

*This research was partially supported by the National Science Foundation.

ASYMPTOTIC FLUCTUATION BEHAVIOR OF SUMS OF
WEAKLY DEPENDENT RANDOM VARIABLES*

by

Walter Philipp

University of Illinois at Champaign-Urbana

and

William F. Stout

University of Illinois at Champaign-Urbana

and

Department of Statistics

University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 950
September 1974

ASYMPTOTIC FLUCTUATION BEHAVIOR OF SUMS OF
WEAKLY DEPENDENT RANDOM VARIABLES

by

Walter Philipp

University of Illinois at Champaign - Urbana

and

William F. Stout

University of Illinois at Champaign - Urbana

and

University of North Carolina at Chapel Hill

1. INTRODUCTION

The purpose of our research is to investigate the asymptotic fluctuation behavior of sums of weakly dependent random variables, such as lacunary trigonometric, mixing, and Gaussian. We present here a brief exposition of the results obtained and a detailed sketch of the method leading to these results. A complete presentation is given in Philipp and Stout (1974).

In essence, $\{x_n\}$ are weakly dependent if

$$E | E(x_{n+k} | x_1, \dots, x_n) | \rightarrow 0$$

as $k \rightarrow \infty$ for each $n \geq 1$. By asymptotic fluctuation behavior, we mean results such as the law of the iterated logarithm, Strassen's functional law of the iterated logarithm, the upper and lower class refinement of the law of the iterated logarithm, and Chung's upper and lower class result for the maxima of partial sums.

We obtain these results by first establishing almost sure invariance principles. The idea of an almost sure invariance principle is due to Strassen (1964, 1965). Strassen proves, among other things, that a martingale with finite variances is with probability one close to Brownian motion on $[0, \infty)$ in a sense made precise in Section 3. The asymptotic fluctuation behavior of Brownian motion has, of course, been thoroughly investigated. In particular, the upper and lower class refinement of the law of the iterated logarithm is known, as well as the functional form of the law of the iterated logarithm [Strassen, (1964)] and the upper and lower class refinement for the maximum of Brownian motion up to time t [Jain and Taylor, (1973)]. Thus, if the approximation of the martingale to Brownian motion given by the almost sure invariance principle is sufficiently close, then all the above fluctuation results for Brownian motion also hold for the martingale.

If we are able to approximate sufficiently closely sums of weakly dependent random variables by a martingale, we can then conclude, in view of the above remarks, that these sums are close to Brownian motion with probability 1. Consequently, the fluctuation results for Brownian motion will continue to hold for the weakly dependent random variables under consideration.

Let us be more specific about this last point. Let $\{x_n\}$ be a sequence of random variables, centered at expectations with finite $(2 + \delta)$ moments.

Put

$$(1.1) \quad S_t = S(t) = \sum_{n \leq t} x_n .$$

Suppose

$$\lim \text{var} \left(n^{-\frac{1}{2}} S_n \right) = \sigma^2 > 0$$

exists and is positive so that without loss of generality we can and do assume $\sigma^2 = 1$. Our goal is to prove for a large class of weakly dependent random variables the basic almost sure invariance principle

$$(1.2) \quad S(t) - X(t) \ll t^{\frac{1}{2}-\eta} \quad \text{a.s.}$$

for some $\eta > 0$. (We use the Vinogradov symbol \ll instead of big O .)

Here $X(t)$ is standard Brownian motion on $[0, \infty)$ and $\eta > 0$ depends on the sequence $\{x_n\}$ considered. Such a result immediately translates almost sure fluctuation results from $\{X(t)\}$ to $\{S(t)\}$ and hence to $\{S_n\}$. Indeed, suppose we have an arbitrary sequence $\{x_n\}$ such that its "bookkeeping function" $\{S(t)\}$ satisfies (1.2). Then the following four theorems are straightforward consequences of known results for Brownian motion.

THEOREM A: Let $\{x_n\}_{n=1}^{\infty}$ be any sequence of random variables satisfying (1.2).

Let $\phi(t)$ be a positive, nondecreasing real-valued function. Then

$$P\{S_n > n^{\frac{1}{2}} \phi(n) \text{ i.o.}\} = 0 \text{ or } 1$$

according as

$$\int_1^{\infty} \frac{\phi(t)}{t} \exp\left(-\frac{1}{2} \phi^2(t)\right) dt$$

converges or diverges.

This result follows from Kolmogorov's test for Brownian motion. For the details of the proof see for example Jain, Jogdeo, and Stout (1974).

Using a recent result on the maximum of Brownian motion due to Jain and Taylor (1973) we get (again, for the details see Jain, Jogdeo, and Stout (1974)).

THEOREM B: Let $\{x_n\}$ and $\phi(t)$ be as in Theorem A. Put

$$M_n = \max_{1 \leq i \leq n} |S_i| .$$

Then

$$P\{M_n < n^{\frac{1}{2}} \phi^{-1}(n) \text{ i.o.}\} = 0 \text{ or } 1$$

according as

$$\int_1^{\infty} \frac{\phi^2(u)}{u} \exp(-8\pi^{-2} \phi^2(u)) du$$

converges or diverges.

Next, let $C[0,1]$ be the space of all real-valued continuous functions $h(t)$, $t \in [0,1]$, with the supremum norm $\|h\| = \sup_{0 \leq t \leq 1} |h(t)|$. Let $K \subset C[0,1]$ be the set of absolutely continuous functions with $h(0) = 0$, $\int_0^1 (h'(t))^2 dt \leq 1$. Let

$$(1.3) \quad S_n(t) = S(nt) = \sum_{k \leq nt} x_k .$$

Define another type of bookkeeping functions $f_n(t) = f_n(t, \omega)$ by

$$(1.4) \quad f_n(t) = (2n \log \log n)^{\frac{1}{2}} S(nt) .$$

THEOREM C: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of random variables satisfying (1.2). Then with probability 1 the sequence of functions $\{f_n(t)\}$ defined by (1.4) is

relatively compact in the topology of uniform convergence and has K as its derived set.

Indeed, by (1.2) and (1.3),

$$S_n(t) - X(nt) \ll (nt)^{\frac{1}{2}-\eta} \ll n^{\frac{1}{2}-\eta} \quad \text{a.s.}$$

uniformly in $0 \leq t \leq 1$. Write $\zeta_n(t) = X(nt)$. Then, of course,

$$(1.5) \quad ||S_n - \zeta_n|| \ll n^{\frac{1}{2}-\eta} \quad \text{a.s., } \eta > 0.$$

Consequently,

$$(n \log \log n)^{-\frac{1}{2}} ||S_n - \zeta_n|| \ll n^{-\eta}.$$

Theorem C follows now from Strassen's (1964) Theorem 1 which states that the conclusion of Theorem C holds for the sequence

$$\{(2n \log \log n)^{-\frac{1}{2}} \zeta_n(t), n \geq 3\}.$$

Similarly, (1.2) implies distribution type invariance principles.

THEOREM D: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of random variables satisfying (1.2).

Then

$$n^{-\frac{1}{2}} S_n(t) \rightarrow W(t)$$

in distribution where $W(t)$ is standard Brownian motion on $[0,1]$ and $S_n(t)$ is defined by (1.4).

Indeed, (1.5) implies that

$$n^{-\frac{1}{2}} (S_n - \zeta_n) \rightarrow 0$$

in probability. But $n^{-\frac{1}{2}} \epsilon_n$ has the same distribution as $\{W(t), 0 \leq t \leq 1\}$ for $n \geq 1$ and the result follows (see e.g. Billingsley (1968) p. 25).

Both Theorems C and D have a large number of corollaries spelled out in detail in Strassen's paper (1964) and Billingsley's book (1968).

In addition to the stationary case, several applications are made to the nonstationary case where

$$\lim n^{-1} E S_n^2 = c \neq 0$$

fails. Then our goal is still to prove the basic relation (1.2), where the definition of $\{S(t)\}$ is modified appropriately:

$$S(t) = \sum_{k=1}^n x_k \quad \text{for} \quad s_n^2 = E S_n^2 \leq t < s_{n+1}^2 .$$

Theorems A-D continue to hold, provided the statements are also modified appropriately.

In Theorem A replace

$$P\{S_n > n^{\frac{1}{2}} \phi(n) \text{ i.o.}\} = 0 \text{ or } 1$$

by

$$P\{S_n > s_n \phi(s_n^2) \text{ i.o.}\} = 0 \text{ or } 1 .$$

Similarly, in Theorem E, replace

$$P\{M_n < n^{\frac{1}{2}} \phi^{-1}(n) \text{ i.o.}\} = 0 \text{ or } 1$$

by

$$P\{M_n < s_n \phi^{-1}(s_n^2) \text{ i.o.}\} = 0 \text{ or } 1 .$$

In Theorems C and D redefine

$$S_n(t) = S(s_n^2 t) , 0 \leq t \leq 1 .$$

In Theorem C, replace (1.3) by

$$f_n(t) = (2 s_n^2 \log \log s_n^2)^{-\frac{1}{2}} S_n(t) .$$

The modifications in the proof of Theorem C are obvious. In Theorem D replace $n^{-\frac{1}{2}} S_n(t)$ by $s_n^{-1} S_n(t)$.

For all cases of weakly dependent random variables that we consider the conclusions of Theorems A, B and C are new. The conclusion of Theorem D, however, is new only in certain cases.

It will be clear that the present method is rather general and is thus applicable to many situations other than those considered in this monograph. We will give an outline of our method in Section 3.

2. STATEMENT OF RESULTS

Let $\{n_k, k \geq 1\}$ be a lacunary sequence of real numbers, i.e. $n_{k+1}/n_k \geq q > 1$ and let $\{a_k\}$ be another sequence of real numbers. Put

$$A_N^2 = \frac{1}{2} \sum_{k \leq N} a_k^2 .$$

Suppose that $A_N \rightarrow \infty$ and that there exists a constant δ with $0 < \delta \leq 1$ such that

$$(2.1) \quad a_k \ll A_k^{1-\delta} .$$

We consider trigonometric series

$$\sum A_k \cos 2\pi n_k \omega + B_k \sin 2\pi n_k \omega = \sum a_k \cos 2\pi n_k (\omega + \beta_k)$$

where for reasons of convenience we put $\beta_k = 0$. For $t \geq 0$ put

$$S(t) = \sum_{k \leq N} a_k \cos 2\pi n_k \omega \quad \text{if} \quad A_N^2 \leq t < A_{N+1}^2 .$$

Then we have the following theorem.

THEOREM 1: Without changing the distribution of the process $\{S(t), t \geq 0\}$ we can redefine the process $\{S(t), t \geq 0\}$ on a richer probability space together with Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}-c\delta} \quad \text{a.s.}$$

for each $c < 1/16$.

Condition (2.1) can quite likely be replaced by a weaker one, say $a_k \ll A_k / \log A_k$. It is also quite likely that the gap condition could be relaxed to $n_{k+1}/n_k \geq 1 + c_k k^{-\frac{1}{2}}$ with $c_k \uparrow \infty$. Perhaps, one could also combine these two possible generalizations.

Theorem 1 can be considered as a refinement of M. Weiss's (1959) law of the iterated logarithm under the restriction (2.1). Also Billingsley's functional central limit theorem for lacunary trigonometric series should be mentioned in this context. (Billingsley (1967)).

Finally, it should be remarked that Theorem 1 includes the unweighted case $a_k = 1$. In this special case, a much simpler proof can be given. Recently and independently from the authors Berkes (1974b) obtained a similar result in the unweighted case, making more stringent assumptions on the sequence $\{n_k\}$.

Given a stochastic sequence $\{x_n, n \geq 1\}$, let F_a^b denote the σ -field generated by x_a, x_{a+1}, \dots, x_b , ($1 \leq a \leq b < \infty$) and F_m^∞ the σ -field generated by x_m, x_{m+1}, \dots . Then the sequence is said to be ϕ -mixing if there exists a sequence $\{\phi(n)\}$ of real numbers with $\phi(n) \downarrow 0$ such that for each $t \geq 1, n \geq 0, A \in F_1^t, B \in F_{t+n}^\infty$ we have

$$(2.2) \quad |P(AB) - P(A)P(B)| \leq \phi(n) P(A) .$$

We assume that

$$(2.3) \quad \sum_{n=1}^{\infty} \phi^2(n) < \infty .$$

This is easily seen to imply the existence of the limit

$$\sigma^2 = \lim_{n \rightarrow \infty} n^{-1} E \left\{ \sum_{k \leq n} x_k \right\}^2 .$$

We assume throughout that $\sigma^2 > 0$, the case $\sigma^2 = 0$ being a degenerate one.

Hence we assume without loss of generality that

$$(2.4) \quad \sigma^2 = 1 .$$

Let

$$S(t) = \sum_{k \leq t} x_k, \quad t \geq 0 .$$

Then we obtain the following almost sure invariance principle.

THEOREM 2: Let $\{x_n, n \geq 1\}$ be a stationary sequence, centered at expectations and satisfying (2.2), (2.3), and (2.4). Suppose that for some $\delta > 0$

$$E|x_1|^{2+\delta} < \infty .$$

Then, without changing the distribution of $\{S(t), t \geq 0\}$, we can redefine the process $\{S(t), t \geq 0\}$ on a richer probability space together with standard Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}\delta/(12+6\delta)+\gamma} \quad \text{a.s.}$$

for each $\gamma > 0$.

REMARK: If $\|x_1\|_\infty < \infty$ then the exponent reduces to $\frac{1}{3} + \gamma$.

Theorem 2 contains one of Reznik's (1968) theorems as a special case. Independently from the authors, Berkes (1974a) establishes a similar result under stronger hypotheses.

It might be interesting to remark that if, instead of $E|x_1|^{2+\delta} < \infty$, only the finiteness of the second moments is assumed then the conclusion of Theorem 2 has to be weakened to

$$S(t) - X(t) = o\left((t \log \log t)^{\frac{1}{2}}\right) \quad \text{a.s.}$$

This result is due to Heyde and Scott (1973).

At the relatively minor cost of strengthening the assumption (2.3) and a more complicated proof, Theorem 2 can be considerably generalized. This is done in the next four theorems. Before stating these four theorems, we make

some preparatory remarks.

Let $\{\xi_n\}$ be a sequence of random variables. Let f be a measurable mapping from the space of doubly infinite sequences $(\dots, \alpha_{-1}, \alpha_0, \alpha_1, \dots)$ of real numbers into the reals. Put

$$(2.5) \quad \eta_n = f(\dots, \xi_{n-1}, \xi_n, \xi_{n+1}, \dots) \quad , \quad n \geq 1 \quad .$$

Instead of ϕ -mixing we shall assume that the ξ_n 's satisfy what we call a retarded strong mixing condition. Denote, as usual, by F_a^b the σ -field generated by the ξ_n 's ($a \leq n \leq b$). We shall assume that

$$(2.6) \quad |P(AB) - P(A)P(B)| \leq \alpha(n) t^{-\kappa}$$

for all $A \in F_1^t$ and all $B \in F_{t+n}^\infty$. Here κ is a nonnegative constant, depending only on the exponent of the moments of the η_n 's and $\alpha(s)$ converges monotonically to zero at a certain rate. The case $\kappa = 0$ gives the so-called strong mixing condition introduced by M. Rosenblatt (1956).

Conditions of the form (2.6) occur in the metric theory of Diophantine approximation (see for example Szűsz (1963) or Philipp (1971), p. 48).

Finally, we shall make no stationarity assumptions. As is customary, we shall assume that η_n can be closely approximated by

$$\eta_{\ell n} = E(\eta_n | F_{n-\ell}^{n+\ell}) \quad .$$

With this notation we have the following theorem:

THEOREM 3: Let $\{\xi_n\}$ be a sequence of random variables and let η_n be defined by (2.5). We shall assume of the function f and the sequence $\{\xi_n\}$

that

$$E\eta_n = 0 .$$

Suppose that there exists constants $0 < \delta \leq 2$ and $C > 0$ such that

$$E|\eta_n|^{2+\delta} \leq C$$

and

$$(2.7) \quad \|\eta_n - \eta_{ln}\|_{2+\delta} \leq C \cdot l^{-(2+7/\delta)}$$

for all $n, l = 1, 2, 3, \dots$. Moreover, suppose that

$$E\left(\sum_{n \leq N} \eta_n\right)^2 = N + o(N^{1-\delta/30})$$

as $N \rightarrow \infty$. Finally, assume that $\{\xi_n\}$ satisfies a retarded strong mixing condition of the form (2.6) with

$$\kappa = \delta/(11 + 4\delta)$$

and

$$(2.8) \quad \alpha(s) \ll s^{-16\delta(1+2/\delta)} .$$

Define a continuous parameter process $\{S(t), t \geq 0\}$ by setting

$$S(t) = \sum_{n \leq t} \eta_n .$$

Then, without changing the distribution of $\{S(t), t \geq 0\}$ we can redefine the process $\{S(t), t \geq 0\}$ on a richer probability space together with standard

Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}-c\delta} \quad \text{a.s. as } t \rightarrow \infty$$

for each $c < 1/588$.

Except for the rates of decay in (2.7) and (2.8), Theorem 3 contains in the stationary case almost all previous results in this direction. For example, it contains Reznik's (1968) and Iosifescu's (1968) theorems on the law of the iterated logarithm for stationary ϕ -mixing random variables except for the convergence rates in (2.7) and (2.8). The same comment holds for Reznik's (1968) law of the iterated logarithm for strong mixing sequences. Notice that in this latter case we only assume a uniform bound on the $(2 + \delta)$ moments of η_n . Reznik either assumes uniform bounds on the η_n 's themselves or on the $(4 + \delta)$ moments.

Theorem 3 also contains recent upper and lower class results of Jain, Jogdeo and Stout (1974) on functionals of a Markov process satisfying Doeblin's condition. Independently of the authors, Berkes (1974a) establishes a result which is essentially the same as Theorem 3 for the special case that $\kappa = 0$.

Again, except for the rates of convergence in (2.7) and (2.8), Theorem 3 also implies a result of Davydov (1970). Moreover, it even implies a non-trivial, though rather weak, remainder term in the central limit theorem for strongly mixing stationary sequences.

Although strict stationarity was not assumed in Theorem 3, restrictions consistent with stationarity on the growth of the variance of the partial

sums of the random variables were imposed. In the same vein we assumed that the $(2 + \delta)$ moments of the random variables were uniformly bounded. In the next two theorems we relax these restrictions. However, for the sake of simplicity we shall restrict ourselves to sequences of random variables satisfying a strong mixing condition without introducing the retardation of Theorem 3.

For the next two theorems let $\{x_n\}$ be a sequence of random variables centered at expectations and with finite $(2 + \delta)$ moments where

$$(2.9) \quad 0 < \delta \leq 2 .$$

Suppose that

$$(2.10) \quad s_N^2 = E \left(\sum_{n \leq N} x_n \right)^2 \rightarrow \infty$$

as $N \rightarrow \infty$ and that

$$(2.11) \quad \max_{k \leq n} E |x_k|^{2+\delta} \ll s_n^{\rho\delta}$$

for some $0 \leq \rho \leq 1$. Moreover, suppose that $\{x_n\}$ satisfies a strong mixing condition with

$$(2.12) \quad \alpha(k) \ll k^{-300(1+2/\delta)} .$$

That is

$$|P(AB) - P(A)P(B)| \leq \alpha(n)$$

for all $A \in F_1^t$ and all $B \in F_{t+n}^\infty$. Here, as usual F_a^b , denotes the σ -field generated by $\{x_n, a \leq n \leq b\}$.

With this notation we have the following two theorems.

THEOREM 4: Let $\{x_n\}$ be a sequence of random variables satisfying (2.9), (2.10), and (2.12). Suppose that there is a constant C such that

$$E|x_n|^{2+\delta} \leq C \quad n = 1, 2, \dots$$

Let F be an arbitrary monotonically increasing function. Suppose that

$$(2.13) \quad \sum_{n=M+1}^{M+N} ||x_n||_{2+\delta} \leq F(A)$$

whenever

$$(2.14) \quad E\left(\sum_{n=M+1}^{M+N} x_n\right)^2 \leq A$$

for some A . In other words suppose that the left-hand side of (2.13) does not exceed F evaluated at the left-hand side of (2.14).

Define

$$S(t) = \sum_{k \leq N} x_k \quad \text{for } s_N^2 \leq t < s_{N+1}^2 .$$

Then, without changing the distribution of $\{S(t), t \geq 0\}$ we can redefine $\{S(t), t \geq 0\}$ on a richer probability space together with standard Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}-c\delta} \quad \text{a.s.}$$

for each $c < 1/588$.

THEOREM 5: Let $\{x_n\}$ be a sequence of random variables satisfying (2.9) - (2.12) with $\rho \leq \frac{1}{4}$. Let σ be a constant with

$$\rho\sigma \leq 1/10 .$$

Suppose that uniformly in $M = 1, 2, \dots$,

$$(2.15) \quad \sum_{n=M+1}^{M+N} \|x_n\|_{2+\delta} << \left(E \left(\sum_{n=M+1}^{M+N} x_n^2 \right)^2 \right)^\sigma$$

as $N \rightarrow \infty$. Then the conclusion of Theorem 4 remains valid for each $c < 1/1232$.

REMARKS: Conditions (2.13), (2.14), and (2.15) are of a similar structure as the Ljapounov condition for the central limit theorem for sequences of independent random variables.

Except for the rate of decay of $\alpha(k)$, Theorem 5 contains a result of Philipp (1969) as a special case. This result of Philipp's is, to the best of our knowledge, so far the only law of the iterated logarithm for

sequences of random variables satisfying a mixing condition where moments are not assumed to be uniformly bounded.

We can considerably relax the retarded mixing condition in the special case where $\eta_n = \xi_n$ in (2.5). In other words let $\{\xi_n\}_{n=1}^{\infty}$ be a sequence of random variables, centered at expectations and with uniformly bounded fifth moments. Suppose that

$$(2.16) \quad E \left(\sum_{n \leq N} \xi_n \right)^2 = N + o(N^{14/15}) .$$

Let κ be a positive constant. We shall assume that

$$(2.17) \quad |P(AB) - P(A)P(B)| \leq \alpha(n t^{-\kappa})$$

for all $A \in F_1^t$ and all $B \in F_{t+n}^{t+n}$. Here $\alpha(s)$ converges monotonically to zero. We call sequences $\{\xi_n\}$ satisfying (2.17) retarded weakly mixing. This condition is less restrictive than the concept of (*)-mixing, introduced by Blum, Hanson and Koopmans (1963) because of the retardation factor $t^{-\kappa}$ and because we do not require that the right-hand side of (2.17) contains the factor $P(A)P(B)$. Of course, (2.17) is also less restrictive than what we called in (2.6) a retarded strong mixing condition since we require less concerning the future of the process $\{\xi_n\}$.

Moreover, we assume that for all $i \leq j < m \leq n$

$$(2.18) \quad |E(\xi_i \xi_j \xi_m \xi_n) - E(\xi_i \xi_j)E(\xi_m \xi_n)| \leq \left\{ \alpha((m-j)j^{-\kappa}) \right\}^{1/5}.$$

Of course, if we would assume (2.6), instead of (2.17), this estimate would immediately follow from a well-known lemma of Ibragimov (1962).

THEOREM 6: Let $\{\xi_n\}$ be a sequence of random variables, centered at expectations, with uniformly bounded fifth moments, and satisfying (2.16), (2.17), and (2.18) with

$$\kappa = 2/19$$

and

$$\alpha(s) \ll s^{-168 \cdot 2}.$$

Then, without changing the distribution of $\{S(t), t \geq 0\}$, we can redefine the process $\{S(t), t \geq 0\}$ on a richer probability space together with

standard Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}-\eta} \quad \text{a.s.}$$

for each $\eta < 1/294$.

Let $\{x_n\}$ denote a Gaussian sequence centered at expectations. Theorems A - D are well-known for the case of stationary uncorrelated (that is, independent identically distributed) x_n . As a matter of fact, for this special case relation (1.2) is rather trivial to prove.

Here we consider Gaussian sequences whose n -th partial sums have variances close to n and that have covariances converging to zero as the distance between indices approaches infinity. In this more general setting the proof of (1.2) is far from being trivial.

We assume that uniformly in m

$$(2.19) \quad E\left(\sum_{k=m+1}^{m+n} x_k\right)^2 = \sigma^2 n + o(n^{1-\varepsilon})$$

for some $\varepsilon > 0$ and some constant $\sigma^2 > 0$, excluding the degenerate case $\sigma = 0$. Hence without loss of generality we assume

$$(2.20) \quad \sigma^2 = 1.$$

Moreover, we assume that uniformly in m

$$(2.21) \quad E(x_m x_{m+n}) = \rho(m, m+n) \ll n^{-2}.$$

Since by (2.19) the variances $E x_k^2$ are uniformly bounded, condition (2.21) is less restrictive than the requirement that the correlations converge to zero

at the given rate.

THEOREM 7: Without changing the distribution of $\{S(t), t \geq 0\}$, we can redefine the process $\{S(t), t \geq 0\}$ on a richer probability space together with standard Brownian motion $\{X(t), t \geq 0\}$ such that

$$S(t) - X(t) \ll t^{\frac{1}{2}-\eta} \quad \text{a.s.}$$

for each $\eta < \min(1/60, 4\epsilon/15)$.

COROLLARY: Let $\{x_n\}$ be a stationary Gaussian sequence, centered at expectations and with covariances

$$E(x_1 x_n) \ll n^{-2}.$$

The conclusion of Theorem 7 holds with

$$S(t) - X(t) \ll t^{\frac{1}{2}-\frac{1}{60}} \quad \text{a.s. .}$$

Similar results are obtained for functionals of certain Markov processes and for the Shannon-McMillan-Breimann theorem in information theory.

3. DESCRIPTION OF THE METHOD

Let $\{x_n\}_{n=1}^{\infty}$ be a weakly dependent sequence with $E x_n = 0$ and $E|x_n|^{2+\delta}$ uniformly bounded. Here $\delta > 0$ is fixed. For ease of explanation suppose that

$$(3.1) \quad \lim_{n \rightarrow \infty} \text{var}(n^{-\frac{1}{2}} S_n) = \sigma^2 > 0.$$

Recall that S_n was defined in (1.1) as the n -th partial sum. Without loss of generality we assume that $\sigma^2 = 1$. As already explained in Section 1 our goal is to establish the fundamental relation (1.2) for various weakly dependent sequences of random variables. The proof of (1.2) is accomplished in two main steps. In Section 3.1 we describe the first step and in Section 3.2 the second step.

3.1. The first step consists of approximating $\{S_n\}$ by a martingale. The following simple lemma concerning such approximations of sums of arbitrary random variables by martingales is useful.

LEMMA 1: Let $\{y_j\}_{j=1}^{\infty}$ be an arbitrary sequence of random variables and let $\{L_j\}_{j=0}^{\infty}$ be a nondecreasing sequence of σ -fields such that y_j is L_j -measurable. (Here L_0 denotes the trivial σ -field.) Suppose that the series

$$(3.2) \quad \sum_{k=0}^{\infty} \mathbb{E}|E(y_{j+k}|L_j)| < \infty \text{ a.s.}$$

for each $j \geq 1$. Then for each $j \geq 1$

$$y_j = Y_j + u_j - u_{j+1}$$

where $\{Y_j, L_j\}_{j=1}^{\infty}$ is a martingale difference sequence and

$$u_j = \sum_{k=0}^{\infty} E(y_{j+k}|L_{j-1}) .$$

Indeed, the desired martingale difference sequence is given by

$$Y_j = \sum_{k=0}^{\infty} (E(y_{j+k}|L_j) - E(y_{j+k}|L_{j-1})) , j \geq 1 ,$$

thus proving the lemma. \square

The idea for this lemma can be traced back at least to Statuljevičius (1969) and Gordin (1969). Both authors gave a martingale representation of certain strict sense stationary sequences, Gordin in terms of unitary operators on Hilbert space.

The significance of the lemma obviously lies in the fact that the partial sums of any sequence of random variables satisfying (3.2) equal a martingale plus a telescoping sum which under certain restrictions can be discarded. However, in most cases Lemma 1 turns out to be a rather weak tool when applied directly to the given sequence of random variables. It only then gains in strength when combined with another essential ingredient of the method which we describe now.

We define new random variables y_j which are sums of progressively larger blocks of the given x_n . Further, we sometimes have to construct sums z_j of blocks of the given x_n smaller than the corresponding blocks defining the y_j . As a typical example, consider the construction of the blocks required for the analysis of the Gaussian random variables of Theorem 6. There we define the y_j and the z_j inductively by adding $[j^{7/8}]$ and $[j^{6/8}]$ consecutive x_n respectively, leaving no gaps between the blocks. For example $y_1 = x_1$, $z_1 = x_2$, $y_2 = x_3$, $z_2 = x_4$, $y_3 = x_5 + x_6$, $z_3 = x_7 + x_8$, $y_4 = x_9 + x_{10} + x_{11}$, The z_j are defined in such a way that they provide enough separation between consecutive y_j 's so that (3.2) holds for each $j \geq 1$. For then, by Lemma 1,

$$y_j = Y_j + u_j - u_{j+1}$$

where $\{Y_j, L_j\}$ is a martingale difference sequence and L_j is the σ -field generated by Y_1, \dots, Y_j . Moreover, and essential to the subsequent analysis, the separation that the z_j are providing between consecutive y_j 's is so large that even

$$(3.3) \quad u_j = \sum_{k=0}^{\infty} E(y_{j+k} | L_{j-1})$$

is small compared to y_j for large j . For then

$$(3.4) \quad y_j \stackrel{\cdot}{\approx} Y_j$$

for large j . Lastly, since the blocks defining the z_j are much smaller than the blocks defining y_j , we have

$$(3.5) \quad \sum_{j=1}^M y_j \stackrel{\cdot}{\approx} \sum_{j=1}^M (y_j + z_j) \text{ a.s.}$$

for large M . In other words the z_j can be discarded without affecting the almost sure behaviour of the partial sums S_N . Of course, we are deliberately vague when saying that u_j is small compared with y_j , etc.

In the construction of the y_j and z_j there is a trade-off involved: If the blocks defining the z_j are too small, (3.2) and (3.4) will fail to hold. However, if these blocks are too large, then (3.5) will fail to hold. That is, the z_j 's could not be neglected in the subsequent analysis.

Assuming that the blocks have been chosen so that (3.2), (3.4), and (3.5) are satisfied, we then concentrate on the sums

$$\sum_{j \leq M} y_j$$

in the remainder of the discussion. However, consideration of these sums will produce fluctuation results only for a certain subsequence of the sequence $\{S_N\}$ of partial sums. Fortunately, it is relatively easy to break into the blocks defining the y_j 's and thus recover the desired fluctuation results for $\{S_N\}$, provided that the blocks defining the y_j 's are not chosen too large.

By the above arguments, noting (3.4), the proof of the fundamental relation (1.2) is reduced by the first step of the method to proving a corresponding almost sure invariance principle for the martingale

$$(3.6) \quad \sum_{j \leq M} Y_j .$$

3.2. The second step of the method is the approximation of the martingale (3.6) in an appropriate manner by Brownian motion. This is accomplished by means of a martingale version of the Skorokhod representation theorem (see Strassen (1965), Theorem 4.3). Application of this representation theorem provides (on a new probability space possibly) stopping times $\{T_j\}$ for Brownian motion $\{X(t)\}$ such that

$$(3.7) \quad \sum_{j \leq M} Y_j = X\left(\sum_{j \leq M} T_j\right) \text{ a.s. ,}$$

$$(3.8) \quad E(T_j | L_{j-1}) = E(Y_j^2 | L_{j-1}) \text{ a.s. ,}$$

and

$$(3.9) \quad E T_j^p \ll E |Y_j|^{2p}$$

for each $p > 1$.

Let M_N denote the index of the y_j or z_j which contains x_N and let $[t]$ denote, as usual, the greatest integer not exceeding t . In order to exploit (3.7) we establish a strong law of large numbers for the T_j in the form

$$(3.10) \quad \sum_{j=1}^{M_N} T_j = N + o(N^{1-\eta_1}) \quad \text{a.s.}$$

where $\eta_1 > 0$. For then we obtain for each $\eta_2 < \frac{1}{2}\eta_1$, omitting some calculations and using (3.10),

$$(3.11) \quad \begin{aligned} X\left[\sum_{j=1}^{M_N} T_j\right] &\doteq X\left[N + o\left(N^{1-\eta_1}\right)\right] \\ &\doteq X(N) + o\left[N^{\frac{1}{2}-\eta_2}\right] \\ &\doteq X(t) + o\left[t^{\frac{1}{2}-\eta_2}\right] \quad \text{a.s.} \end{aligned}$$

for $N \leq t < N + 1$. Thus if we set

$$S^*(t) = \sum_{j=1}^{M[t]} Y_j$$

we obtain, using (3.7) and (3.11),

$$(3.12) \quad S^*(t) - X(t) \ll t^{\frac{1}{2}-\eta_2} \quad \text{a.s.}$$

as $t \rightarrow \infty$.

Now we are almost done since, in view of (3.4), the reduction to the martingale case described in Section 3.1 results in the estimate

$$\begin{aligned}
 S(t) - S^*(t) &= \sum_{n \leq t} x_n - S^*(t) \\
 (3.13) \qquad &= \sum_1^{M[t]} (y_j - Y_j) \ll t^{\frac{1}{2} - \eta_3} \text{ a.s.}
 \end{aligned}$$

for some $\eta_3 > 0$. We combine (3.12) and (3.13) and get

$$S(t) - X(t) \ll t^{\frac{1}{2} - \eta} \text{ a.s.}$$

for some $\eta > 0$, which is the desired fundamental relation (1.2). This completes the second step of the method.

How can we get (3.10) which is, as we just have seen, the key to establishing the fundamental relation (1.2)? Consider the decomposition

$$\begin{aligned}
 \sum_{j=1}^{M_N} T_j - N &= \sum_{j=1}^{M_N} (T_j - E(T_j | L_{j-1})) \\
 (3.14) \qquad &+ \sum_{j=1}^{M_N} (E(Y_j^2 | L_{j-1}) - Y_j^2) \\
 &+ \sum_{j=1}^{M_N} Y_j^2 - N.
 \end{aligned}$$

Here we have used (3.8). We first establish a strong law of large numbers for the Y_j^2 in the form

$$(3.15) \qquad \sum_{j=1}^{M_N} Y_j^2 = N + o\left(N^{1-\eta_4}\right) \text{ a.s.}$$

for some $\eta_4 > 0$. This takes care of the third sum on the right-hand side of (3.14). As a by-product of the proof of (3.15) we obtain a bound on $E|Y_j|^{2+\delta}$ which, in view of (3.9), yields a bound on $E T_j^{1+\frac{1}{2}\delta}$. Using these two bounds

and a standard martingale convergence theorem it is not difficult to show that the first two sums on the right-hand side of (3.14) are $\ll N^{1-\eta_5}$ almost surely, proving (3.10). (In general, T_j is not measurable with respect to L_j . But this is only a technicality which we shall not worry about at this point.)

For general martingales, (3.15) does not hold. But using (3.14) it follows that it is enough to prove

$$(3.16) \quad \sum_{j=1}^{M_N} y_j^2 = N + o\left(N^{1-\eta_6}\right) \text{ a.s.}$$

for some $\eta_6 > 0$. In view of the weak dependence of the y_j 's, relation (3.16) might appear to be straight-forward. But typically, proving (3.16) provides the most difficulty.

REFERENCES

- [1] Berkes, I. (1974a). Almost sure invariance principles for mixing processes. Submitted for publication.
- [2] Berkes, I. (1974b). An almost sure invariance principle for lacunary trigonometric series. Submitted for publication.
- [3] Billingsley, Patrick (1967). Unpublished manuscript.
- [4] Billingsley, Patrick (1968). Convergence of Probability Measures. Wiley, New York.
- [5] Blum, J. R., Hansen, D. L., Koopmans, L.H. (1963). On the strong law of large numbers for a class of stochastic processes. *Z. Wahrscheinlichkeitstheorie*, 2, pp. 1 - 11.
- [6] Chow, Y. S. (1965). Local convergence of martingales and the law of large numbers. *Ann. Math. Statist.*, 36, pp. 552 - 558.
- [7] Chung, Kai Lai (1967). Markov Chains with Stationary Transition Probabilities. 2nd ed. Springer-Verlag. Berlin and New York.
- [8] Chung, Kai Lai (1968). A Course in Probability Theory. Harcourt, Brace, and World, New York.
- [9] Cohn, Harry (1965). On a class of dependent random variables. *Rev. Roum. Math. Pures et Appl.*, 10, pp. 1593 - 1606.
- [10] Davydov, Yu. A. (1970). The invariance principle for stationary processes. *Theor. Probability Appl.*, 15, pp. 487 - 493.
- [11] Deo, Chandrakant M. (1973). A note on empirical processes of strong-mixing sequences. *Ann. of Prob.*, 1, pp. 870 - 875.
- [12] Erdős, P. and Gál, I. S. (1955). On the law of the iterated logarithm. *Proc. Koninkl. Nederl. Akad. Wetensch. Ser. A*, 58, pp. 65 - 84.
- [13] Gaal, S. and Gaal, L., (1964). The discrepancy of the sequence $\{(2n_x)\}$. *Proc. Koninkl. Nederl. Akad. Wetensch. Ser. A*, 67, pp. 129 - 143.
- [14] Gál, I. S. and Koksma, J. F. (1950). Sur l'ordre de grandeur des fonctions sommables. *Proc. Koninkl. Nederl. Akad. Wetensch. Ser. A*, 53, pp. 638 - 653.
- [15] Gaposhkin, V. F. (1966). Lacunary series and independent functions, *Russian Math. Surveys*, 21, pp. 3 - 82.
- [16] Gordin, M. I. (1969). The central limit theorem for stationary processes. *Soviet Math. Doklady*, 10, pp. 1174 - 1176.
- [17] Heyde, C. C. and Scott, D. J. (1973). Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments. *Ann. of Prob.*, 1, pp. 428 - 437.

- [18] Ibragimov, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.*, 7, pp. 349 - 382.
- [19] Iosifescu, M. (1968). La loi du logarithme itéré pour une classe de variables aleatoires dependent. *Teorija Verroj.*, 13, pp. 315 - 325.
- [20] Jain, Naresh J., Jogdeo, Kumar, and Stout, William F. (1974). Upper and lower functions for martingales and mixing processes. *Ann. of Prob.*, 2, ***.
- [21] Jain, Naresh J. and Taylor, S. J. (1973). Local asymptotic laws for Brownian motion. *Ann. of Prob.*, 1, pp. 527 - 549.
- [22] Morrison, Donald F. (1967). Multivariate Statistical Methods. McGraw-Hill, New York.
- [23] Codaira, Hiroshi (1973). The law of the iterated logarithm for Gaussian processes. *Ann. Prob.*, 1, pp. 954 - 967.
- [24] Philipp, Walter (1969). The law of the iterated logarithm for mixing stochastic processes. *Ann. Math. Statist.*, 40, pp. 1985 - 1991.
- [25] Philipp, Walter (1971). Mixing Sequences of Random Variables and Probabilistic Number Theory. Amer. Math. Soc. Mem. No. 114, Amer. Math. Soc., Providence, Rhode Island.
- [26] Philipp, Walter and Stout, William F. (1974). Almost Sure Invariance Principles for Sums of Dependent Random Variables. Submitted for publication.
- [27] Reznik, M. Kh. (1968). The law of the iterated logarithm for some classes of stationary processes. *Theor. Probability Appl.*, 8, pp. 606 - 621.
- [28] Rosenblatt, M. (1956). A central limit theorem and a mixing condition. *Proc. Nat. Acad. USA*, 42, pp. 412 - 413.
- [29] Stout, W. F. (1974). Almost Sure Convergence. Academic Press, New York.
- [30] Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *A. Wahrscheinlichkeitstheorie*, 3, pp. 211 - 226.
- [31] Strassen, V. (1965). Almost sure behavior of sums of independent random variables and martingales. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, 2, pp. 315 - 343.
- [32] Szusz, P. (1963). Über die metrische Theorie der Diophanthischen Approximation. *Acta Arith.*, 8, pp. 225 - 241.
- [33] Takahashi, S. (1972). On the law of the iterated logarithm for lacunary trigonometric series, *Tohoku Math. J.*, 24, pp. 319 - 329.
- [34] Vervaat, W. (1972). Success epochs in Bernoulli trials. Mathematisch centrum - Amsterdam.