E}$  occurs at the  $n$ th trial always tends to a limit except when  $\mathcal{E}$  can occur only after  $t$ ,  $2t$ ,  $3t$ ,  $\dots$  trials, in which case  $u_{nt}$  tends to a limit. As is well known, a great many papers were and are being written to prove this theorem under various conditions and we have here the most general result. It permits many applications. In particular, we shall show that the ergodic properties of finite and denumerable Markov chains follow almost at once (§10). As has often been emphasized, the two cases are very different from both an algebraic and an analytic standpoint. The new method applies equally to both cases and means a simplification even in the case of finite chains. Most of §10 on Markov chains is independent of the remainder and should be readable directly after Theorem 1.

§4 contains further theorems of the “renewal” type. Theorem 2 finds many applications, in particular for a new method of finding the second moment of the recurrence times of finite or infinite Markov chains (Theorem 16). The remaining theorems of §4 are not used in the sequel, and Theorem 2 only in an incidental manner.

In §5 we turn to the study of the random variable  $N_n$  which gives the number of occurrences of  $\mathcal{E}$  in  $n$  trials. If the recurrence time of  $\mathcal{E}$  has finite

variance,  $N_n$  obeys the central limit theorem, the law of the iterated logarithm, and so on. More challenging and also of greater physical interest is the case where the recurrence time of  $\mathcal{E}$  has infinite expectation, or finite expectation but infinite variance. We obtain all possible limiting distributions of  $N_n$ . They are related to the stable laws, but the relationship shows a curious difference according as the index of the stable law is greater or smaller than 1 (Theorem 7). As a by-product we obtain the result that the Mittag-Leffler function is the Laplace transform of a distribution related to the stable laws of index  $\alpha < 1$  (Theorem 11 of §7). The new identity was verified by direct calculations by Pollard [12]. The lemma of §5 refers to the convergence of a sum of independent random variables towards a stable law. In amplification of Doeblin's results [3] we find the norming constants and give, incidentally, a direct proof of the sufficiency of Doeblin's necessary and sufficient conditions. In §6 the generating function and moments of  $N_n$  are found.

§8 treats "strong" laws. In the case of a finite variance of the recurrence time we obtain the law of the iterated logarithm in a trivial way. More interesting is the determination of precise lower bounds for  $N_n$  in the case of infinite moments of the recurrence time. This limit theorem is of a new type. Unfortunately we are at present unable to give *upper* bounds which are governed by limit laws of a quite different type. The only special case known is that where  $\mathcal{E}$  means the equalization of the numbers of heads and tails in coin tossing (Chung and Hunt [2]).

In §9 the theory is illustrated by its application to the theory of success-runs. In particular, the Poisson distribution for long runs discovered by von Mises [10] is derived in an easy way. The theory of success-runs can be treated by the method of difference equations (cf. Uspensky [16]). The new method has the advantage that it applies without change to other recurrent patterns where the difference equation method leads to difficulties or is inapplicable. For example, the difference equation method requires essential modifications even for the next simple case where runs of either kind are counted. §10 contains applications to Markov chains.

*Formal definition of recurrent events.* This paper and the usual applications of our theory require only formal manipulations with probabilities, but the mathematically minded reader has a right to a rigorous definition. Underlying the theory is the sample space whose points are formed by the infinite sequences  $(j_1, j_2, \dots)$  where the  $k$ th place is occupied by a positive integer (representing a "possible outcome of the  $k$ th trial"). In this space a probability measure is defined<sup>(4)</sup>. A repetitive event  $\mathcal{E}$  is defined by a rule which for every point and every  $n$  permits to decide whether or not  $\mathcal{E}$  has occurred

---

(4) There is no loss of generality in admitting *all* sequences  $\{j_k\}$  since we can attribute probability zero to the undesirable ones. Our model includes the case of variable trials, for example, the case where tossing a coin and throwing dice follow each other. The sample space is a product space, but *the measure is not necessarily a product measure*.

at the  $n$ th trial. This means simply a classification of all finite sequences  $\{j_1, \dots, j_n\}$  into an  $\mathcal{E}$ -class and an  $\bar{\mathcal{E}}$ -class. Now  $\mathcal{E}$  is recurrent in our sense if for every  $(j_1, j_2, \dots, j_n)$  of the  $\mathcal{E}$ -class and every sequence  $(j_{n+1}, j_{n+2}, \dots, j_{n+v})$  the following two conditions hold. First  $\Pr\{j_1, j_2, \dots, j_n, j_{n+1}, \dots, j_{n+v}\} = \Pr\{j_1, \dots, j_n\} \Pr\{j_{n+1}, j_{n+2}, \dots, j_{n+v}\}$ , and, second,  $(j_1, \dots, j_{n+v})$  and  $(j_{n+1}, \dots, j_{n+v})$  belong to the same class.

**2. Definitions and notations.** For a given recurrent event  $\mathcal{E}$  we shall consider three sets of random variables. The number of trials up to and including the  $n$ th occurrence of  $\mathcal{E}$  will be denoted by  $S_n$ . Then

$$(2.1) \quad S_n = X_1 + \dots + X_n$$

where  $X_k$  is one plus the number of trials between the  $(k-1)$ th and the  $k$ th occurrence of  $\mathcal{E}$ . By definition the  $X_k$  are mutually independent and have a common probability distribution so that

$$(2.2) \quad \Pr\{X_k = v\} = p_v, \quad v = 1, 2, \dots,$$

is independent of  $k$ . The words random variable and probability distribution are here used in a wide sense: it can happen that  $\sum p_v < 1$  implies a positive probability for the nonoccurrence of  $\mathcal{E}$ . The  $X_k$  will be called the *recurrence times*;  $\mathcal{E}$  will be called *certain* or *uncertain* according as  $\sum p_v = 1$  or  $\sum p_v < 1$ . We shall say that  $\mathcal{E}$  is *periodic* with period  $t > 1$  if it can occur only at times  $t, 2t, 3t, \dots$ . In this case

$$(2.3) \quad p_v = 0 \quad \text{whenever } v \neq jt, \quad t > 1.$$

In the case of certain recurrent events we shall have to consider the mean, second moment, and variance of the recurrence times, which can all be finite or infinite. For the *mean recurrence time*  $E(X_k)$  we shall write

$$(2.4) \quad \mu = \sum v p_v.$$

The second factorial moment will be denoted by

$$(2.5) \quad M = \sum v(v-1)p_v$$

and the variance  $\text{Var}(X_k)$  by

$$(2.6) \quad \sigma^2 = M + \mu - \mu^2.$$

These quantities have no meaning for uncertain events.

The common distribution function of  $X_j$  will be denoted by  $F(x)$ ; thus

$$(2.7) \quad F(x) = \sum_{v=0}^{[x]} p_v.$$

The last set of random variables to be considered is  $N_k$ , the *number of realizations of  $\mathcal{E}$  in  $k$  trials*. We shall see that in the case  $\mu = \infty$  the  $N_k$  have no resemblance to sums of independent variables.

In addition to the probabilities  $p_n$ , we shall study the probability  $u_n$  that  $\mathcal{E}$  occurs at the  $n$ th trial (for the first time or not). We shall put  $u_0=1$ . Then, from the definition,

$$(2.8) \quad u_n = u_0 p_n + u_1 p_{n-1} + \cdots + u_{n-1} p_1 \quad (n \geq 1).$$

### 3. The asymptotic behavior of $\{u_n\}$ .

**THEOREM 1.** (a) *In order that  $\mathcal{E}$  be uncertain it is necessary and sufficient that*

$$(3.1) \quad \sum u_n < \infty.$$

(b) *If  $\mathcal{E}$  is certain and not periodic then*

$$(3.2) \quad \lim_{n \rightarrow \infty} u_n = \frac{1}{\mu}$$

*where  $\mu$  is the mean recurrence time (in particular,  $u_n \rightarrow 0$  if, and only if,  $\mu = \infty$ ).*

(c) *If  $\mathcal{E}$  is certain and has period  $t$  then*

$$(3.3) \quad \lim_{n \rightarrow \infty} u_{nt} = \frac{t}{\mu}$$

*while  $u_n = 0$  when  $n$  is not a multiple of  $t$ .*

Many special cases and theorems on mean convergence are known from renewal theory. The most general result seems to be due to Täcklind [14]. He assumes the existence of higher moments and uses deep tools of Fourier analysis and complex variables. The theorem and the method of proof were generalized by Blackwell [1] to the case of a continuous time parameter.

**Proof.** For the generating functions

$$(3.4) \quad P(s) = \sum_{r=1}^{\infty} p_r s^r, \quad U(s) = \sum_{r=0}^{\infty} u_r s^r$$

we find from (2.8) that  $U(s) - 1 = U(s)P(s)$  or

$$(3.5) \quad U(s) = \frac{1}{1 - P(s)}.$$

(a) If  $\mathcal{E}$  is uncertain then  $P(1) < 1$  and the power series of  $U(s) = \sum P^r(s)$  converges for  $|s| \leq 1$  by the fundamental theorem on double series. (b) In this case  $P(1) = 1$ , and the statement is identical with a theorem proved for the present purposes and by elementary methods by Erdős, Feller, and Pollard [4]. (c) In this case both  $P(s)$  and  $U(s)$  are power series in  $s^t$  and the statement reduces to (b) if  $s^t$  is introduced as independent variable.

*Applications.* The ergodic properties of denumerable Markov chains are a simple corollary to Theorem 1 (cf. §10). The probability that in  $2n$  Bernoulli trials the number of successes equals the number of failures is  $u_{2n}$

$= (C_{2n,n})p^nq^n$ . If  $p \neq 1/2$  we have  $\sum u_{2n} = U(1) = \{1 - 4pq\}^{-1/2} = |p - q|^{-1} < \infty$ ; thus the event is uncertain and the probability of its realization  $P(1) = 1 - U^{-1}(1) = 1 - |p - q|$ . If  $p = q = 1/2$  then  $u_{2n} \sim (\pi n)^{-1/2}$  so that the event is certain but has infinite mean recurrence time. More generally we have the following result due to Pólya [13]. In a symmetric unrestricted random walk in  $r$  dimensions the return to the origin is a certain or uncertain event according as  $r \leq 2$  or  $r > 3$ . We had this result for  $r = 1$ . For  $r = 2$  we have

$$\begin{aligned} u_{2n} &= \sum_{k=0}^n C_{2n,2k} C_{2k,k} C_{2n-2k, n-k} 4^{-2n} \\ &= 4^{-2n} C_{2n,n} \sum_{k=0}^n (C_{n,k})^2 = 4^{-2n} (C_{2n,n})^2 \sim (\pi n)^{-1} \end{aligned}$$

so that  $\sum u_n$  diverges. In  $r$  dimensions  $u_{2n}$  is of the order of magnitude  $n^{-r/2}$  so that  $\sum u_n$  converges for  $r \geq 3$ . Similarly, if  $r$  coins are tossed the probability of ever finding them all in the same phase is 1 if  $r \leq 3$  but less than 1 if  $r > 3$ .

**4. Further “renewal” theorems.** We shall now derive further information concerning the asymptotic behavior of  $u_n$  under the additional hypothesis that the moment  $M$  defined in (2.5) is finite. We shall formulate all theorems for non-periodic events. If  $\mathcal{E}$  has period  $t > 1$  the theorems require a trivial reformulation.

**THEOREM 2.** Suppose that  $\mathcal{E}$  is certain and non-periodic and that  $M < \infty$ . Let

$$(4.1) \quad U_n = u_0 + u_1 + \cdots + u_n.$$

Then

$$(4.2) \quad U_n - \frac{n+1}{\mu} \rightarrow \frac{M}{2\mu^2}.$$

The theorem was proved by Täcklind [14] assuming the existence of higher moments than  $M$ . He uses Fourier analysis whereas our proof is elementary.

**Proof.** It is convenient to use the following notations which will in part be used also in the following sections. We put formally

$$(4.3) \quad q_k = \sum_{v=k+1}^{\infty} p_v, \quad r_k = \sum_{v=k+1}^{\infty} q_v$$

and

$$(4.4) \quad Q(s) = \sum_{k=0}^{\infty} q_k s^k, \quad R(s) = \sum_{k=0}^{\infty} r_k s^k.$$

Then

$$(4.5) \quad 1 - P(s) = (1 - s)Q(s),$$

$$(4.6) \quad \mu - Q(s) = (1 - s)R(s)$$

and hence

$$(4.7) \quad \mu = P'(1) = Q(1) = \sum_0^{\infty} q_k,$$

$$(4.8) \quad M = P''(1) = 2Q'(1) = 2R(1) = 2 \sum_0^{\infty} r_k.$$

Accordingly,  $M < \infty$  implies that all  $q_k$  and  $r_k$  are finite.

Now the left side of (4.2) is the coefficient of  $s^n$

$$(4.9) \quad \frac{1}{1 - s} \left\{ U(s) - \frac{1}{\mu(1 - s)} \right\} = \frac{R(s)}{\mu \{ 1 - P(s) \}}.$$

Theorem 1(b) guarantees that the coefficient of  $s^n$  in the right-hand member tends to  $\mu^{-2}R(1)$  and in view of (4.8) this proves the assertion.

**COROLLARY.** *Under the conditions of Theorem 2 we have*

$$(4.10) \quad \lim_{s \rightarrow 1} \left\{ U(s) - \frac{1}{\mu(1 - s)} \right\} = \frac{M}{2\mu^2}.$$

**Proof.** Theorem 2 states that the power series of the term within braces converges, for  $s = 1$ , to the right side. The assertion (4.10) follows from Abel's theorem.

**THEOREM 3<sup>(6)</sup>.** *Under the conditions of Theorem 2*

$$(4.11) \quad \sum \left| u_n - \frac{1}{\mu} \right| < \infty.$$

**Proof.** We have to prove that the power series for

$$(4.12) \quad U(s) - \frac{1}{\mu(1 - s)} = \frac{R(s)}{\mu Q(s)}$$

converges absolutely for  $s = 1$ . Now  $R(s)$  has positive coefficients and converges for  $s = 1$ . The same is true of  $Q(s)$ . Moreover,  $Q(s) \neq 0$  for  $|s| \leq 1$ . In fact, from  $Q(s_0) = 0$  it would follow that  $1 - P(s_0) = (1 - s_0)Q(s_0) = 0$ . However,  $P(s)$  has non-negative coefficients and therefore  $|P(s_0)| \leq P(|s_0|) \leq 1$ . Now  $P(1) = 1$ . For  $|s_0| \leq 1$  we can have  $P(|s_0|) = 1$  only if all terms in the expansion are positive, that is, if for every  $\nu$  with  $p_\nu \neq 0$  we have  $s_0^\nu = 1$ . Then

---

<sup>(6)</sup> Theorems 3 and 4 will not be used in the sequel.

$s_0$  is a root of unity. If  $t$  is the smallest occurring index  $P(s)$  is a power series in  $s^t$  and  $\mathcal{E}$  periodic against our hypothesis. It follows<sup>(6)</sup> from a theorem of Wiener on the absolute convergence of the reciprocal of a Fourier series that the power series for  $Q^{-1}(s)$  converges absolutely on the unit circle.

**THEOREM 4.** *Under the hypotheses of Theorem 2*

$$(4.13) \quad u_n = \frac{1}{\mu} + o\left(\frac{1}{n}\right).$$

**Proof.** Obviously  $n(u_n - 1/\mu)$  is the coefficient of  $s^{n-1}$  in

$$(4.14) \quad U'(s) - \frac{1}{\mu(1-s)^2} = \frac{1}{1-P(s)} \left\{ \frac{R(s)}{\mu} - \frac{Q'(s)}{Q(s)} \right\}.$$

We have just seen that the power series of the term within braces converges absolutely for  $s=1$  and its value is 0 by (4.8). The coefficients of the factor  $\{1-P(s)\}^{-1}$  converge to  $\mu^{-1}$ , and therefore the coefficient of  $s^n$  tends to zero.

**5. Limit theorems for  $N_n$ .** We start from the obvious remark that

$$(5.1) \quad \Pr\{N_n \geq k\} = \Pr\{S_k \leq n\}.$$

Now  $S_k$  is, by (2.1), the sum of mutually independent random variables. If these have finite variance  $\sigma^2$ , then all desirable information is provided by the central limit theorem. In fact, if  $\sigma^2 < \infty$ , then for every fixed  $x$

$$(5.2) \quad \Pr\{S_k \leq k\mu + k^{1/2}\sigma x\} \rightarrow \Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x e^{-y^2/2} dy.$$

Putting  $n=k\mu+k^{1/2}\sigma x$  we find

$$(5.3) \quad k = \frac{n}{\mu} - \frac{x\sigma}{\mu} k^{+1/2} \sim \frac{n}{\mu} - \frac{x\sigma}{\mu^{3/2}} n^{1/2}$$

and hence we have the following theorem.

**THEOREM 5.** *If the recurrence times have finite variance then for every fixed  $x$*

$$(5.4) \quad \Pr\left\{N_n \geq \frac{n}{\mu} - \frac{x\sigma}{\mu^{3/2}} n^{1/2}\right\} \rightarrow \Phi(x).$$

Note that this does not prove that the mean and variance of  $N_n$  are asymptotically equal  $n/\mu$  and  $\sigma\mu^{-3/2}$ : the truth of this statement will be proved in the following section. Simple as the remark leading to Theorem 5 is, it replaces several proofs given by various methods in special cases.

The situation becomes less obvious and more interesting when we pass to the case where the recurrence time has infinite variance  $\sigma^2$ . We saw that

---

(6) Cf. Erdős, Pollard, Feller [4].

this case is of particular importance even in the simplest applications. Now (5.1) shows that  $N_n$  will have a limiting distribution if, and only if,  $S_n$  has one. Necessary and sufficient conditions for this case were given by Doeblin [3, Theorem 5]. Transposed to the present case they can be formulated as:

**THEOREM 6.** *If  $\mathcal{E}$  is a certain event then in order that there exist two sequences of norming constants such that*

$$(5.5) \quad \Pr \{N_n < \mu_n + x a_n\} \rightarrow V(x)$$

*where  $V(x)$  is a distribution function different from  $\Phi(x)$  it is necessary and sufficient that the distribution  $F(x) = \Pr \{X_i \leq x\}$  of the recurrence times has the property*

$$(5.6) \quad 1 - F(x) = x^{-\alpha} h(x)$$

*where  $0 < \alpha < 2$  is a constant and*

$$(5.7) \quad \lim_{x \rightarrow \infty} \frac{h(cx)}{h(x)} \rightarrow 1$$

*for every constant  $c$ .*

Note that there exist exceptional cases where  $\sigma^2 = \infty$  and the central limit theorem holds. They can be dealt with in the same way but present little interest. Presumably all distributions of interest in applications have either finite variance or the property (5.6). In the latter case the limiting distribution  $V(x)$  will naturally be related to quasi-stable laws of index  $\alpha$  and we now proceed to investigate this relationship.

We shall denote by  $G_\alpha(x)$  the *stable law whose characteristic function is given by*

$$(5.8) \quad \gamma_\alpha(z) = \exp - \left\{ |z|^\alpha \left( \cos \frac{\pi\alpha}{2} - i \sin \frac{\pi\alpha}{2} \operatorname{sgn} z \right) \Gamma(1 - \alpha) \right\}.$$

It is well known that when  $\alpha < 1$  we have  $G_\alpha(x) = 0$  for  $x \geq 0$  while  $G_\alpha(x) > 0$  for all  $x$  when  $\alpha \geq 1$  (cf. Lévy [9]).

**THEOREM 7.** *Suppose that the distribution of the recurrence times is of the form (5.6). Then:*

(1) *If  $0 < \alpha < 1$*

$$(5.9) \quad \Pr \left\{ N_k \geq \frac{x}{c_k} \right\} \rightarrow G_\alpha(x^{-1/\alpha})$$

*where*

$$(5.10) \quad c_k = 1 - F(k) = \Pr \{X_i \geq k\}.$$

(2) *If  $1 < \alpha < 2$*

$$(5.11) \quad \Pr \left\{ N_k \geq \frac{k}{\mu} - \frac{b_k}{\mu^{(\alpha+1)/\alpha}} x \right\} \rightarrow G_\alpha(x)$$

where  $b_k$  is such that

$$(5.12) \quad 1 - F(b_k) \sim \frac{1}{k}.$$

These are the only possible non-normal limiting distributions for  $N_n$ .

The second part of the theorem applies also in the case where  $\alpha=1$  and  $\mu < \infty$ . If  $\alpha=1$  and  $\mu = \infty$  a limit theorem of the form (5.12) holds, but the norming terms  $k/\mu$  have to be replaced by a sequence  $\mu_n$  which increases more slowly. This case is rather uninteresting and the explicit determination of  $\mu_n$  requires lengthy computations which are here omitted. The proof of Theorem 7 is based on the following lemma which amplifies Doeblin's criterion.

LEMMA<sup>(7)</sup>. Let  $\{X_j\}$  be an arbitrary sequence of mutually independent positive random variables with a common distribution function  $F(x)$  which satisfies the condition (5.6)–(5.7). Let  $S_n = X_1 + \dots + X_n$  and define  $b_n$  by (5.12). Then:

(1) If  $\alpha < 1$

$$(5.13) \quad \Pr \{S_n \leq xb_n\} \rightarrow G_\alpha(x).$$

(2) If  $\alpha > 1$  and  $\mu = E(X_j) < \infty$

$$(5.14) \quad \Pr \{S_n \leq n\mu + xb_n\} \rightarrow G_\alpha(x).$$

**Proof.** Let  $\alpha < 1$ . The characteristic function of  $X_j$  is

$$(5.15) \quad \phi(z) = \int_0^\infty e^{izx} dF(x) = 1 + iz \int_0^\infty e^{izx} \{1 - F(x)\} dx.$$

For simplicity let us take  $z > 0$ . From (5.7)

$$(5.16) \quad \begin{aligned} \phi\left(\frac{z}{b_n}\right) &= 1 + iz \int_0^\infty e^{izx} (b_n x)^{-\alpha} h(b_n x) dx \\ &= 1 + iz \{1 - F(b_n)\} \int_0^\infty e^{izx} x^{-\alpha} dx \frac{h(b_n x)}{h(b_n)}. \end{aligned}$$

Now here  $1 - F(b_n) \sim 1/n$  by definition. The factor  $h(b_n x)/h(b_n) \rightarrow 1$  for every fixed  $x$ . To see that a formal passage to the limit under the integral sign is permitted it suffices to note that a truncation of the last integral at  $L > 0$  introduces an error,

---

(7) This lemma implies the sufficiency of Doeblin's conditions which is here proved in a direct way.

$$(5.17) \quad \frac{iz}{b_n} \int_{Lb_n}^{\infty} e^{izx/b_n} \{1 - F(x)\} dx.$$

Since  $F(x)$  is monotonic, this error is smaller in absolute value than

$$(5.18) \quad \frac{2z}{b_n} \frac{\pi b_n}{z} \{1 - F(Lb_n)\} = 2\pi(Lb_n)^{-\alpha} h(Lb_n) \sim 2\pi L^{-\alpha} \{1 - F(b_n)\}$$

which is  $o(1 - F(b_n))$  as  $L \rightarrow \infty$ . Hence, putting

$$(5.19) \quad A = iz \int_0^{\infty} e^{izx} x^{-\alpha} dx = -z^{\alpha} \Gamma(1 - \alpha) e^{-i\pi\alpha/2}$$

(cf. P. Lévy [9, p. 200]) we have from (5.16)

$$(5.20) \quad \phi(z/b_n) = 1 + A/n + o(1/n).$$

The characteristic function of  $S_n/b_n$  is  $\phi^n(z/b_n)$  and tends to  $e^A$ . This proves (5.13). In the case  $\alpha \geq 1$  we must take for  $\phi(z)$  the characteristic function of  $X_j - \mu$  and write it in a form which will lead in the limit to a convergent integral thus

$$(5.21) \quad \begin{aligned} \phi(z) &= \int_0^{\infty} e^{iz(x-\mu)} dF(x) \\ &= e^{-iz\mu} + iz\mu - iz \int_0^{\infty} (1 - e^{iz(x-\mu)}) \{1 - F(x)\} dx. \end{aligned}$$

Now

$$(5.22) \quad e^{-iz\mu/b_n} + iz\mu/b_n = 1 + O(b_n^{-2}) = 1 + o(n^{-1}),$$

and the remainder of the proof is not essentially different from the proof in the first case.

To prove (5.9) we let the integers  $n$  and  $k$  be functions of each other so that

$$(5.23) \quad k \sim b_n x.$$

Then by (5.12) and (5.6)–(5.7)

$$(5.24) \quad \begin{aligned} \frac{1}{n} &\sim 1 - F(b_n) \sim 1 - F\left(\frac{k}{x}\right) \sim h(k) \left(\frac{k}{x}\right)^{-\alpha} \\ &\sim \{1 - F(k)\} x^{+\alpha} = c_k x^{+\alpha}. \end{aligned}$$

Therefore

$$(5.25) \quad \Pr \{S_n \leq b_n x\} \sim \Pr \{S_n \leq k\} = \Pr \{N_k \geq n\} \sim \Pr \left\{N_k \geq \frac{x^{-\alpha}}{c_k}\right\}.$$

The extreme left member tends to  $G_\alpha(x)$  and (5.9) follows by a change of variables.

To prove (5.11) note that (5.7) implies that  $h(x)=o(x^\epsilon)$  for every  $\epsilon>0$ . It follows then from (5.12) that  $k=o(b_k^{-\alpha+\epsilon})$  or, since  $\alpha>1$ , that  $b_k=o(k)$ . Let

$$(5.26) \quad k \sim n\mu + b_n x.$$

Then  $k/n \rightarrow \mu$  and therefore

$$(5.27) \quad \mu \sim \frac{k}{n} \sim \frac{1 - F(b_n)}{1 - F(b_k)} \sim \left( \frac{b_k}{b_n} \right)^\alpha$$

or

$$(5.28) \quad n \sim \frac{k}{\mu} - \frac{b_n}{\mu} x \sim \frac{k}{\mu} - \frac{b_k x}{\mu^{(\alpha+1)/\alpha}}.$$

The proof is now accomplished by a reasoning analogous to (5.25).

*Illustrative examples.* (a) Theorem 5 contains as special cases the known central limit theorems for runs and for finite Markov chains. It applies also to the number of passages through a state of an infinite Markov chain provided that the corresponding recurrence time has finite variance. It can easily be shown that every distribution  $\{p_n\}$  is possible as the distribution of the recurrence time of a state in an infinite Markov chain so that in general the number of passages through a given state need not have an asymptotic distribution. However, such things will occur in artificially constructed chains and in general our limit theorems will apply. As random walk problems show the case  $\mu=\infty$  must be considered normal for infinite chains.

(b) An explicit formula for  $G_\alpha(x)$  is known<sup>(8)</sup> for  $\alpha=1/2$ :

$$(5.29) \quad G_{1/2}(x) = 2 \left\{ 1 - \Phi \left( \left( \frac{\pi}{2x} \right)^{1/2} \right) \right\} \quad \text{for } x > 0$$

with  $G_{1/2}(x)=0$  for  $x \leq 0$ . This limiting form occurs in the coin tossing problem mentioned in the introduction. Let  $\mathcal{E}$  be the event of an equalization of the number of heads and tails. Then  $\mathcal{E}$  has period 2 and

$$(5.30) \quad U(s) = \sum_0^\infty C_{2n,n}(s/4)^{2n} = (1 - s^2)^{-1/2}.$$

Hence  $P(s) = 1 - (1 - s^2)^{1/2}$ , whence  $p_{2n} \sim 1/(2\pi^{1/2} n^{3/2})$  and  $1 - F(x) \sim (2/\pi x)^{1/2}$ . Therefore we can take  $c_k^{-1} = (\pi k/2)^{1/2}$  and obtain from (5.9) and (5.29) for  $x > 0$

$$(5.31) \quad \Pr \{ N_k \geq xk^{1/2} \} \rightarrow 2 \{ 1 - \Phi(x) \}.$$

---

(8) For  $z > 0$  the characteristic function is  $\exp \{ -(1-i)(\pi z/2)^{1/2} \}$  and therefore the Laplace transform  $\exp \{ -(\pi z)^{1/2} \}$  and (5.19) can be found in all tables. Explicit formulas for  $G_\alpha(x)$ ,  $\alpha < 1$ , in terms of infinite series of Whittaker functions are found in Pollard [11].

This, then, is the *limiting distribution for the number  $N_k$  of returns to equilibrium (or waves) in tossing a coin*. It is curious to find the *truncated normal distribution* but it must be remembered that this will occur only when the critical exponent  $\alpha = 1/2$ . Note that the mean  $E(N_k)$  increases like  $k^{1/2}$  and not like  $k$  as "common sense" would expect it. In the case of a random walk in three dimensions the index of the limiting distribution is  $\alpha = 1$ .

### 6. Moments and generating function for $N_n$ . Let

$$(6.1) \quad p_{n,k} = \Pr \{S_k \leq n\} - \Pr \{S_{k+1} \leq n\} = \Pr \{N_n = k\}$$

be the probability that  $n$  trials produce exactly  $k$  events  $\mathcal{E}$ . The familiar manipulations with generating functions show that

$$(6.2) \quad \sum_{n=1}^{\infty} p_{n,k} s^n = P^k(s) \frac{1 - P(s)}{1 - s}.$$

For an investigation of  $N_n$  we have to fix  $n$  and require, instead of (6.2), the generating function

$$(6.3) \quad P_n(z) = \sum_{k=0}^{\infty} p_{n,k} z^k.$$

Comparing (6.2) and (6.3) we find:

**THEOREM 8.** *The generating function (6.3) of the distribution of  $N_n$  is the coefficient of  $s^n$  in*

$$(6.4) \quad \frac{1 - P(s)}{(1 - s)(1 - zP(s))}.$$

*The moments  $E(N_n) = \sum k p_{n,k}$  and  $E(N_n^2) = \sum k^2 p_{n,k}$  are, respectively, the coefficients of  $s^n$  in*

$$(6.5) \quad \frac{P(s)}{(1 - s)(1 - P(s))}$$

*and*

$$(6.6) \quad \frac{P(s) + P^2(s)}{(1 - s)(1 - P(s))^2}.$$

The presence of the factor  $(1 - s)^{-1}$  in these expressions has the effect that we need only to determine the summatory functions of the coefficients corresponding to the other factor. Thus, to find  $E(N_n)$  we need only to find the partial sums of the coefficients of  $P(s)\{1 - P(s)\}^{-1} = -1 + \{1 - P(s)\}^{-1}$ , that is, of  $u_k$ . Hence with the notation (4.1) we have  $E(N_n) = U_n - 1$ . If  $\sigma^2 < \infty$  we can apply Theorem 2 and find

$$(6.7) \quad E(N_n) = \frac{n}{\mu} + \frac{\sigma^2 + \mu + \mu^2}{2\mu^2} - 1 + o(1).$$

Similarly it follows from (6.6) easily that  $E(N_n^2) \sim n^2/\mu^2$ . However, for an appraisal of  $\text{Var}(N_n^2) = E(N_n^2) - E^2(N_n)$  we require terms of order  $n$ . For this purpose let us use the notations (4.4) to put (6.6) into the form

$$(6.8) \quad \frac{2}{\mu^2(1-s)^3} + \frac{1}{1-P(s)} \frac{1}{1-s} \cdot \left\{ \frac{2R(s)(\mu + Q(s))}{\mu^2Q(s)} - 3 \right\} + \frac{1}{1-s}$$

which can be obtained by subtracting successively principal parts. The power series of the term in braces converges for  $s=1$  to  $4R(1)\mu^{-2}-3 = 2(\sigma^2+\mu^2-\mu)\mu^{-2}-3$ . This is then the limit of the coefficient of  $s^n$  in the Taylor series of  $(1-s)^{-1}$  times the term in braces. It follows that

$$(6.9) \quad E(N_n^2) = \frac{(n+2)(n+1)}{\mu^2} + n \left\{ \frac{2(\sigma^2 + \mu^2 - \mu)}{\mu^3} - \frac{3}{\mu} \right\} + o(n).$$

Combining (6.7) and (6.9) we find

$$(6.10) \quad \text{Var}(N_n) \sim n\sigma^2/\mu^3.$$

Thus we have the following theorem.

**THEOREM 9.** *If the recurrence time of  $\mathcal{E}$  has finite mean and variance  $\sigma^2$ , then  $E(N_n)$  and  $\text{Var}(N_n)$  are given by (6.7) and (6.10).*

If  $\sigma^2 = \infty$  the mean and variance of  $N_n$  remain finite and their determination is of particular interest. In the extreme cases where  $N_n$  has no limiting distribution the behavior of  $E(N_n)$  can be rather unpleasant. The following theorem is typical for the case of smooth distributions and presumably covers most special cases of interest.

**THEOREM 10.** *Suppose that the distribution function  $F(x)$  of the recurrence time satisfies*

$$(6.11) \quad 1 - F(x) \sim Ax^{-\alpha}$$

where  $0 < \alpha < 2$ . Then:

(a) If  $1 < \alpha < 2$

$$(6.12) \quad E(N_n) = \frac{n}{\mu} + \frac{A}{(\alpha-1)(2-\alpha)\mu^2} n^{2-\alpha} + o(n^{2-\alpha})$$

and

$$(6.13) \quad \text{Var}(N_n) \sim \frac{2A}{(2-\alpha)(3-\alpha)\mu^3} n^{3-\alpha}$$

where  $\mu$  is the mean recurrence time.

(b) If  $0 < \alpha < 1$

$$(6.14) \quad E(N_n) \sim \frac{\sin \alpha\pi}{A\alpha\pi} n^\alpha$$

and

$$(6.15) \quad \text{Var}(N_n) \sim \left\{ \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha + 1/2)} \pi^{1/2} 2^{1-2\alpha} - 1 \right\} \frac{\sin^2 \alpha\pi}{A^2 \alpha^2 \pi^2} n^{2\alpha}.$$

Note that in the second case  $E(N_n)$  increases more slowly than  $n$  and that the standard deviation is of the same order of magnitude.

**Proof.** Let  $\alpha > 1$ . Using again the notation (4.4) we write (6.5) in the form

$$(6.16) \quad \frac{1}{\mu(1-s)^2} + \frac{1}{1-s} \cdot \frac{R(s)}{\mu^2} - \frac{1}{1-s} + \frac{R^2(s)}{\mu^2 Q(s)}.$$

The coefficient of  $s^n$  in the first term is  $(n+1)/\mu$ . The coefficient of  $s^n$  in the next term is  $r_0 + r_1 + \dots + r_n$ . Now  $q_n = 1 - F(n)$  or  $An^{-\alpha}$  and therefore  $r_n = q_{n+1} + \dots \sim An^{-(\alpha-1)}/(\alpha-1)$  and, since  $0 < \alpha - 1 < 1$ , we find  $r_0 + \dots + r_n \sim An^{2-\alpha}/(\alpha-1)(2-\alpha)$ . As for the last term, we know that the Taylor series of  $Q^{-1}(s)$  converges absolutely for  $s=1$ . The coefficient of  $s^n$  in  $R^2(s)$  is obviously  $O(n^{3-2\alpha}) = o(n^{2-\alpha})$  so that (6.13) is true. To prove (6.13) it should be verified that (6.6) can be written in the form

$$(6.17) \quad \frac{2}{\mu^2(1-s)^3} + \frac{R(s)}{(1-s)^2} \cdot \frac{4}{\mu^3} + \rho(s)$$

where  $\rho(s)$  is a power series whose  $n$ th coefficient is  $O(n)$ . The coefficients of  $(1-s)^{-2}R(s)$  are obtained by summing  $r_k$  twice and we find

$$(6.18) \quad E(N_n^2) = \frac{(n+2)(n+1)}{\mu^2} + \frac{4An^{3-\alpha}}{\mu^3(3-\alpha)(2-\alpha)(\alpha-1)} + O(n).$$

A comparison of (6.12) and (6.18) yields (6.13).

The proof in the case  $\alpha < 1$  is less elementary. From  $q_n = 1 - F(n) \sim An^{-\alpha}$  we can conclude that  $Q(s) \sim A\Gamma(1-\alpha)(1-s)^{\alpha-1}$  as  $s \rightarrow 1$  (cf. Titchmarsh [15, p. 225]). Hence

$$(6.19) \quad \frac{1}{1 - P(s)} \sim \frac{1}{A\Gamma(1-\alpha)(1-s)^\alpha}.$$

Using a well known Tauberian theorem (Titchmarsh [15, p. 242]) we conclude that the coefficient of  $s^n$  in  $\{(1-s)(1-P(s))\}^{-1}$  is asymptotically

$$(6.20) \quad \frac{n^\alpha}{A\Gamma(1-\alpha)\Gamma(1+\alpha)} = \frac{n^\alpha \sin \alpha\pi}{A\alpha\pi}.$$

This proves (6.14) and (6.15) follows by a similar argument.

**7. A formula for stable distributions.** Formula (6.4) is useful for many purposes. As an application we shall show that it permits a direct derivation of the limiting distribution (5.9). We obtain an explicit formula for the characteristic function or Laplace transform of the limiting distribution. In this way we are led to the following theorem.

**THEOREM 11.** *Let  $0 < \alpha < 1$  and let  $\phi(\lambda)$  be the Laplace transform of the distribution function  $1 - G_\alpha(x^{-1/\alpha})$*

$$(7.1) \quad \phi(\lambda) = - \int_0^\infty e^{-\lambda x} dG_\alpha(x^{-1/\alpha}).$$

Then

$$(7.2) \quad \phi(\lambda \Gamma(1 - \alpha)) = \sum_{v=0}^{\infty} (-1)^v \frac{\lambda^v}{\Gamma(1 + v\alpha)}.$$

**Proof.** We prove the theorem indirectly<sup>(9)</sup> by showing that (7.2) is the limiting form of the Laplace transform  $\phi_k(\lambda)$  of the distribution  $\Pr\{N_k c_k \leq x\}$ . Now  $\phi_k(\lambda)$  is obtained from the generating function (6.3) on replacing  $z$  by  $\exp(-\lambda c_k)$ . Accordingly

$$(7.3) \quad \phi_k(\lambda) = \text{coef. of } s^k \text{ in } \frac{1 - P(s)}{(1 - s)\{1 - e^{-\lambda c_k} P(s)\}}.$$

For simplicity we consider the case where  $F(x)$  is a distribution function with  $1 - F(x) \sim x^{-\alpha}$  (in which case  $\mathcal{E}$  is certain). Then we may put  $c_k = k^{-\alpha}$  (cf. (5.10)). The fraction to the right admits of an expansion

$$(7.4) \quad \phi_k(\lambda) = \text{coef. of } s^k \text{ in } \frac{1}{1 - s} \sum_{v=0}^{\infty} (-1)^v \left\{ \frac{P(s)}{1 - P(s)} (1 - e^{-\lambda c_k}) \right\}^v.$$

Now  $1 - F(n) = p_{n+1} + p_{n+2} + \dots = q_n \sim n^{-\alpha}$ . Therefore as  $s \rightarrow 1$

$$(7.5) \quad \frac{P(s)}{1 - P(s)} = \frac{P(s)}{(1 - s)Q(s)} \sim \frac{1}{\Gamma(1 - \alpha)(1 - s)^\alpha}$$

(Titchmarsh [15, p. 225]). Hence (Titchmarsh [15, p. 242])

$$(7.6) \quad \text{coef. of } s^k \text{ in } \frac{1}{1 - s} \left( \frac{P(s)}{1 - P(s)} \right)^v \sim \frac{k^{v\alpha}}{\Gamma(1 + v\alpha)\Gamma^v(1 - \alpha)}.$$

Now  $(1 - e^{-\lambda c_k})^v \sim \lambda^v k^{-v\alpha}$  and we obtain from (7.4) formally

$$(7.7) \quad \phi_k(\lambda) \sim \sum_{v=0}^{\infty} (-1)^v \frac{1}{\Gamma(1 + v\alpha)} \left( \frac{\lambda}{\Gamma(1 - \alpha)} \right)^v.$$

---

(9) A direct analytical verification is due to Pollard [12].

A simple appraisal of the remainder justifies the termwise passage to the limit. Since  $\phi(\lambda) = \lim \phi_k(\lambda)$  this proves the theorem.

**8. Strong laws.** The method of §5 permits us to translate every strong law on  $S_n$  into a strong law on  $N_n$ . In the case of a finite variance  $\sigma^2$  we remain within the domain of the law of the iterated logarithm. In the case of infinite variance we have at present only sharp *upper* bounds for  $S_n$  which yield *lower* bounds for  $N_n$ . These are of a new and unfamiliar type and are different according as  $\mu$  is, or is not, finite. One remarkable feature is that this *strong law holds even if  $N_n$  has no limiting distribution*. The situation is probably different for upper bounds for  $N_n$  which presumably obey a law of the iterated logarithm type.

The existence of  $\sigma^2$  is sufficient to assert that  $\lim \sup \frac{(S_k - k\mu)}{\{2k\sigma^2 \log \log k\}^{-1/2}} = 1$  with probability 1 (Hartman and Wintner [7]). Using (5.1) it is easily seen that this implies the following theorem.

**THEOREM 12.** *Whenever the recurrence times have finite variance we have*

$$(8.1) \quad \liminf \frac{N_n - n\mu^{-1}}{\{2\sigma^2\mu^{-3}n \log \log n\}^{1/2}} = -1$$

(by the same token the upper limit is +1).

We can also generalize the more precise laws but shall here consider only the case where not only  $\sigma^2 < \infty$  but

$$(8.2) \quad \sum (j^2 \log \log j) p_j < \infty.$$

If  $\phi(k) \uparrow \infty$  then it is known (Feller [5]) that with probability 1 the inequality

$$(8.3) \quad S_k > k\mu + \sigma k^{1/2} \phi(k)$$

will hold for infinitely or at most finitely many  $k$  according as

$$(8.4) \quad \sum \frac{\phi(k)}{k} e^{-\phi^2(k)/2}$$

diverges or converges. An easy computation shows that this is equivalent to the following theorem.

**THEOREM 13.** *If the distribution of the recurrence time is such that (8.2) holds and if  $\phi(n) \uparrow \infty$  then with probability 1 the inequality*

$$(8.5) \quad N_n - n/\mu < \sigma \mu^{-3/2} n^{1/2} \phi(n)$$

*will be satisfied for infinitely many or at most finitely many  $n$  according as (8.4) diverges or converges.*

Similar statements can be made when (8.2) does not hold but  $\sigma^2 < \infty$ ; in this case further terms appear in the exponent of (8.4) and the formulation

of the law is less pleasing. Needless to say that in (8.5) the inequality may be reversed if  $\phi(n)$  is replaced by  $-\phi(n)$ .

The following theorems are of a less familiar type.

**THEOREM 14.** *Suppose that the distribution  $F(x)$  of the recurrence time has infinite expectation. Let  $\phi(n) \uparrow$  be such that  $\phi(n)/n \downarrow$ . Then with probability 1*

$$(8.6) \quad N_n < \phi(n)$$

*will be satisfied for infinitely or for at most finitely many  $n$  according as*

$$(8.7) \quad \int \{1 - F(x)\} d\phi(x)$$

*diverges or converges.*

No analogue of this theorem with the inequality reversed is known. In the special case where  $\mathcal{E}$  means a return to equilibrium in a coin tossing game Chung and Hunt [2] found both upper and lower bounds. It can easily be verified that (8.7) is equivalent to their condition.

**Proof.** It is known (Feller [6]) that if  $\psi(k)/k \uparrow$  then with probability 1 the inequality  $S_k > \psi(k)$  will be satisfied for infinitely or at most finitely many  $n$  according as

$$(8.8) \quad \sum \{1 - F(\psi(k))\}$$

diverges or converges. Since  $1 - F(\psi(x))$  decreases monotonically the sum can be replaced by an integral. The theorem now follows by letting  $\psi(y)$  be the inverse function of  $\phi(x)$ .

**THEOREM 15.** *Suppose that the distribution  $F(x)$  of the recurrence times has finite expectation but that for some  $\alpha < 2$  the moment of order  $\alpha$  is infinite. Let  $\phi(n)$  be monotonic such that  $\phi(n)/n$  decreases but  $\phi(n)/n^\lambda$  increases for some  $\lambda < 1$ . Then with probability 1 the inequality*

$$(8.9) \quad N_n < n/\mu - \phi(n)/\mu$$

*will be satisfied for infinitely or at most finitely many  $n$  according as (8.7) diverges or converges.*

No counterpart for upper bounds is known. The conditions on  $\phi(n)$  can easily be relaxed but they seem to take care of all interesting cases. If  $\phi(n)$  is of the order of magnitude  $n^{1/\alpha}$  then (8.9) will certainly be satisfied for infinitely many  $n$ .

**Proof.** Let  $k$  be an integer such that  $k\mu \sim n - \phi(n)$  and put  $\psi(k) = \phi(n)$ . Then of the inequalities (8.9) and

$$(8.10) \quad S_k \geq k\mu + \psi(k)$$

either both or none will be satisfied for infinitely many subscripts. For the lat-

ter the criterion is given by the convergence or divergence of  $\sum \{1 - F(\psi(k))\}$  (Feller [6]) and the assertion of the theorem is that this series and (8.7) converge and diverge together. This follows from the integral criterion.

**9. Application to runs.** As an example of the simplicity with which various theorems can be obtained we shall outline the theory of success-runs of length  $r$  in a series of Bernoulli trials with probabilities  $p$  and  $q$ . According to the definition given in §1 with examples, the probability that a run terminates at the  $n$ th trial is, when  $n$  is not a multiple of  $r$ ,

$$(9.1) \quad u_n = q \{ p^r + p^{2r} + \cdots + p^{[n/r]r} \};$$

if  $n \geq r$  is a multiple of  $r$  the last term should be  $p^n$  instead of  $qp^n$ . Hence

$$(9.2) \quad U(s) = \left\{ \frac{qs}{1-s} p^r s^r + 1 \right\} \sum_{r=0}^{\infty} (ps)^{r^r} = \frac{1-s+qp^r s^{r+1}}{(1-s)(1-p^r s^r)}.$$

From (3.5) we have for the *generating function of the recurrence time* (or first appearance)

$$(9.3) \quad P(s) = \frac{(1-ps)p^r s^r}{1-s+qp^r s^{r+1}}$$

in agreement with the classical result usually derived from difference equations (for example Uspensky [16]). From (9.1) we have the *mean recurrence time*

$$(9.4) \quad \mu = \lim (1/u_n) = (1-p^r)/qp^r,$$

and the second moment is obtained from (4.10). We have thus without further work the moments, the central limit theorem, and even the law of the interated logarithm for the number  $N_n$  of runs in  $n$  trials. However, the method yields also deeper results as, for example, von Mises' [10] discovery that the distribution of "long runs" is nearly a Poisson distribution. To deduce this theorem we start from (6.4) which in our case states that the generating function of the number  $N_n$  of runs in  $n$  trials is

$$(9.5) \quad P_n(z) = \text{coef. of } s^n \text{ in } \frac{1-p^r s^r}{1-s+qp^r s^{r+1}-z(1-ps)p^r s^r}.$$

This is a rational function and the coefficient of  $s^n$  can be found from the partial fraction expansion. After eliminating the extraneous root  $s=1/p$  which is common to numerator and denominator we seek the roots of the reduced denominator. This leads to the equation

$$(9.6) \quad p^r s^r z + qs(1+ps+\cdots+p^{r-1}s^{r-1}) = 1.$$

Since all coefficients to the left are positive this equation has a unique positive root  $s_0$  and all other roots are greater in absolute value. Hence

$$(9.7) \quad P_n(z) \sim A/s_0^{n+1}$$

where

$$(9.8) \quad A = \frac{1 - p^r s_0^r}{1 - (r + 1)p^r s_0^r(q + pz) + rzp^r s_0^{r-1}}.$$

Now if  $r \rightarrow \infty$  then  $A \rightarrow 1$  and

$$(9.9) \quad s_0 = 1 + (1 - z)q p^r + O(p^{2r})$$

as can be seen from the original form of the denominator. Accordingly, if we let  $n \rightarrow \infty$  and  $r \rightarrow \infty$  so that

$$(9.10) \quad n q p^r \rightarrow \lambda$$

then formally from (9.7)

$$(9.11) \quad P_n(z) \sim e^{-\lambda(1-z)}$$

which is the generating function of the Poisson distribution with mean  $\lambda$ . This confirms von Mises' result. It remains to show that (9.7) holds even when  $r \rightarrow \infty$ , that is, that the remaining roots remain uniformly without influence. Now it is easily seen that all roots except  $s_0$  are greater in absolute value than  $1/p$  and their contributions are readily estimated<sup>(10)</sup>.

**10. Application to Markov chains.** Every state  $E_j$  of a denumerable Markov chain with stationary transition probabilities  $p_{jk}$  can be considered as a *recurrent event*. States are therefore automatically classified into certain and uncertain states or, to use the usual terminology, *transient* and non-transient ones. The non-transient states are *null-states* if their mean recurrence time  $\mu$  is infinite. Another and independent classification of states is into *periodic* and non-periodic states.

Let  $p_{jk}^{(n)}$  denote the  $n$ -step transition probabilities, that is, the elements of the matrix  $P^n$  where  $P = (p_{jk})$ . If the state  $E_j$  has period  $t$  then  $p_{jj}^{(nt)} = 0$  for all  $n$  which are not divisible by  $t$ . If  $E_j$  is not periodic then put  $t = 1$ . Theorem 1 tells us that

$$(10.1) \quad \lim_{n \rightarrow \infty} p_{ii}^{(nt)} = p_i$$

always exists. The state  $E_j$  is transient if, and only if,

$$(10.2) \quad \sum_n p_{ii}^{(n)} < \infty$$

<sup>(10)</sup> Computations which give much more than the required result can be found in Uspensky [16] although he considers only (9.3). Now this is essentially a special case of (9.5) for  $z=0$  and the generalization necessitates only few changes.

and a *null-state* if this series diverges but  $p_{jj}^{(n)} \rightarrow 0$ .

Our theory gives us further information. We know that  $1/p_j$  is the mean recurrence time of  $\mathcal{E}_j$  (possibly  $\infty$ ). Theorem 2 gives us an apparently new result. If the recurrence time has finite second moment  $m_2$ , then the second factorial moment  $M = m_2 - \mu$ . Theorem 2 can therefore now be formulated as:

**THEOREM 16.** *For any state  $\mathcal{E}_j$  with finite mean recurrence time  $\mu_j$  we have*

$$(10.3) \quad \lim_{n \rightarrow \infty} \left\{ \sum_{k=0}^n p_{jj}^{(k)} - \frac{n+1}{\mu_j} \right\} = \frac{m_{2j}}{2\mu_j^2} - \frac{1}{2\mu_j}$$

where  $m_{2j}$  is the (possibly infinite) second moment of the recurrence time. In particular, in a finite chain the limit (10.3) exists for every non-transient state (for a transient state the sum converges).

We have thus a new possibility of finding the higher moments of the recurrence times. An alternative method for *finite* chains is as follows. Instead of a single generating function we consider the generating *matrix*

$$(10.4) \quad U(s) = \sum_{k=0}^{\infty} P^k s^k = (sP - I)^{-1}$$

where  $I$  is the identity matrix. For  $s=1$  we have a singularity since 1 is a characteristic value of  $P$ . However, if the chain is not periodic, then  $P^\infty = \lim P^n$  exists and is a matrix whose rows are all equal. The matrix  $U - 1/(1-s)P^\infty$  is regular for  $s=1$  and its diagonal elements are  $(m_{2j} - \mu_j)/2\mu_j^2$  (cf. (4.10)). If  $P$  has period  $t > 1$ , a similar statement applies to the matrix  $P^t$ .

It should be noted that we have proved the central limit theorem and the law of the iterated logarithm not only for the individual states of finite chains but also for all states of infinite chains for which the recurrence time has finite variance. For other states of infinite chains our new limit theorems apply although it should be borne in mind that artificial chains can be constructed for which the number of passages through a given state has no limiting distribution.

*Ergodic properties of Markov chains.* We want finally to show that the ergodic properties of finite and infinite Markov chains alike follow directly from Theorem 1. It suffices to consider *indecomposable* chains; that is, we suppose that it is impossible to find a “closed subset”  $C$  of states such that  $p_{jk}=0$  whenever  $\epsilon_j \in C$ ,  $\epsilon_k \notin C$ . (A closed subset forms a new Markov chain.) Now all states which can be reached from  $\epsilon_j$  (that is, all  $\epsilon_k$  for which there exists an  $n$  such that  $p_{jk}^{(n)} > 0$ ) form a closed set. Therefore every state can be reached from every other.

We know for every  $j$  the limit (10.1) exists. We want to investigate the behavior of  $p_{jk}^{(n)}$ . There exist two integers  $a, b$  such that  $p_{jk}^{(a)} > 0$  and  $p_{kj}^{(b)} > 0$ . Now  $p_{jj}^{(a+n+b)} \geq p_{jk}^{(a)} p_{kk}^{(n)} p_{kj}^{(b)}$  and  $p_{kk}^{(a+n+b)} \geq p_{kj}^{(b)} p_{jj}^{(n)} p_{jk}^{(a)}$ . Therefore the two series  $\sum p_{jj}^{(n)}$  and  $\sum p_{kk}^{(n)}$  converge or diverge together; if  $p_j=0$  then  $p_k=0$  and vice

versa. Hence: in an indecomposable chain *all states are of the same character*: all transient, or all null-states, or all non-null-states; all have the same period.

If all states have period  $t$ , then  $P^t$  is the matrix of a non-periodic chain. It suffices therefore to study the non-periodic case. Let then  $f_{jk}^{(n)}$  be the probability that starting from  $\mathcal{E}_j$  the system will reach  $\mathcal{E}_k$  for the first time at the  $n$ th step (first passage time distribution). Clearly

$$(10.5) \quad p_{jk}^{(n)} = f_{jk}^{(1)} p_{kk}^{(n-1)} + \cdots + f_{jk}^{(n)} p_{kk}^{(0)}.$$

If the states are not transient then a first passage is certain and  $\sum_n f_{jk} = 1$ . Since  $p_{kk}^{(n)} \rightarrow p_k$  it follows from (10.5) that also  $p_{jk}^{(n)} \rightarrow p_k$ . If the states are transient then  $p_{kk}^{(n)} \rightarrow 0$  and hence clearly also  $p_{jk}^{(n)} \rightarrow 0$ . Therefore  $P^n$  tends to the matrix whose rows are  $(p_1, p_2, \dots)$ . This is the theorem to be proved.

#### LITERATURE

2. D. Blackwell, *A renewal theorem*, Duke Math. J. vol. 15 (1948) pp. 145–150.
2. K. L. Chung and G. A. Hunt, *On the zeros of  $\sum_i \pm 1$* , Ann. of Math. vol. 50 (1949) pp. 383–400.
3. W. Doeblin, *Sur l'ensemble de puissances d'une loi de probabilité*, Studia Math. vol. 9 (1941) pp. 71–96.
4. P. Erdős, H. Pollard, and W. Feller, *A property of power series with positive coefficients*, Bull. Amer. Math. Soc. vol. 55 (1949) pp. 201–204.
5. W. Feller, *The law of the iterated logarithm for identically distributed random variables*, Ann. of Math. vol. 47 (1946) pp. 631–638.
6. ——, *A limit theorem for random variables with infinite moments*, Amer. J. Math. vol. 68 (1946) pp. 257–262.
7. P. Hartman and A. Wintner, *On the law of the iterated logarithm*, Amer. J. Math. vol. 63 (1941) pp. 169–176.
8. P. Lévy, *Sur certains processus stochastiques homogènes*, Compositio Mathematica vol. 7 (1939) pp. 283–339.
9. ——, *Théorie de l'addition des variables aléatoires*, Paris, Gauthier-Villars, 1937.
10. R. von Mises, *Das Problem der Iterationen*, Zeitschrift für Angewandte Mathematik und Mechanik, vol. 1 (1921) pp. 298–307.
11. H. Pollard, *The representation of  $e^{-\lambda}$  as a Laplace integral*, Bull. Amer. Math. Soc. vol. 52 (1946) pp. 908–910.
12. ——, *The completely monotonic character of the Mittag-Leffler function  $E_\alpha(-x)$* , Bull. Amer. Math. Soc. vol. 54 (1948) pp. 1115–1116.
13. G. Pólya, *Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassennetz*, Math. Ann. vol. 84 (1921) pp. 149–160.
14. S. Täcklind, *Fourieranalytische Behandlung vom Erneuerungsproblem*, Skandinavisk Aktuariedtskrift (1945) pp. 68–105.
15. E. C. Titchmarsh, *The theory of functions*, 2d ed., Oxford University Press, 1939.
16. J. V. Uspensky, *Introduction to mathematical probability*, New York, McGraw-Hill, 1937.

CORNELL UNIVERSITY,  
ITHACA, N. Y.