The Annals of Probability
1973, Vol. 1, No. 2, 272-296

SOME STRASSEN-TYPE LAWS OF THE ITERATED LOGARITHM FOR MULTIPARAMETER STOCHASTIC PROCESSES WITH INDEPENDENT INCREMENTS[1]

BY MICHAEL J. WICHURA

University of Chicago

For processes with independent increments and multi-dimensional time parameters, analogues of Strassen's version of the law of the iterated logarithm are proven using standard large-deviation and truncation techniques. Applications to empirical processes are included.

1. Introduction. In 1964 Strassen introduced the following striking form of the law of the iterated logarithm. Let $X = (X(t))_{t \geqq 0}$ be a standard Brownian motion process. For each $n \geqq 3$, let H_n be the random function defined on $[0, 1]$ by setting $H_n(s) = X(ns)/(2n \log \log n)^{\frac{1}{2}}$ for each s. Then with probability one, the sequence (H_n) is relatively compact in the uniform topology on $C[0, 1]$, and its limit points coincide with the class of absolutely continuous functions which vanish at 0 and whose derivatives lie in the unit ball of $L_2([0, 1])$ under Lebesgue measure. Using the Skorokhod imbedding theorem, Strassen also deduced the corresponding law for the partial sums of independent identically distributed random variables with zero means and finite variances.

In this paper we discuss Strassen-type laws of the iterated logarithm for processes with independent increments and multi-dimensional time parameters. Since the Skorokhod imbedding theorem is not available in this context (cf. Pyke (1972) Section 8), we develop all our results by making use of large-deviation techniques due to Kolmogorov (1929), and truncation techniques due to Hartman and Wintner (1941) and Feller (1968). In conjunction with an imbedding theorem involving Poisson processes, our theorems yield a Strassen-type law of the iterated logarithm for empirical processes. This law in turn has a whole family of corollaries, including Finkelstein's (1971) result.

Our main results are stated in Section 2. Some spade work for the proofs is done in Section 3; the proofs themselves are given in Sections 4, 5, and 6.

We conclude this section by establishing some notation and terminology. Let q be a positive integer, and put $T = [0, \infty)^q$. A point t in T is written explicitly as $(t^{(1)}, \cdots, t^{(q)})$, or simply $(t^{(p)})$; $\mathbf{c}$ denotes the point t such that $t^{(p)} = c$,

Received March 27, 1972; revised August 22, 1972.

[1] This research was carried out in the Department of Statistics, University of Chicago, under partial sponsorship of the Statistics Branch, Office of Naval Research, Grant No. N00014-67-A-0285-0009 and Research Grant No. GP 32037x from the Division of Mathematical Sciences of the National Science Foundation.

AMS 1970 *subject classifications.* Primary 60F15, 60J30; Secondary 60G17, 60F10.

Key words and phrases. Law of the iterated logarithm, processes with independent increments, multi-parameter stochastic processes, empirical processes.

$1 \leqq p \leqq q$. For points t, u in T, $|t|$ denotes $\prod_{1 \leqq p \leqq q} t^{(p)}$, and tu denotes the point $(t^{(1)}u^{(1)}, \cdots, t^{(q)}u^{(q)})$. An *interval* in T is a rectangle, all of whose sides are parallel to the coordinate axes. $[a, b]$ denotes the interval $\prod_{1 \leqq p \leqq q} [a^{(p)}, b^{(p)}]$; a is its *lower left corner*, b its *upper right corner*.

Each t in T determines 2^q *quadrants*

$$Q_R(t) = \{u \in T : t^{(p)} R_p u^{(p)},\ 1 \leqq p \leqq q\}$$

as R varies over the 2^q q-tuples $(R_1, \cdots, R_q)$, in which each R_p is one of the relations $\leqq$ and $>$. $D(T)$ denotes the space of real-valued functions x on T such that

(1.1) $\quad x_Q(t) \equiv \lim_{u \to t, u \in Q} x(u)$ exists for each of the 2^q quadrants $Q = Q_R(t)$

(1.2) $$x(t) = x_{Q(\leqq, \cdots, \leqq)}(t)$$

for each t in T, and

(1.3) $\quad x(t) = 0$ whenever any coordinate of t equals 0.

The *increment*, $x(A)$, of x in $D(T)$ over a subinterval A of T is the qth difference of x around the corner points of A (e.g., for $q = 1$ and $A = [a, b]$, $x(A) = x(b) - x(a)$). Similarly, the *jump*, $J_t(x)$ of x in $D(T)$ at the point t in T is the qth difference of the 2^q limits $x_{Q_R(t)}$; in other words, $J_t(x)$ is the limit of $x(A)$ as the diameter of A tends to 0, with A varying among intervals having t as an interior point.

U_q denotes the *unit cube* $[\mathbf{0}, \mathbf{1}]$. $L_q = \{t \in U_q : \min_{1 \leqq p \leqq q} t^{(p)} = 0\}$ is the *lower boundary* of U_q. Suppose that ν is a measure on the Borel subsets of U_q and that x and y are two real valued functions on U_q, such that y is integrable with respect to ν. x is said to be the *indefinite integral* of y, and y the *derivative* of x, with respect to ν, written

$$x = \mathscr{I}_\nu(y), \qquad \text{or equivalently,} \quad y = (\dot{x})_\nu,$$

if

$$x(t) = \int_{[\mathbf{0}, t]} y(s)\nu(ds) \qquad \text{for all } t \text{ in } U_q.$$

When ν is Lebesgue measure, the dependence on ν is suppressed and we write, e.g.

(1.4) $$x = \mathscr{I}(y), \qquad y = \dot{x}.$$

$L_2(U_q, \nu)$ denotes the set of y's which are square-integrable with respect to ν; $L_2(U_q)$ is $L_2(U_q, \nu)$ when ν is Lebesgue measure. $\mathscr{D}(U_q)$ denotes the space of functions x on U_q which satisfy (1.1) and (1.2); it is to be noted that $x \in \mathscr{D}(U_q)$ is not required to satisfy (1.3).

T is *partially ordered* by stipulating that $s \leqq t$ if and only if $s^{(p)} \leqq t^{(p)}$ for each p. This partial ordering applies also to N_q, the set of points in T all of whose coordinates are positive integers.

The indicator function of a set A is written I_A. λ denotes *Lebesgue measure*, on whatever domain is pertinent at the time. For $c \geqq 0$, $\log c$ and $\log_2 c = \log(\log c)$ have their usual meanings except near 0; we take $\log c$ (resp. $\log_2 c$) to be 1 over $[0, e)$ (resp. over $[0, e^e)$).

2. Statement of results. THROUGHOUT THIS SECTION, $X = (X(t))_{t \in T}$ DENOTES A PROCESS, DEFINED ON SOME PROBABILITY SPACE $(\Omega, \mathscr{B}, P)$, SUCH THAT (i) THE INCREMENTS OF X OVER DISJOINT INTERVALS ARE INDEPENDENT, (ii) EACH $X(t)$ HAS ZERO MEAN AND FINITE VARIANCE, AND (iii) ALL SAMPLE PATHS OF X LIE IN $D(T)$.

We are interested here in the almost sure sample path behavior of X as t "grows large." Let t_n, $n \geqq 1$, be an increasing sequence of points in T. To keep track of the "history" of X up to time t_n, put $\sigma_n^2 = \mathrm{Var}\,(X(t_n))$ and define a random function H_n on U_q by

$$H_n(u) = X(ut_n)/(2\sigma_n^2 \log_2 \sigma_n^2)^{\frac{1}{2}} . \tag{2.1}$$

Let

$$K = \{\mathscr{I}(y) : \textstyle\int_{U_q} y^2(t)\, dt \leqq 1\}$$

be the class of functions on U_q which are indefinite integrals (cf. (1.4)) of functions in the unit ball of $L_2(U_q)$ under Lebesgue measure. Let d denote the uniform metric for functions on U_q; thus $d(x, y) = \sup_{u \in U_q} |x(u) - y(u)| \equiv ||x - y||$. The following is the simplest of several theorems, each of which asserts that (under appropriate conditions) almost all sample points ω in Ω have the property that the functions $H_n(\cdot)(\omega)$, $n \geqq 1$, eventually resemble *only* functions in K, and repeatedly resemble *each* x in K:

THEOREM 1. *Let H_n be defined by* (2.1). *If X has stationary increments, and if*

$$\sigma_n^2 \to \infty \qquad \textit{and} \qquad \sigma_{n+1}^2/\sigma_n^2 \to 1 , \tag{2.2}$$

then

$$P(\{\limsup_n d(H_n, K) = 0\}) = 1 \tag{2.3}$$

$$P(\textstyle\bigcap_{x \in K} \{\liminf_n d(H_n, x) = 0\}) = 1 . \tag{2.4}$$

This result contains that of Pyke (1972) for the standard Brownian motion process on T; it applies equally well, e.g., to a mean-centered homogeneous Poisson process. The requirement in (2.2) that $\sigma_n^2 = \mathrm{Var}\,(X(t_n))$ not tend to infinity too fast is not needed for (2.3), but is for (2.4) (for example, when X is the standard Brownian motion process in univariate time ($q = 1$), $d(H_n, 0) \to 0$ wp 1 for $t_n = \sigma_n^2 = e^{e^n}$, $n \geqq 1$). In view of the compactness of K (see Section 3), (2.3) and (2.4) are equivalent to the statement that wp 1 the sequence (H_n) is relatively compact and has K as its set of limit points. Also, (2.3) implies that the H_n's are uniformly bounded and asymptotically equicontinuous; it follows that the discrete sequence (t_n) can be replaced by any increasing family $(t_\theta)_{\theta > 0}$ of points in T for which $\mathrm{Var}\,(X(t_\theta))$ tends continuously to ∞ as $\theta \uparrow \infty$ (cf. Lemma 4.4 below).

In order to state analogues of Theorem 1 in the case of non-stationary increments, we shall have to make the structure of the process X more explicit. In doing so, we will call upon the following lemma, which can be proved along the lines of Doob (1953) page 415:

LEMMA 2.1. *Let S be a countable subset of T such that* $\min_p s^{(p)} > 0$ *for each s in S. Let $(\xi_s)_{s\in S}$ be a family of independent random variables having zero means and finite variance; suppose moreover that*

$$\sum_{s\in B\cap S} \mathrm{Var}\,(\xi_s) < \infty$$

for every bounded subset B of T. Then there exists an independent-increment process $\Xi = (\Xi(t))_{t\in T}$, defined on the same probability space as the ξ_s's, such that

(a) $P(\{\Xi(t) = \sum_{n=1}^{\infty} \xi_{s_n}\}) = 1$, *for each t in T and each enumeration $(s_n)_{n\geq 1}$ of the points of S in* $[\mathbf{0}, t]$,
(b) *the sample paths of Ξ lie in $D(T)$*,
(c) $J_t(\Xi) = 0$ *for all $t \notin S$, and* $P(\{J_t(\Xi) = \xi_t\}) = 1$ *for all $t \in S$,*
(d) $E(\Xi(t)) = 0$ *and* $\mathrm{Var}\,(\Xi(t)) < \infty$ *for all t in T.*

We will refer to the process Ξ in Lemma 2.1 as the *partial sum process* formed from the ξ_s's; it is, of course, only unique up to an almost sure equivalence. Our assumptions about the sample path and moment properties of X imply that for each t in T, the limits $X_{Q_{R(t)}}$ (cf. (1.1)) can be taken in the sense of $L_2(P)$ convergence (cf. Doob (1953) page 108), and thus that the jump $J_t(X)$ has zero mean and finite variance. Let

$$\Delta = \{t \in T\colon \mathrm{Var}\,(J_t(x)) > 0\} \tag{2.5}$$

be the "fixed discontinuity" set of X. The $J_t(X)$'s with t in Δ are independent random variables and

$$\sum_{t\in B\cap\Delta} \mathrm{Var}\,(J_t(X)) < \infty \tag{2.6}$$

for any bounded subset B of T. Thus we can use Lemma 2.1 to form a partial sum process Y from the variables $J_t(X)$, $t \in \Delta$. We will call Y the *partial sum component* of X. Now put $Z = X - Y$. Z is a process with independent increments which

(a) is independent of Y,
(b) has sample paths in $D(T)$, and for which
(c) $J_t(Z) = 0$ wp 1,
(d) $E(Z(t)) = 0$ and $\mathrm{Var}\,(Z(t)) < \infty$

for all t in T. It follows (e.g., by using the techniques of Gikhman and Skorokhod (1969) Chapter 6) that all the increments of Z have infinitely divisible distributions, so we may call Z the *infinitely divisible component* of X. Let Π be the Lévy measure associated with Z, defined on the Borel sets of $T \times R^1$ by

$$\Pi(A) = E(\textstyle\sum_{t\in T} I_A((t, J_t(Z)))) ,$$

and for measurable $B \subset T$ let Π_B be the measure defined on the Borel sets of R^1 by

(2.7) $$\Pi_B(C) = \Pi(B \times C) .$$

$\Pi_B(C)$ is just the expected number of jumps of Z which occur at times in B and which lie in C. Write Π_t for $\Pi_{[0,t]}$. The log characteristic function of $Z(t)$ can be written in the form

(2.8) $$\log E(e^{i\zeta Z(t)}) = -\zeta^2 v_t^2/2 + \int_{\zeta \neq 0} (e^{i\zeta\xi} - 1 - i\zeta\xi)\Pi_t(d\xi)$$

where v_t^2 is the variance of the Gaussian component of $Z(t)$. The variance of $Z(t)$ itself is $v_t^2 + \int_{\xi \neq 0} \xi^2 \Pi_t(d\xi)$.

We are now going to give extensions of Theorem 1 under conditions similar to those used by Kolmogorov (1929), and by Hartman and Wintner (1941). In each case, we will require the variance structure of X to be homogeneous in the limit, in the sense that

(2.9) $$\lim_{n\to\infty} \operatorname{Var}(X(t_n A))/\sigma_n^2 = \lambda(A)$$

for each interval $A \subset U_q$. To illustrate this condition, suppose that $q = 1$ and that X is constructed in the usual way from the partial sums S_n of independent random variables by setting $X_t = S_n$ for t in $[\operatorname{Var}(S_n), \operatorname{Var}(S_{n+1}))$, $n \geqq 0$ $(S_0 = 0)$. Then (2.9) is satisfied if we take $t_n = \operatorname{Var}(S_n)$ and assume that (2.2) holds.

For the "Kolmogorov" version of Theorem 1, we impose bounds on the magnitude of the jumps of X. Write $||J_t(X)||_\infty$ for the usual essential supremum of $J_t(X)$, and $||\Pi_t||_\infty$ for $\inf\{b > 0 : \Pi_t\{\xi : |\xi| \geqq b\} = 0\}$. Recall that Δ is defined by (2.5).

THEOREM 2. *Let H_n be defined by* (2.1). *Suppose that* (2.2) *and* (2.9) *hold, and that*

(2.10) $$\sup_{t \leqq t_n, t \in \Delta} ||J_t(X)||_\infty/\sigma_n = o(1/(\log_2 \sigma_n^2)^{\frac{1}{2}})$$

(2.11) $$\sup_{t \leqq t_n} ||\Pi_t||_\infty/\sigma_n = ||\Pi_{t_n}||_\infty/\sigma_n = o(1/(\log_2 \sigma_n^2)^{\frac{1}{2}})$$

as $n \to \infty$. Then (2.3) *and* (2.4) *hold.*

If X is a pure partial sum process, then Theorem 2 gives us a Strassen-type law of the iterated logarithm corresponding to Kolmogorov's classical result. If X is purely Gaussian, then (2.10) is vacuous (Δ is empty) and (2.11) is trivially satisfied ($\Pi_t = 0$ for all t in T), so we get an extension of Pyke's result. The hypotheses of Theorem 2 can be modified so as to make applicable various convergence rate results for the central limit theorem (cf. Chover (1966) and Tompkins (1971b)).

Now put $F_1 = [\mathbf{0}, t_1]$ and set $F_n = [\mathbf{0}, t_n] - [\mathbf{0}, t_{n-1}]$ for $n \geqq 1$; note that $\sum_{m \leqq n} F_m = [\mathbf{0}, t_n]$. Let Π_{F_n} be defined by (2.7). A "Hartman–Wintner" type truncation scheme and Theorem 2 lead to the following result, in which constraints are imposed on the large jumps of X:

THEOREM 3. *Let H_n be defined by* (2.1). *Suppose that* (2.2) *and* (2.9) *hold, and that there exists a positive number δ and a positive measure Λ on $[\delta, \infty]$ such that*

$$\sum_{t \in F_n \cap \Delta} P(\{|J_t(X)| \geqq c\}) \leqq \operatorname{Var}(X(F_n))\Lambda([c, \infty)) \tag{2.12}$$

$$\Pi_{F_n}\{\xi : |\xi| \geqq c\} \leqq \operatorname{Var}(X(F_n))\Lambda([c, \infty)) \tag{2.13}$$

for all $c \geqq \delta$ and all $n \geqq 1$, and

$$\int \xi^2\Lambda(d\xi) = \delta^2\Lambda([\delta, \infty)) + 2\int_\delta^\infty \xi\Lambda([\xi, \infty))\, d\xi < \infty . \tag{2.14}$$

Then (2.3) *and* (2.4) *hold.*

When X has stationary increments, (2.12) is vacuous, and (2.13) and (2.14) hold for any $\delta > 0$ with Λ defined by

$$\Lambda([c, \infty)) = (\Pi_{U_q}([c, \infty)) + \Pi_{U_q}((-\infty, -c]))/\operatorname{Var}(X(U_q))$$

($c \geqq \delta$); (2.14) is satisfied because of our assumption of finite variances. Thus Theorem 1 is a special case of Theorem 3. In view of (2.14), condition (2.12) is only slightly stronger than the bound resulting from Chebychev's inequality. When the partial sum component Y of X has only finitely many jump points (t's in Δ) in each F_n, condition (2.12) can be replaced by the pair

$$\liminf_n \sigma_n^2/N_n > 0 \tag{2.15}$$

$$P(\{|J_t(X)| \geqq c\}) \leqq \Lambda([c, \infty)) \qquad \text{for all} \quad c \geqq \delta \quad \text{and all} \quad t \in \Delta\,; \tag{2.16}$$

here N_n denotes the number of points t in $[\mathbf{0}, t_n] \cap \Delta$. Conditions (2.2) (2.9), (2.14), (2.15), and (2.16) are essentially those used by Hartman and Wintner for their classical law. The last three of these conditions are satisfied, e.g., if Δ is a lattice and if the $J_t(X)$'s for $t \in \Delta$ are independent and identically distributed (with zero means and finite variance).

The above results yield limit theorems for various functionals of the H_n's. Most of these are based on the following lemma, which is analogous to the so-called mapping theorem in weak-convergence (cf. Billingsley (1968) page 34):

LEMMA 2.2. *Let $(\Omega, \mathscr{B}, P)$ be a probability space, let S and S' be metric spaces, and let C be a compact subset of S. Let $(\mathscr{H}_n)$ be a sequence of mappings from Ω into S such that for P-almost all ω in Ω, the sequence $(\mathscr{H}_n(\omega))$ is relatively compact in S and has C as its set of limit points. Let (Φ_n) be a sequence of functions from $\Omega \times S$ into S', and let Φ map S into S'. Let G be the set of sample points ω in Ω such that*

$$\Phi_n(\omega, x_n) \to \Phi(x) \tag{2.17}$$

whenever $x_n \in$ range $(\mathscr{H}_n(\omega))$, $x \in C$, and $x_n \to x$. Suppose that $G \in \mathscr{B}$ and $P(G) = 1$. Then for P-almost all ω, the sequence $(\Phi_n(\omega, \mathscr{H}_n(\omega)))$ is relatively compact in S' and its limit points coincide with the compact set $C' \equiv \Phi(C)$; moreover, if $(\Phi_{n'}(\omega, \mathscr{H}_{n'}(\omega)))$ is a subsequence which converges to a point y in C', then the distance between $\mathscr{H}_{n'}(\omega)$ and the set $\{x \in C: \Phi(x) = y\}$ tends to zero.

The statement of Lemma 2.2 simplifies somewhat if the Φ_n's are nonrandom, i.e., do not depend on ω. Here are some consequences of Theorems 1, 2, and 3, each of which follows from Lemma 2.2. In each case we suppose that the conclusions (2.3) and (2.4) in Theorem 1 are known to hold.

I. Wp 1,

$$\limsup_n H_n(\mathbf{1}) = 1 ;$$

moreover, for large n, $H_n(\mathbf{1})$ is near 1 iff H_n is near the function

$$u \to |u| \qquad (u \in U_q) .$$

II. Wp 1,

$$\limsup_n \int |(H_n(u)|\, du = (1/3^{\frac{1}{2}})^q \qquad \text{and} \qquad \limsup_n (\int H_n(u)^2\, du)^{\frac{1}{2}} = (2/\pi)^q ;$$

moreover, for large n, $||H_n||_1$ is near $(1/3^{\frac{1}{2}})^q$ iff $|H_n|$ is near the function

$$u \to \prod_{1 \leqq p \leqq q} (\tfrac{3}{4})^{\frac{1}{2}}(1 - (1 - u^{(p)})^2)$$

and $||H_n||_2$ is near $(2/\pi)^q$ iff $|H_n|$ is near the function

$$u \to \prod_{1 \leqq p \leqq q} (8/\pi)^{\frac{1}{2}} \sin((\pi/2)u^{(p)}) .$$

In general, for any $a \geqq 1$, the almost sure $\limsup_n$ of $||H_n||_a$ is the qth power of the corresponding value in the univariate case (Strassen (1964) page 219) and the nonnegative function in K that delivers this value is a "product" function, each of whose components is the corresponding function in the univariate case (Strassen (1964) page 221).

III. Let μ_n, $n \geqq 1$, and μ be finite signed measures on U_q. Suppose that μ_n converges weakly to μ, in the sense that $\mu_n(f) \to \mu(f)$ for all continuous real-valued functions f on U_q; this is the case if, e.g., one has (i) $\mu_n([\mathbf{0}, \mu]) \to \mu([\mathbf{0}, u])$ for all continuity points $u \in U_q$ of μ, and (ii) $\sup_n ||\mu_n||_\infty < \infty$ (here $||\ ||_\infty$ denotes total variation). Condition (ii) is also necessary (cf. Royden (1963) page 171 and 256). Wp 1, we have

$$\limsup_n \int H_n\, d\mu_n = (\int M(u)^2\, du)^{\frac{1}{2}} = ||M||_2 \tag{2.18}$$

where M is the function defined on U_q by

$$M(u) = \mu([u, \mathbf{1}]) ;$$

moreover $\int H_n\, d\mu_n$ is near $||M||_2$ iff H_n is near the indefinite integral of $M/||M||_2$. These conclusions even hold for random μ_n's which converge weakly to μ wp 1.

This result has obvious applications to Riemann-integral type functionals of the H_n's. Another class of limit theorems arises when the sample paths of the H_n's are of bounded variation (see Gikhman and Skorokhod (1969) for necessary and sufficient conditions relative to the infinitely divisible component of X). In this case we can define a (random) measure η_n on the Borel sets of U_q by stipulating that $\eta_n([\mathbf{0}, u]) = H_n(u)$ for all $u \in U_q$. Define M_n on U_q by $M_n(u) = \mu_n([u, \mathbf{1}])$. Fubini's theorem implies that $\int H_n\, d\mu_n = \int M_n\, d\eta_n$, and so (2.18) becomes

$$\limsup_n \int M_n\, d\eta_n = ||M||_2 . \tag{2.19}$$

For example, suppose that $q = 2$ and that $\xi_{j,k}$, $j, k \geqq 1$, are (independent) random variables such that the process X with

$$X(s, t) = \sum_{j\leqq s, k\leqq t} \xi_{j,k}$$

satisfies the hypotheses of one of the above theorems, with t_n taken to be the point $(n, n) \in T$, for $n \geqq 1$. Then for any function M of bounded variation on U_q, (2.19) (with $M_n = M$ for all n) implies

$$\limsup_n \sum_{j,k\leqq n} M(j/n, k/n)\xi_{j,k}/(2\sigma_n^2 \log_2 \sigma_n^2)^{\frac{1}{2}} = ||M||_2 \tag{2.20}$$

wp 1. The univariate version ($q = 1$) of (2.20) extends a series of results due to Gál (1951), Stackelberg (1964), Gaposhkin (1965), and Tompkins (1971a). It is easy to go on to give a Strassen-type version of (2.20).

The above results have all dealt with the sample path behavior of the process X as the time parameter t grows large along a fixed sequence. It is natural to ask what happens if t simply grows large, without any restrictions. The two theorems below deal with the situation in which t tends to infinity in the sense that each coordinate of t tends to infinity. In this connection, we shall write $\lim_{t\to\infty}$ for $\lim_{t(p)\to\infty; 1\leqq p\leqq q}$. Consider the random process, H_t, $t \in T$, defined on U_q by

$$H_t(u) = X(ut)/(2q\sigma_t^2 \log_2 \sigma_t^2)^{\frac{1}{2}} \tag{2.21}$$

($\sigma_t^2 = \mathrm{Var}\,(X(t))$). Comparing (2.21) with (2.1), one sees that t_n in (2.1) has been replaced by t, the manner of indexing H has been changed, and, most important, a factor of $q^{\frac{1}{2}}$ has been introduced into the normalizing constants. The following theorems show that under conditions somewhat more stringent than those used above, the H_t's are wp 1 relatively compact as $t \to \infty$ and have K as their limit points.

Here is the analogue of Theorem 2:

THEOREM 4. *Let H_t be defined by* (2.21). *Suppose that*

$$\lim_{t\to\infty} \sigma_t^2/|t| \quad \textit{exists and is finite and positive} \tag{2.22}$$

$$\lim_{t\to\infty} ((\log_2 |t|)/|t|)^{\frac{1}{2}} \sup_{s\leqq t; s\in\Delta} ||J_s(X)||_\infty = 0 \tag{2.23}$$

$$\lim_{t\to\infty} ((\log_2 |t|)/|t|)^{\frac{1}{2}} ||\Pi_t||_\infty = 0\,. \tag{2.24}$$

Then

$$P(\{\lim_{t\to\infty} \sup_{u\geqq t} d(H_u, K) = 0\}) = 1 = P(\textstyle\bigcap_{x\in K} \{\lim_{t\to\infty} \inf_{u\geqq t} d(H_u, x) = 0\})\,. \tag{2.25}$$

The conditions of Theorem 4 are met when X has stationary imcrements; thus for standard Brownian motion on T, Theorem 4 gives a Strassen-type version of Zimmerman's (1972) law of the iterated logarithm. Condition (2.22) is a somewhat more stringent version of (2.9); we point out though that the process used to illustrate (2.9) statisfies (2.22) as well. It is important to note that Theorem 4 would no longer be valid were the limits on t to be taken simply as

$|t| \to \infty$, instead of as each coordinate of t tends to ∞. To see this, take $q = 2$ and let X be Brownian motion. If (t_n) is a sequence of points in T such that $|t_n| \to \infty$ sufficiently slowly while the corresponding first coordinates $t_n^{(1)} \to \infty$ sufficiently rapidly, then $\limsup H_{t_n}(\mathbf{1})/(|t_n| \log_2 |t_n|)^{\frac{1}{2}} = \infty$ wp 1, in view of the near independence of the standard normal random variables $X(t_n)/(|t_n|)^{\frac{1}{2}}$, $n \geqq 1$. This example suggests, and it is indeed the case, that Theorem 4 remains in force when all the passages to the limit on t are taken as $|t| \to \infty$ while $\min_p |t^{(p)}|$ remains bounded away from 0. Under this kind of limiting procedure, Orey and Pruitt (1973) have established the strong (i.e., integral-test) form of the law of the iterated logarithm for Brownian motion on T (actually, Orey and Pruitt deal with the path behavior near $\mathbf{0}$; a time inversion gives the corresponding result near ∞).

For each k in N_q, put $C_k = \{t \in T: k^{(p)} - 1 \leqq t^{(p)} \leqq k^{(p)}; 1 \leqq p \leqq q\}$. Theorem 4 and a modification of Feller's (1968) truncation scheme lead to

THEOREM 5. *Let $q \geqq 2$, and H_t be defined by* (2.21). *Suppose that* (2.22) *holds and that there exists a number $\delta > 0$ and a positive measure Λ on $[\delta, \infty)$ such that*

$$\max\left(\textstyle\sum_{t \in C_k \cap \Delta} P(\{|J_t(X)| \geqq c\}),\quad \Pi_{C_k}\{\xi : |\xi| \geqq c\}\right) \leqq \Lambda([c, \infty)) \tag{2.26}$$

for all $c \geqq \delta$ and all k in N_q, and

$$\textstyle\int_\delta^\infty (\xi^2 \log^{q-1} \xi)/\log_2 \xi \; \Lambda(d\xi) < \infty\,. \tag{2.27}$$

Then (2.25) *holds.*

To put this result in perspective, let V_k, $k \in N_q$, be a family of i.i.d. random variables, and define a process $\mathscr{V}$ on N_q by setting

$$\mathscr{V}(k) = \textstyle\sum_{j \leq k} V_j\,.$$

When $q \geqq 2$, Theorem 5 and Lemma 5.1 below imply that

$$P(\{\limsup_k |\mathscr{V}(k)|/(|k| \log_2 |k|)^{\frac{1}{2}} < \infty\}) = 1 \tag{2.28}$$

holds if and only if

$$E(V_k) = 0 \qquad \text{and} \qquad E((V_k^2 \log^{q-1} |V_k|)/\log_2 |V_k|) < \infty\,. \tag{2.29}$$

However, when $q = 1$, (2.28) is known to be equivalent to

$$E(V_k) = 0 \qquad \text{and} \qquad E(V_k^2) < \infty \tag{2.30}$$

(cf. Strassen (1966) or Feller (1968)). A glance at (2.29) and (2.30) shows why one needs $q \geqq 2$ in the statement of Theorem 5. The difference between the cases $q = 1$ and $q \geqq 2$ arises primarily because it is in precisely the latter case that one can deduce the finiteness of $E(V_k^2)$ from (2.29). By way of comparison, it is of interest to note that there is no such discontinuity with regard to q in the strong law of large numbers; Smythe (1973) has shown that $\lim_k \mathscr{V}(k)/|k| = 0$ wp 1 if and only if $E(V_k) = 0$ and $E(|V_k| \log^{q-1} |V_k|) < \infty$.

Theorems 4 and 5 imply almost sure limit results for various functionals of the H_t's. The general form of these derived results is given by the "net" version of Lemma 2.2. In particular, applications I, II, and III above carry over with obvious modifications to the present setting.

We turn now to a consideration of the law of the iterated logarithm for empirical distribution functions. Let μ be an arbitrary probability measure on U_q, and let $V_1, V_2, \cdots$ be independent U_q-valued random vectors, each with distribution μ. Define processes Y_n on U_q by

$$Y_n(u) = I_{[0,u]}(V_n)\,, \tag{2.31}$$

and set

$$S_n = \textstyle\sum_{m \leq n} Y_m\,.$$

Define H_n^μ on $U_1 \times U_q$ by

$$\begin{aligned} H_n^\mu(s, u) &= (S_{[ns]}(u) - ES_{[ns]}(u))/(2n \log_2 n)^{\frac{1}{2}} \\ &= [ns](F_{[ns]}(u) - F(u))/(2n \log_2 n)^{\frac{1}{2}}\,, \end{aligned} \tag{2.32}$$

where $F_k = (1/k)S_k$ is the sample empirical distribution function based on $V_1, \cdots, V_k$, and F is the distribution function of μ. H_n^μ takes values in $\mathscr{D}(U_1 \times U_q)$. Let

$$\nu = \lambda \times \mu \tag{2.33}$$

be the product of Lebesgue measure on U_1 and μ on U_q. Let K^μ denote the set of functions x in $\mathscr{D}(U_1 \times U_q)$ for which there exists an $\dot{x}$ in $L_2(U_1 \times U_q, \nu)$ such that

$$x = \mathscr{I}_\nu(\dot{x}) \tag{2.34}$$

$$\textstyle\int \dot{x}^2\, d\nu \leq 1 \tag{2.35}$$

$$x(s, 1) = 0 \qquad \text{for all} \quad s \quad \text{in} \quad U_1\,. \tag{2.36}$$

Then we have

THEOREM 6. *Let H_n^μ be defined by* (2.32). Wp 1 *the sequence* (H_n^μ) *is relatively compact* (*with respect to the uniform metric d*) *and its limit points coincide with K^μ.*

Theorem 6 is not a direct consequence of our earlier results, because the Y_n's, though independent and identically distributed, are not real-valued. To prove Theorem 6 we will imbed (S_n) in a suitable Poisson process on $[0, \infty) \times U_q$, make use of (a variant of) Theorem 2, and then use Lemma 2.2. We note that

$$H_n(1, \cdot) = n(F_n - F)/(2n \log_2 n)^{\frac{1}{2}}\,,$$

so Theorem 6 (with $q = 1$ and F continuous) stands in the same relation to Finkelstein's (1971) law as Strassen's (1964) law does to the classical law of the iterated logarithm. For some related results, see Kiefer (1971) and Müller (1970). Here are some applications of Theorem 6.

IV. Let $D_n = \sup_{u \in U_q} |F_n(u) - F(u)|$ be the Kolmogorov–Smirnov distance between F_n and F. Put

$$(2.37)\qquad c(\mu) = \sup_{u\in U_q} [F(u)(1 - F(u))]^{\frac{1}{2}} = \sup_{A\in\mathscr{R}} [\mu(A)(1 - \mu(A))]^{\frac{1}{2}},$$

where $\mathscr{R}$ denotes the class of all intervals in U_q which contain the point $\mathbf{0}$; if F is continuous, $c(\mu) = \frac{1}{2}$. Let $\mathscr{R}_0$ be the class of A's in $\mathscr{R}$ which realize the supremum in (2.37), and let $\mathscr{S}(\mu)$ be the class of functions x in $\mathscr{D}(U_1 \times U_q)$ of the form

$$(2.38)\qquad x(s, u) = s\psi(u)$$

where ψ is the indefinite integral, relative to μ, of the function $(I_A - \mu(A))/c(\mu)$, for some A in $\mathscr{R}_0$. When $q = 1$ and μ is the uniform distribution on U_1, there is just one x in $\mathscr{S}(\mu)$; the corresponding ψ is linear on $[0, \frac{1}{2}]$ and on $[\frac{1}{2}, 1]$, with $\psi(0) = 0$, $\psi(\frac{1}{2}) = \frac{1}{2}$, and $\psi(1) = 0$. In general, we have

$$(2.39)\qquad \limsup_n nD_n/(2n \log_2 n)^{\frac{1}{2}} = c(\mu)$$

wp 1. For $q = 1$ and F continuous, (2.39) goes back to Smirnov (1944) and Chung (1949); Pyke (1971) conjectured that (2.39) holds for arbitrary F. Kiefer (1961) has (2.39) for $q \geqq 1$ and F continuous. Our result gives the added information that, when n is large, $nD_n/(2n \log_2 n)^{\frac{1}{2}}$ is close to $c(\mu)$ iff H_n^μ or $-H_n^\mu$ is close to the set $\mathscr{S}(\mu)$.

We may strengthen (2.39) as follows. Let $\mathscr{R}^*$ be the class of all intervals in U_q. Then wp 1

$$(2.40)\qquad \limsup_n \sup_{A\in\mathscr{R}^*} (n|F_n(A) - \mu(A)|/(2n \log_2 n)^{\frac{1}{2}} - [\mu(A)(1 - \mu(A))]^{\frac{1}{2}}) = 0 .$$

This extends a series of results due to Cassels (1951), Phillip (1969), and Zaremba (1971); see also Phillip (1971).

V. Our next application of Theorem 6 has as a corollary a Strassen-type version of (2.39). Let Γ be a (semi-) norm on $\mathscr{D}(U_q)$ which is majorized by the uniform norm $\|\ \|$. For example, Γ may be the uniform norm itself, or the L_a norm $x \to (\int |x(u)|^a \mu(du))^{1/a}$ $(a \geqq 1)$. Define processes H_n^Γ on U_1 by

$$(2.41)\qquad H_n^\Gamma(s) = \Gamma(H_n^\mu(s, \cdot)) .$$

Put $c_\Gamma(\mu) = \sup\{\Gamma(w) : w \in W_\mu\}$, where W_μ consists of the functions w in $\mathscr{D}(U_q)$ for which there exists a $\dot{w}$ in $L_2(\mu)$ such that

$$w = \mathscr{I}_\mu(\dot{w}) , \qquad \int_{U_q} \dot{w}^2\, d\mu \doteq 1 , \qquad \text{and} \qquad w(\mathbf{1}) = 0 .$$

Then wp 1, (H_n^Γ) is relatively compact with respect to the uniform metric and its limit points coincide with the set

$$(2.42)\qquad K_\Gamma^\mu = \{z \in \mathscr{D}(U_1) : z \geqq 0 \text{ and } z = \mathscr{I}(\dot{z}) \text{ for some function } \dot{z} \text{ satisfying } \textstyle\int \dot{z}^2(\xi)\, d\xi \leqq c_\Gamma^2(\mu)\} .$$

We note that when $\Gamma = \|\ \|$, $c_\Gamma(\mu) = c(\mu)$, defined by (2.37), and

$$H_n^\Gamma(s) = [ns]D_{[ns]}/(2n \log_2 n)^{\frac{1}{2}}$$

($s \in U_1$). Most of the applications which Strassen (1964) gives for his theorem carry over with trivial changes to the present context. For example, for each θ in [0, 1], the proportion of m's $\leqq n$ for which

$$mD_m/(2m \log_2 m)^{\frac{1}{2}} \geqq \theta c(\mu)$$

has an almost sure lim sup on n equal to $1 - \exp(-4(\theta^{-2} - 1))$.

3. Preliminaries. We begin by discussing some properties of the set K. We shall make use of the following terminology. For each $p \leqq q$, let $\mathscr{A}^{(p)}$ be a finite partition of [0, 1] into left-closed, right-open intervals (adopt the convention of always closing up the interval containing 1), and let $\mathscr{A} = \{\prod_p A_p : A_p \in \mathscr{A}^{(p)}, 1 \leqq p \leqq q\}$ be the corresponding "rectangular" partition of U_q. Let $x \in \mathscr{D}(U_q)$ be a function which vanishes along the lower boundary L_q of U_q. Define $x_{\mathscr{A}}$ to be the indefinite integral (cf. (1.4)) of the function $\sum_{A \in \mathscr{A}} (x(A)/\lambda(A)) I_A$. For each A in $\mathscr{A}$, $x_{\mathscr{A}}$ and x coincide on the corner points of A, and at each inner point of A, $x_{\mathscr{A}}$ is an average of the values of x at the corner points.

Now suppose that x in K is the indefinite integral of $\dot{x}$. From the Cauchy–Schwarz inequality

$$|\textstyle\int_B \dot{x}(u)\, du| \leqq [\lambda(B)]^{\frac{1}{2}}(\int_B \dot{x}^2(u)\, du)^{\frac{1}{2}} \tag{3.1}$$

holding for Borel subsets B of U_q, it follows that

$$||x|| \equiv \sup_u |x(u)| \leqq 1, \tag{3.2}$$

$$\omega_\delta(x) \equiv \sup\{|x(u) - x(t)| : |u^{(p)} - t^{(p)}| \leqq \delta, 1 \leqq p \leqq q\} \leqq (q\delta)^{\frac{1}{2}} \tag{3.3}$$

for $\delta > 0$, and

$$||\dot{x}_{\mathscr{A}}||_2^2 \equiv \textstyle\int \dot{x}_{\mathscr{A}}^2(u)\, du = \sum_{A \in \mathscr{A}} x(A)^2/\lambda(A) \leqq \int \dot{x}^2(u)\, du \leqq 1, \tag{3.4}$$

for each rectangular partition $\mathscr{A}$ of U_q. Relation (3.4) is crucial in the proof of the "lower class" result (2.4).

Next suppose $x \in \mathscr{D}(U_q)$ vanishes on L_q. Let $\mathscr{A}$ be a rectangular partition of U_q. Since the function $x_{\mathscr{A}}/||\dot{x}_A||_2$ is in K, (3.2) implies

$$d(x, K) \leqq d(x, x_{\mathscr{A}}) + d(x_{\mathscr{A}}, K) \leqq 2 \max_{A \in \mathscr{A}} w_A(x) + (||\dot{x}_{\mathscr{A}}||_2 - 1)^+; \tag{3.5}$$

here

$$w_A(x) = \sup_{u \in \bar{A}} |x(u) - x({}_A u)|$$

(${}_A u$ denotes the lower-left corner of A, and $\bar{A}$ the closure of A). Relation (3.5) is the basis for the proof of the "upper class" result (2.3).

Finally, we observe that the converse to (3.4) holds; if $x \in \mathscr{D}(U_q)$ vanishes on L_q and if $x_{\mathscr{A}}$ is in K for all rectangular partitions $\mathscr{A}$, then x is in K. This may be proved by showing that x is the indefinite integral of L_2-limit (cf. Doob (1953) page 319) of the martingale $(\dot{x}_{\mathscr{A}_n})_{n \geqq 1}$, where $\mathscr{A}_n$ is the partition of U_q obtained by taking each $\mathscr{A}_n^{(p)}$ to be $\{[(i-1)2^{-n}, i2^{-n}) : 1 \leqq i \leqq 2^n\}$. It follows that K is closed and therefore compact by (3.2), (3.3), and the Arzela–Ascoli theorem.

The following fluctuation inequality is taken from Wichura (1969):

LEMMA 3.1. *Let $n \in N_q$, and let $(W_m)_{m \leq n}$ be a q-dimensional array of independent random variables with 0 means and finite variances. Put $S_m = \sum_{k \leq m} W_k$, and set $M_n = \max_{m \leq n} |S_m|$. Then*

$$(3.6) \qquad P(\{M_n \geqq 2^q a\}) \leqq (1 - (\sigma/a)^2)^{-q} P(\{|S_n| \geqq a\})$$

for all a such that $a^2 \geqq \sigma^2 \equiv \mathrm{Var}\,(S_n)$.

In the sequel, we will make use of the obvious generalization of (3.6) to (separable) processes with a continuous (as opposed to discrete) time index. For some related inequalities see Kiefer (1961), Cabaña (1972), Pyke (1971), and Orey and Pruitt (1973).

The following large deviation result will be sufficient for our purposes. For related, and more extensive, results see, e.g., Feller (1943) and (1969), Statulevicius (1966), and Bahadur (1971).

LEMMA 3.2. *Let (a_n) be a sequence of positive constants tending to ∞. Let (S_n) be a sequence of random variables whose cumulant generating functions* (cgf's)

$$C_n : \zeta \to \log\,(E(\exp\,(\zeta S_n)))$$

($\zeta \in R^1$) satisfy

$$(3.7) \qquad C_n''(\zeta) = 1 + o(1)$$

uniformly for $|\zeta| \leqq 2a_n$, as $n \to \infty$. Then

$$(3.8) \qquad P(\{S_n \geqq a_n\}) = \exp[-(\tfrac{1}{2})a_n^2(1 + o(1))]$$

as $n \to \infty$.

PROOF. Let C_{Q_n} be the cgf of the distribution, Q_n, of $S_n - a_n$. Let τ_n be the root of the equation $C'_{Q_n}(\zeta) = 0$ and put $\rho_n = -C_{Q_n}(\tau_n)$. Set $\gamma_n^2 = C''_{Q_n}(0)$ and let Q_n^* be the distribution whose cgf is given by $C_{Q_n^*}(\zeta) = C_{Q_n}(\zeta/\gamma_n + \tau_n) - C_{Q_n}(\tau_n)$. According to the argument on page 5 of Bahadur (1971),

$$(3.9) \qquad \log\,(P(\{S_n \geqq a_n\})) = -(1 + o(1))\rho_n$$

provided $\gamma_n \tau_n = o(\rho_n)$ and no subsequence of (Q_n^*) converges weakly to the distribution degenerate at 0. Using (3.7), it is easily checked that $\sigma_n \sim a_n$, $\rho_n \sim a_n^2/2$, $\gamma_n^2 \to 1$, and $C''_{Q_n^*}(\zeta) \to 1$ for all ζ, so Q_n^* converges weakly to the standard normal distribution. Consequently, (3.8) follows from (3.9). □

We will make use of Lemma 3.2 to prove

LEMMA 3.3. *Suppose that the hypotheses of Theorem 2 hold. Let $(A)_{A \in \mathscr{A}}$ be a finite collection of disjoint subintervals of U_q, and let $(\theta_A)_{A \in \mathscr{A}}$ be a unit vector (so $\sum_A \theta_A^2 = 1$). Then for any $\beta > 0$,*

$$(3.10) \qquad P(\{\textstyle\sum_{A \in \mathscr{A}} \theta_A X(t_n A)/(\sigma_n(\lambda(A))^{\frac{1}{2}}) \geqq (2\beta \log_2 \sigma_n^2)^{\frac{1}{2}}\}) = (\log \sigma_n^2)^{-(1+o(1))\beta}\,.$$

PROOF. Write $X = Y + Z$ as the sum of its partial-sum and infinitely-divisible components, as in Section 2. Let $A \in \mathscr{A}$, and let $C_{\mathscr{Y}_n}$ and $C_{\mathscr{X}_n}$ be the cgf's of

$$\mathscr{Y}_n = Y(t_n A)/(\sigma_n(\lambda(A))^{\frac{1}{2}}) \qquad \text{and} \qquad \mathscr{X}_n = Z(t_n A)/(\sigma_n(\lambda(A))^{\frac{1}{2}})$$

respectively. In view of Lemma 3.2, it suffices to show

$$C''_{\mathscr{Y}_n}(\zeta) + C''_{\mathscr{X}_n}(\zeta) = 1 + o(1) \tag{3.11}$$

uniformly for ζ's which are $O((\log_2 \sigma_n^2)^{\frac{1}{2}})$.

The cgf of $Z(t_n A)$ is (cf. (2.8)) given by

$$C_{Z(t_n A)}(\zeta) = v(t_n A)^2\zeta^2/2 + \int_{\xi\neq 0} (e^{\zeta\xi} - 1 - \zeta\xi)\Pi_{t_n A}(d\xi) ,$$

where $\Pi_{t_n A}$ is defined by (2.7), and $v.^2$ is the variance of the Gaussian component of Z. Thus

$$C''_{\mathscr{X}_n}(\zeta) = (\sigma_n^2\lambda(A))^{-1}(v(t_n A)^2 + \int_{\xi\neq 0} \xi^2 \exp(\zeta\xi/\sigma_n(\lambda(A))^{\frac{1}{2}}))\Pi_{t_n A}(d\xi)) .$$

By (2.11), $\Pi_{t_n A}$ gives measure 0 to ξ's of magnitude exceeding $o(1)(\sigma_n/(\log_2 \sigma_n^2)^{\frac{1}{2}})$, so we have

$$C''_{\mathscr{X}_n}(\zeta) = (\sigma_n^2\lambda(A))^{-1}\operatorname{Var}(Z(t_n A))(1 + o(1)) \tag{3.12}$$

as $n \to \infty$, uniformly for $|\zeta| \leqq b(\log_2 \sigma_n^2)^{\frac{1}{2}}$ (b fixed, but arbitrary).

Now consider $C_{\mathscr{Y}_n}$. Put $\Delta_n = \Delta \cap t_n A$ and set $U_{t;n} = J_t(X)/(\sigma_n(\lambda(A))^{\frac{1}{2}})$ for t in Δ_n. By (a) of Lemma 2.1, $\mathscr{Y}_n = \sum_{m=1}^{\infty} U_{t_m;n}$ wp 1 for any fixed enumeration $(t_m)_{m\geqq 1}$ of the points of Δ_n. Each $U_{t;n}$ has mean 0, and by (2.10) we have

$$B_n \equiv \sup_{t\in\Delta_n} ||U_{t;n}||_\infty = o(1/(\log_2 \sigma_n^2)^{\frac{1}{2}})$$

as $n \to \infty$. Let $M_{t;n}$ be the moment generating function of $U_{t;n}$. By expanding $M_{t;n}$, $M'_{t;n}$, and $M''_{t;n}$ in power series and dominating terms of the form $E|U_{t;n}|^j$ by $\operatorname{Var}(U_{t;n})B_n^{j-2}$ for $j \geqq 3$, one easily finds that

$$C''_{\mathscr{Y}_n}(\zeta) = (\sigma_n^2\lambda(A))^{-1}\operatorname{Var}(Y(t_n A))(1 + o(1)) \tag{3.13}$$

uniformly for $|\zeta| \leqq b(\log_2 \sigma_n^2)^{\frac{1}{2}}$. Adding (3.12) and (3.13), and using (2.9) we get (3.11). □

4. Proof of Theorems 2 and 4. (a) Proof of Theorem 2. We first prove the upper class statement (2.3), along the following lines. Consider (3.5), with x replaced by H_n; we would like to guarantee that with probability 1 the right-hand side is small for all large n. Along any "geometric" subsequence, the first Borel Cantelli lemma allows us to deduce that the second term is eventually small for any $\mathscr{A}$ (Lemma 4.1) and that the first term is eventually small provided $\mathscr{A}$ is fine enough (Lemma 4.2). It then remains only to note (Lemma 4.4) that there is enough continuity in the mapping $n \to H_n$ to deduce (2.3) from the corresponding statement for geometric subsequences.

Proceeding to the details, let c be a positive number, (slightly) greater than 1. Using (2.2), find indices n_k such that $\sigma_{n_k}^2 \equiv \operatorname{Var}(X(t_{n_k})) \sim c^k$ as $k \to \infty$. For each integer $m \geqq 1$, let $\mathscr{A}_m$ be the rectangular partition of U_q with each $\mathscr{A}_m^{(p)} = \{[(i-1)/m, i/m) \colon 1 \leqq i \leqq m\}$, and put $H_{n;m} = (H_n)_{\mathscr{A}_m}$.

LEMMA 4.1. *For each m*, $\limsup_k ||\dot{H}_{n_k;m}||_2 \leqq 1$ wp 1.

PROOF. It suffices to show that

$$\limsup_k \sum_{A\in\mathscr{A}_m} \theta_A X(t_{n_k}A)/((\sigma_{n_k}(\lambda(A))^{\frac12})(2\log_2\sigma_{n_k}^2)^{\frac12}) \leqq 1 \tag{4.1}$$

wp 1 for each unit vector $(\theta_A)_{A\in\mathscr{A}_m}$. But (4.1) follows easily from Lemma 3.3 and from the first Borel Cantelli lemma. □

LEMMA 4.2. Wp 1, $\limsup_m \limsup_k \max_{A\in\mathscr{A}_m} w_A(H_{n_k}) = 0$.

PROOF. Let $m \geqq 1$, $A \in \mathscr{A}_m$. Let u^A (resp. ${}_Au$) be the upper-right (resp. lower-left) corner point of A, and put $L_A = [\mathbf{0}, u^A] - [\mathbf{0}, {}_Au]$. By (2.9) we have $\sigma_n^{-2}\operatorname{Var}(X(t_n u^A) - X(t_n {}_Au)) \to \lambda(L_A)$; note that

$$\lambda(L_A) \leqq q/m\ . \tag{4.2}$$

Lemmas 3.1 and 3.3 imply that for each $\varepsilon > 0$,

$$P(\{\max_{A\in\mathscr{A}_m} w_A(H_n) \geqq 2^q\varepsilon\}) \leqq 4m^q(\log\sigma_n^2)^{-m\varepsilon^2/2q}$$

holds for all large n. Lemma 4.2 now follows from the first Borel Cantelli lemma. □

LEMMA 4.3. Wp 1, $\limsup_k d(H_{n_k}, K) = 0$.

PROOF. This follows from (3.5) and Lemmas 4.1 and 4.2. □

LEMMA 4.4. Wp 1, $\limsup_k \max_{n_{k-1}<n\leqq n_k} d(H_n, H_{n_k}) \leqq (c^{\frac12} - 1) + (q(c - 1/c))^{\frac12}$.

PROOF. Let $n_{k-1} < n \leqq n_k$. Put $\beta_n = ((\sigma_{n_k}^2 \log_2\sigma_{n_k}^2)/(\sigma_n^2\log_2\sigma_n^2))^{\frac12}$ and let α_n be the point in U_q whose pth coordinate is $\alpha^{(p)} = t_n^{(p)}/t_{n_k}^{(p)}$, $1 \leqq p \leqq q$. Then

$$H_n(u) = \beta_n H_{n_k}(\alpha_n u)$$

for all u in U_q. It follows from (2.9) that $\min_p \alpha_n^{(p)} \geqq c^{-1} - (c-1)$ provided k is sufficiently large. Lemma 4.4 now follows from Lemma 4.3 and inequalities (3.2) and (3.3). □

LEMMA 4.5. Wp 1, $\limsup_n d(H_n, K) = 0$.

PROOF. This follows by combining Lemmas 4.3 and 4.4, and letting $c \downarrow 1$. □

We turn now to the proof of the lower class statement (2.4). Since

$$\bigcap_{x\in K}\{\liminf_n d(H_n, x) = 0\} = \bigcap_{j\geqq 1}\{\liminf_n d(H_n, x_j) = 0\}$$

for any countable dense sequence of points x_j of K, it suffices to show that

$$\liminf_n d(H_n, x) = 0 \tag{4.3}$$

wp 1 for any x in K satisfying

$$\|\dot{x}\|_2 < 1\ . \tag{4.4}$$

Roughly speaking, the plan is to get $d((H_n)_{\mathscr{A}}, x_{\mathscr{A}})$ frequently small along suitable geometric subsequences using the second Borel Cantelli lemma (Lemma 4.6), and then to use equicontinuity considerations (here we need the partition $\mathscr{A}$ to be fine) to deduce (4.3) (Lemma 4.7). The program has to be modified a little so as to make the second Borel Cantelli lemma applicable.

Suppose then that x in K satisfies (4.4). Let c be a (large) positive integer, and, using (2.2), find indices n_k such that $\sigma^2_{n_k} \sim (c^q)^k$ as $k \to \infty$. Put

(4.5) $$B_c = \bigcup_p U_{p-1} \times [0, 2/c) \times U_{q-p} \subset U_q ;$$

condition (2.9) implies that the point

(4.6) $$(t^{(p)}_{n_{k-1}}/t^{(p)}_{n_k})_{1 \leq p \leq q} \in B_c$$

for all large k. Let $\mathscr{A}_c$ be the rectangular partition of U_q obtained by taking each $\mathscr{A}_c^{(p)} = \{[(i-1)/c, i/c) : 1 \leqq i \leqq c\}$. Put $\mathscr{A}_c^* = \{A \in \mathscr{A}_c : A$ does not intersect $B_c\}$.

Lemma 4.6. *For each* $\delta > 0$, $\liminf_k \max_{A \in \mathscr{A}_c^*} |H_{n_k}(A) - x(A)| \leqq \delta$ wp 1.

Proof. Fix δ, and put

$$G_k = \bigcap_{A \in \mathscr{A}_c^*} \{|H_{n_k}(A) - x(A)| \leqq \delta\} .$$

By (4.6), the G_k's (with k large) are mutually independent events, and so the lemma will follow provided

(4.7) $$\sum_{k=1}^\infty P(G_k) = \sum_{k=1}^\infty \prod_{A \in \mathscr{A}_c^*} P(\{|H_{n_k}(A) - x(A)| \leqq \delta\}) = \infty .$$

Using Chebychev's inequality to handle the terms in the product for which $|x(A)| < \delta$, and Lemma 3.3 to handle the other terms, one finds that, for any given $\varepsilon > 0$,

$$P(G_k) \geqq 2^{-1}(\log \sigma^2_{n_k})^{-\sum_A (x(A)^2/\lambda(A))(1+\varepsilon)}$$

for all large k. In view of (4.4) and (3.4), it follows (upon choosing ε sufficiently small) that (4.7) holds. □

Lemma 4.7. *For each* $\delta > 0$, $\liminf_k d(H_{n_k}, x) \leqq 2(q/c)^{\frac{1}{2}} + c^q\delta + 2(q(2/c))^{\frac{1}{2}}$ wp 1.

Proof. Let $u \in U_q$ and let t be the lower-left corner point of the element of $\mathscr{A}_c$ containing u. Then $H_{n_k}(u)$ can be written in the form

$$H_{n_k}(u) = (H_{n_k}(u) - H_{n_k}(t)) + \sum_{A \in \mathscr{A}_c^*, A \subset [0,t]} H_{n_k}(A) + \int_F \dot{H}_{n_k;c}(s)\, ds$$

where $F = F(u)$ is a sum of elements of $\mathscr{A}_c$, each of which lies in B_c. A similar decomposition holds for x, and so the desired bound follows from (3.1), (3.3), and Lemmas 4.1, 4.5, and 4.6. □

Lemma 4.8. Wp 1, $\liminf_n d(H_n, x) = 0$.

Proof. This follows from Lemma 4.7. □

In view of (4.3), Lemmas 4.5 and 4.8 yield Theorem 2.

(b) Proof of Theorem 4. The proof of Theorem 4 is similar to that of Theorem 2. Lemma 3.2 is actually valid when stated for arbitrary nets, and so under the conditions of Theorem 4, the conclusion (3.10) of Lemma 3.3 is valid with

t_n replaced by t, and σ_n^2 by σ_t^2. The role of the points t_{n_k}, $k \geqq 1$, in the proof in part (a) above is taken over by the "exponential grid" $(t_k)_{k \in N_q}$, where now t_k is defined to be the point $(c^{k_1}, \cdots, c^{k_q})$. In the analogues of Lemmas 4.1, 4.2, and 4.6 one uses the fact that for each j in N_q

$$\sum_{k \geqq j, k \in N_q} (k_1 + \cdots + k_q)^{-\beta q}$$

converges if $\beta > 1$, and diverges if $\beta < 1$.

5. Proof of Theorems 3 and 5. (a) Proof of Theorem 3. We begin by introducing a Hartman–Wintner type truncation scheme which will put us in the domain of applicability of Theorem 2. Let $v \to \theta(v)$ and $v \to \varepsilon(v)$ be any two positive functions of v in $(0, \infty)$, related by the identity

$$\theta(v)/v^{\frac{1}{2}} = \varepsilon(v)/(\log_2 v)^{\frac{1}{2}}, \tag{5.1}$$

and satisfying $\varepsilon(v) \downarrow 0$ and $\theta(v) \uparrow \infty$ as $v \uparrow \infty$.

Write X as the sum of its partial sum component Y and infinitely-divisible component Z, as in Section 2. Let Z^* be the process obtained from Z by first removing, for each $n \geqq 1$, all jumps $J_t(Z)$ which occur at (random) time points t in F_n and which exceed $\theta(\sigma_n^2)$ in magnitude, and by then centering to zero means. Then (cf. Gikhman and Skorokhod (1969) Chapter 6) Z^* is a $D(T)$-valued process with independent increments, whose Lévy measure Π^* satisfies (by construction) $\Pi^*_{F_n}\{\xi : |\xi| \geqq \theta(\sigma_n^2)\} = 0$ for all $n \geqq 1$. Moreover, $Z^*(t)$ has mean 0 and finite variance for each t.

For each t in $\Delta \cap F_n$ put $J_t^*(X) = J_t'(X) - E(J_t'(X))$, where $J_t'(X) = J_t(X) I_{\{|J_t(X)| \leqq \theta(\sigma_n^2)\}}$. Then for all t in Δ, $\mathrm{Var}\,(J_t^*(X)) \leqq \mathrm{Var}\,(J_t(X))$, and so in view of (2.6) we can use Lemma 2.1 to construct a partial sum process Y^* out of the $J_t^*(X)$'s ($t \in \Delta$).

Put $X^* = Y^* + Z^*$; then Y^* and Z^* are respectively the partial-sum and infinitely-divisible components of X^* (in particular, $J_t(X^*) = J_t^*(X)$ wp 1, for each t in Δ). Let H_n^* be defined in terms of X^* in the same way H_n was obtained from X:

$$H_n^*(u) = X^*(t_n u)/(2\sigma_n^{*2} \log_2 \sigma_n^{*2})^{\frac{1}{2}}$$

($u \in U_q$) with $\sigma_n^{*2} = \mathrm{Var}\,(X^*(t_n))$. Theorem 3 is proved by using Theorem 2 to show that (2.3) and (2.4) hold H_n replaced by H_n^*, and by then showing that the function ε appearing in (5.1) can be chosen to decrease to zero so slowly that $d(H_n, H_n^*) \to 0$ wp 1. The argument for this is similar to Stout's (1970) version of the Hartman–Wintner proof, and will be omitted.

(b) Proof of Theorem 5. One consequence of the "divergence" half of the following preliminary lemma is that (2.28) implies (2.29):

LEMMA 5.1. *Let L be a random variable, and set $l = E((L^2 \log^{q-1} |L|)/\log_2 |L|)$. If $l < \infty$, then*

$$\sum_{k \in N_q} P(\{|L| \geqq \zeta(|k| \log_2 |k|)^{\frac{1}{2}}\}) < \infty$$

for all $\zeta > 0$. *Conversely, if* $l = \infty$, *then*

$$\sum_{k \in N_q, k \geqq j} P(\{|L| \geqq \zeta(|k| \log_2 |k|)^{\frac{1}{2}}\}) = \infty$$

for all $\zeta > 0$ *and all* j *in* N_q.

PROOF. For any $c \geqq 1$, the Lebesgue measure of $\{t \in T: |t| \leqq c, t \geqq \mathbf{1}\}$ is easily shown to be

$$\sum_{j=1}^{q} (-1)^{j-1} (c \log^{q-j} c)/(q-j)! + (-1)^q$$

using Fubini's theorem and induction on q. This and standard methods give the lemma. □

We now turn to the proof of the main result. To alleviate the notational burden, we shall assume that X is a pure partial sum process; the general case, in which the infinitely-divisible component of X does not vanish, is handled by the technique used in the proof of Theorem 3 in part (a) above. We rescale X so that (2.22) becomes

$$\lim \sigma_t^2/|t| = 1 \,. \tag{5.2}$$

Following Feller (1968), we introduce several truncation levels in order to get into the range of applicability of the "Kolmogorov" type Theorem 4. Put

$$\alpha(s) = \delta s^{\frac{1}{2}} (\log_2 s)^{-1}$$
$$\beta(s) = \delta s^{\frac{1}{2}} (\log_2 s)^{\frac{3}{8}}$$
$$\gamma(s) = \delta s^{\frac{1}{2}} (\log_2 s)^{\frac{1}{2}} \,.$$

Note that for each k in N_q, $\delta \leqq \alpha(|k|) \leqq \beta(|k|) \leqq \gamma(|k|)$. For simplicity, write J_t for $J_t(X)$. For each k in N_q and each t in $\Delta \cap C_k$, set

$$J_t' = J_t I_{\{|J_t| \leqq \alpha(|k|)\}} \,, \qquad J_t^* = J_t' - E(J_t')$$
$$J_t'' = J_t I_{\{\alpha(|k|) < |J_t| \leqq \beta(|k|)\}} \,, \qquad J_t^{**} = J_t'' - E(J_t'')$$
$$J_t''' = J_t I_{\{\beta(|k|) < |J_t|\}} \,, \qquad J_t^{***} = J_t''' - E(J_t''') \,.$$

Use Lemma 2.1 to construct, on the same probability space on which X is defined, partial sum processes X^* (resp. X^{**}) corresponding to the J_t^*'s (resp. J_t^{**}'s) for $t \in \Delta$. Define X^{***} by means of the equation $X = X^* + X^{**} + X^{***}$; X^{***} then serves as the partial sum process associated with the J_t^{***}'s, $t \in \Delta$. For each $t \in T$, define H_t^* on U_q by

$$H_t^*(u) = X^*(ut)/(2q|t| \log_2 |t|)^{\frac{1}{2}} \,.$$

To prove Theorem 5, it suffices to show that wp 1,

(5.3) the net $(H_t^*)_{t \in T}$ is relatively compact as $t \to \infty$ and has K as its set of limit points,

$$\lim_{t \to \infty} \sup_{s \leqq t} |X^{**}(s)|/\gamma(|t|) = 0 \,, \quad \text{and} \tag{5.4}$$

$$\lim_{t \to \infty} \sup_{s \leqq t} |X^{***}(s)|/\gamma(|t|) = 0 \,. \tag{5.5}$$

LEMMA 5.2. (5.3) *holds.*

PROOF. For each k in N_q

$$|\mathrm{Var}\,(X(C_k)) - \mathrm{Var}\,(X^*(C_k))| \leqq 2\sum_{t\in C_k\cap\Delta}\int_{\{|J_t|\geqq\alpha(|k|)\}} J_t^2 \leqq 2\int_{\alpha(|k|)}^{\infty}\xi^2\Lambda(d\xi)$$

by (2.26). (5.3) follows easily from (2.27), (5.2), and Theorem 4. □

For the next three lemmas, we make use of the following notation. Choose and fix a moderately large integer θ. For each k in N_q, set

$$t_k = (\theta^{k_1}, \cdots, \theta^{k_q}), \qquad \tau_k = (\theta^{k_1-1}, \cdots, \theta^{k_q-1}).$$

Partition T into intervals B_k, $k \in N_q$, such that, for each k, t_k is the upper right corner of B_k.

LEMMA 5.3. (5.4) *holds.*

PROOF. For each k in N_q, put

$$M_k = \sup_A |X^{**}(A)|,$$

where the supremum is taken over all subintervals A of B_k having the same lower left corner point as B_k. Then for all t in B_k,

$$|X^{**}(t)| \leqq \sum_{j\leqq k} M_j.$$

We will show below that for any given $\varepsilon > 0$,

$$P(\{M_k \geqq \varepsilon\gamma(|t_k|) \text{ for infinitely many } k \text{ in } N_q\}) = 0. \tag{5.6}$$

Given this, there exists an almost surely finite random variable G_ε such that for all k in N_q and all $t \leqq t_k$,

$$|X^{**}(t)| \leqq \varepsilon\sum_{j\leqq k}\gamma(|t_j|) + G_\varepsilon. \tag{5.7}$$

As ε is arbitrary, (5.4) follows easily.

Using the first Borel Cantelli lemma, Lemma 3.1, and the 4th moment form of Chebychev's inequality, argue (as in Lemma 5.1 of Feller (1968)) that (5.6) holds provided

$$\sum_{k\in N_q}(\beta^2(|t_k|)\gamma^{-4}(|t_k|)\sum_{t\in B_k\cap\Delta}\int_{\{\alpha(|\tau_k|)\leqq|J_t|\leqq\beta(|t_k|)\}} J_t^2)$$

is finite. That this sum in fact converges is a consequence of the following lemma (which is similar to Lemma 5.2 of Feller (1968)). □

LEMMA 5.4. *Let $g < h$ be two numbers in $[-1, \frac{1}{2}]$, and put $a(s) = \delta s^{\frac{1}{2}}(\log_2 s)^g$, $b(s) = \delta s^{\frac{1}{2}}(\log_2 s)^h$. Let $\rho > 1$. Then* (2.26) *and* (2.27) *imply*

$$\sum_{k\in N_q}((\log_2|t_k|)^{-\rho}|t_k|^{-1}\sum_{t\in B_k\cap\Delta}\int_{\{a(|\tau_k|)\leqq|J_t|\leqq b(|t_k|)\}} J_t^2) < \infty. \tag{5.8}$$

PROOF. We make use of the estimate

$$\int_{\{a\leqq|J_t|\leqq b\}} J_t^2 \leqq a^2P(\{|J_t| \geqq a\}) + 2\int_a^b \xi P(\{|J_t| \geqq \xi\})\,d\xi.$$

In view of this and (2.26), (5.8) will hold provided

$$\int(\sum_{m:a(\theta^m)\leqq\xi} m^{q-1}\theta^m(\log_2\theta^m)^{2g-\rho})\Lambda(d\xi) < \infty \tag{5.9}$$

$$\int(\sum_m I_{\mathscr{I}_m}(\xi))(\xi\log^{q-1}\xi)/(\log_2\xi)^\rho\Lambda([\xi, \infty))\,d\xi < \infty, \tag{5.10}$$

where $\mathscr{J}_m$ denotes the interval $[a(\theta^m), b(\theta^{m+q})]$, $m \geqq 1$. Because of the presence of the exponential factors θ^m, the sum in (5.9) is of the same order of magnitude as the last term, which is itself $0(\xi^2 \log^{q-1} \xi / \log_2^{\rho} \xi)$ as $\xi \to \infty$. Thus (2.27) implies (5.9). Moreover, the number of m's for which $\mathscr{J}_m$ contains ξ is on the order of $\log_3 \xi$ as $\xi \to \infty$ and so (2.27) implies (5.10) because $\rho > 1$. □

LEMMA 5.5. (5.5) *holds.*

PROOF. For each k in N_q,

(5.11) $$\sum_{t \in C_k \cap \Delta} E(|J_t'''|) = \sum_{t \in C_k \cap \Delta} \int_{\{|J_t| \geqq \beta(|k|)\}} |J_t| \leqq \int_{\beta(|k|)}^{\infty} \xi \Lambda(d\xi) < \infty$$

by (2.26). Consequently, we can define a $D(T)$-valued, independent-increment process $R = (R(t))_{t \in T}$ by setting

$$R_t = X_t^{***} + \sum_{s \leqq t, s \in \Delta} E(J_s''') ;$$

for each fixed enumeration $(s_n)_{n \geqq 1}$ of the points of Δ in $[\mathbf{0}, t]$,

$$R_t = \sum_{n=1}^{\infty} J_{s_n}'''$$

wp 1. Let $\varepsilon > 0$ and put $N_q(\varepsilon) = \{k \in N_q : \beta(|k|) \leqq \varepsilon \gamma(|k|)\}$. For $k \in N_q(\varepsilon)$ and t in $C_k \cap \Delta$, set

$$V_t = J_t I_{\{\beta(|k|) < |J_t| \leqq \varepsilon\gamma(|k|)\}} , \qquad W_t = J_t I_{\{\varepsilon\gamma(|k|) < |J_t|\}} .$$

Since $P(\{W_t \neq 0$ for some t in $C_k \cap \Delta\} \leqq \Lambda([\varepsilon\gamma(|k|), \infty))$, the "convergence" half of Lemma 5.1 implies that wp 1 only finitely many of the intervals C_k, $k \in N_q(\varepsilon)$, contain a nonzero W_t. To get bounds on the V_t's, put

$$F_k = \{V_t \neq 0 \text{ for at least two } t\text{'s in } B_k \cap \Delta\}$$

for each $k \in N_q(\varepsilon)$. Chebychev's inequality and (2.26) imply that

$$P(F_k) \leqq (|t_k| \beta^{-2}(|\tau_k|) \int \xi^2 \Lambda(d\xi))(\beta^{-2}(|\tau_k|) \sum_{t \in B_k \cap \Delta} \int_{\{\beta(|\tau_k|) \leqq |J_t| \leqq \varepsilon\gamma(|t_k|)\}} J_t^2)$$

and so Lemma 5.4 implies that wp 1 only finitely many F_k occur. Because $|V_t| \leqq \varepsilon\gamma(|k|)$ for all t in $B_k \cap \Delta$, we arrive at (5.7) with X^{**} replaced by R (and with a new G_ε). As ε is arbitrary, this proves

$$\lim_{t \to \infty} \sup_{s \leqq t} |R(s)| / \gamma(|t|) = 0$$

wp 1. Thus to get (5.5) we need only show

$$\sum_{t \leqq k, t \in \Delta} E(|J_t'''|) / \gamma(|k|)$$

tends to 0 as $k \to \infty$ through N_q. But this follows from (5.11) because

$$\sum_{j \leqq k} \beta^{-1}(|j|) \int_{\beta(|j|)}^{\infty} \xi^2 \Lambda(d\xi) \leqq \kappa^2 \delta^{-1} \sum_{j \leqq k} (j_1 j_2 \cdots j_q)^{-\frac{1}{2}} \leqq 2^q \delta^{-1} \kappa^2 |k|^{\frac{1}{2}}$$

where $\kappa^2 = \int \xi^2 \Lambda(d\xi)$. □

6. Proof of Theorem 6 and its corollaries. (a) Proof of Theorem 6. Let $\mu^{(p)}$ be the marginal of μ on the pth component, U_1, of $U_q = U_1 \times \cdots \times U_1$. We give the proof of Theorem 6 first in the case that each $\mu^{(p)}$ is the uniform distribution.

Let $N = (N(s, u))_{s \in [0,\infty), u \in U_q}$ be a Poisson process with mean function $(s, u) \to E(N(s, u)) = sF(u)$, where F is the distribution function of μ. Let

$$0 < \tau_1 < \tau_2 < \tau_3 < \cdots \tag{6.1}$$

be the (random) s-coordinates of the jump-times of N. The differences $\tau_n - \tau_{n-1}$, $n \geqq 1$, are independent random variables, each exponentially distributed with mean 1. Also, the differences $N(\tau_n, \cdot) - N(\tau_{n-1}, \cdot)$, $n \geqq 1$, are independent random processes, each having the same distribution as Y_1 (cf. (2.31)). Thus we may and will represent the sequence $(S_n)_{n \geqq 1}$ by $(N(\tau_n, \cdot))_{n \geqq 1}$; under this representation the definition of H_n^μ (cf. (2.32)) becomes

$$H_n^\mu(s, u) = Z(\tau_{[ns]}, u)/(2n \log_2 n)^{\frac{1}{2}} \tag{6.2}$$

($s \in U_1$, $u \in U_q$), where

$$Z(s, u) = X(s, u) - F(u)X(s, \mathbf{1})$$

with

$$X(s, u) = N(s, u) - E(N(s, u)) = N(s, u) - sF(u)$$

($s \in [0, \infty)$, $u \in U_q$).

LEMMA 6.1. *For $n \geqq 1$, let H_n^* be defined on $U_1 \times U_q$ by*

$$H_n^*(s, u) = X(ns, u)/(2n \log_2 n)^{\frac{1}{2}} .$$

Then wp 1 *the sequence* (H_n^*) *is relatively compact (with respect to the uniform metric* (d) *and its limit points coincide with*

$$K^* \equiv \{\mathscr{I}_\nu(y) : \textstyle\int_{U_1 \times U_q} y^2 \, d\nu \leqq 1\} ,$$

where $\nu = \lambda \times \mu$ (cf. (2.33)).

PROOF. This follows from a slight modification of the proof of Theorem 2. For t_n, use the point $(n, \mathbf{1}) \in U_1 \times U_q$, $n \geqq 1$. In condition (2.9), replace the role of Lebesgue measure on U_q by ν on $U_1 \times U_q$. In Lemma 3.3, use the fact that for any G, Π_G puts all its mass at the point $\mathbf{1}$. The assumption that μ has uniform 1-dimensional marginals implies that (4.2) (with λ replaced by ν, and q by $q + 1$) is valid. In Lemma 4.4, use the fact that the functions in K^* are equicontinuous with respect to their first argument ($s \in U_1$). Finally, replace B_c, as defined by (4.5), by $[0, 2/c) \times U_q$. □

LEMMA 6.2. *For $n \geqq 1$, let H_n^{**} be defined on $U_1 \times U_q$ by*

$$H_n^{**}(s, u) = Z(ns, u)/(2n \log_2 n)^{\frac{1}{2}} = H_n^*(s, u) - F(u)H_n^*(s, \mathbf{1}) .$$

Then wp 1, (H_n^{**}) *is relatively compact and its limit points coincide with* K^μ *(defined by* (2.34), (2.35), *and* (2.36)).

PROOF. For x in $\mathscr{D}(U_1 \times U_q)$, define $\Phi(x) \equiv \hat{x}$ in $\mathscr{D}(U_1 \times U_q)$ by

$$\hat{x}(s, u) = x(s, u) - F(u)x(s, \mathbf{1}) .$$

Since Φ is continuous (with respect to the uniform metric), it suffices by Lemma 2.2 to show that $K^\mu = \Phi(K^*)$. Suppose then that x is in K^*. Clearly $\hat{x}$ satisfies (2.36). Also, since

$$\hat{x}(s, u) = \int_{\xi \leqq s, \eta \leqq u} (\dot{x}(\xi, \eta) - (\int \dot{x}(\xi, \zeta)\mu(d\zeta)))\nu(d\xi, d\eta) ,$$

$\hat{x}$ is the indefinite integral, relative to ν, of the function $\dot{\hat{x}}$ whose section at $s \in U_1$ is

$$\dot{\hat{x}}_s = \dot{x}_s - \langle \dot{x}_s, I_{U_q} \rangle_\mu ,$$

where $\langle \cdot, \cdot \rangle_\mu$ denotes the inner product of $L_2(U_q, \mu)$; thus (2.34) holds. Finally, (2.35) is satisfied because $\dot{\hat{x}}$ is just the projection of $\dot{x}$ in $L_2(U_1 \times U_q, \nu)$ onto the subspace of functions z such that $\langle z_s, I_{U_q} \rangle_\mu = 0$ for all s in U_1. This proves that $\Phi(K^*) \subset K^\mu$; the opposite inclusion is immediate.

LEMMA 6.3. *Let H_n^μ be defined by* (6.2). *Then* wp 1, (H_n^μ) *is relatively compact and its limit points coincide with K^μ.*

PROOF. We can write $Z(\tau_{[ns]}, u)$ as $Z(n(\tau_{[ns]}/n), u)$. The strong law of large numbers implies that $\max_{k \leqq n} |\tau_k/n - k/n| \to 0$ wp 1, while Lemma 6.2 implies that the H_n^{**}'s are asymptotically equicontinuous with respect to their first argument ($s \in U_1$). Thus Lemma 6.3 follows from Lemma 6.2. □

This proves Theorem 6 under the assumption that μ has uniform 1-dimensional marginals. For the general case, let F_p be the distribution function of the marginal of μ on the pth component, U_1, of U_q, and let G_p be the inverse of F_p, defined by

$$G_p(a) = \inf \{b : a \leqq F_p(b)\}$$

($a \in U_1$). Define $\mathscr{G}: U_q \to U_q$ by

$$\mathscr{G}((u^{(1)}, \cdots, u^{(q)})) = (G_1(u^{(1)}), \cdots, G_q(u^{(q)})) .$$

We can find (using an easy extension of the univariate procedure) a probability μ^* on U_q having uniform 1-dimensional marginals and inducing μ via $\mathscr{G}$ (i.e., $\mu = \mu^* \mathscr{G}^{-1}$). Let $(V_n^*)_{n \geqq 1}$ be independent U_q-valued random vectors, each having distribution μ^*. Then we may and will represent the sequence (V_n) by $(\mathscr{G}(V_n^*))$. Since $a \leqq F_p(b)$ iff $G_p(a) \leqq b$, the definition of H_n^μ (cf. (2.32)) under this representation becomes

$$H_n^\mu(s, u) = H_n^{\mu^*}(s, \mathscr{F}(u)) ,$$

where $\mathscr{F}: U_q \to U_q$ is given by

$$\mathscr{F}((u^{(1)}, \cdots, u^{(q)})) = (F_1(u^{(1)}), \cdots, F_q(u^{(q)})) .$$

For x in $\mathscr{D}(U_1 \times U_q)$, define $\Phi(x) \equiv \hat{x}$ in $\mathscr{D}(U_1 \times U_q)$ by

$$\hat{x}(s, u) = x(s, \mathscr{F}(u)) .$$

Since Φ is continuous (with respect to the uniform metric) Theorem 6 will be completely proved once we establish

LEMMA 6.4. $K^\mu = \Phi(K^{\mu^*})$.

PROOF. Let x be in K^{μ^*}. Then

$$\begin{aligned}\hat{x}(s, u) &= \textstyle\int_{\xi\leq s, \eta\leq \mathscr{T}(u)} \dot{x}_\xi(\eta)\mu^*(d\eta)\, d\xi \\ &= \textstyle\int_{\xi\leq s} (\int_{[0,u]} (\mathscr{G}(\eta))\dot{x}_\xi(d\eta)\mu^*(d\eta))\, d\xi \\ &= \textstyle\int_{\xi\leq s} (\int I_{[0,u]} (\mathscr{G}(\eta))E(\dot{x}_\xi \mid \mathscr{G} = \mathscr{G}(\eta))\mu^*(d\eta))\, d\xi \\ &= \textstyle\int_{\xi\leq s} (\int_{\zeta\leq u} E(\dot{x}_\xi \mid \mathscr{G} = \zeta)\mu(d\zeta))\, d\xi ,\end{aligned}$$

where $E(\dot{x}_\xi \mid \mathscr{G} = \cdot)$ is the conditional expectation (defined on the target space of $\mathscr{G}$) of $\dot{x}_\xi$ in $L_2(U_q, \mu^*)$ with respect to the mapping $\mathscr{G}$. Since

$$\textstyle\int E(\dot{x}_\xi \mid \mathscr{G} = \zeta)^2\mu(d\zeta) \leqq \int \dot{x}_\xi(\eta)^2\mu^*(d\eta) ,$$

the $L_2(\nu)$ norm of the mapping $(\xi, \zeta) \to E(\dot{x}_\xi \mid \mathscr{G} = \zeta)$ cannot be greater than the $L_2(\nu^*)$ norm of $\dot{x}$. Thus (2.34) and (2.35) hold for $\hat{x}$, and (2.36) is immediate. This proves $\Phi(K^{\mu^*}) \subset K^\mu$; the opposite inclusion is easily established (check that y in K^μ is the image of x in K^{μ^*}, where $\dot{x}_\xi(\eta) = \dot{y}_\xi(\mathscr{G}(\eta))$). □

(b) Proof of IV, Section 2. Let x be in K^μ. By (2.36) we have $\int_0^s \langle \dot{x}_\xi, I_{U_q}\rangle_\mu\, d\xi = 0$ for all s in U_1, so $\langle \dot{x}_\xi, I_{U_q}\rangle_\mu = 0$ for almost all ξ. Now let A be in $\mathscr{R}$ (notation of IV). Then

$$\begin{aligned}|\langle \dot{x}, I_{U_1\times A}\rangle_\nu| &= |\textstyle\int \langle \dot{x}_\xi, I_A\rangle_\mu\, d\xi| \\ &= |\textstyle\int \langle \dot{x}_\xi, I_A - \mu(A)\rangle_\mu\, d\xi| \\ &\leqq \textstyle\int ||\dot{x}_\xi||_\mu ||I_A - \mu(A)||_\mu\, d\xi \\ &\leqq (\mu(A)(1 - \mu(A)))^{\frac{1}{2}}(\textstyle\int ||\dot{x}_\xi||_\mu^2\, d\xi)^{\frac{1}{2}} \\ &\leqq (\mu(A)(1 - \mu(A))))^{\frac{1}{2}} .\end{aligned}$$

This gives (2.39) (use $\sup_u |x(1, u)| = \sup_{A\in\mathscr{R}} |\langle \dot{x}, I_{U_1\times A}\rangle_\nu|$). A similar argument yields (2.40).

(c) Proof of V, Section 2. Here it suffices to show that $K_\Gamma{}^\mu$ is the image of K^μ under the (continuous) mapping $\Phi : x \to \hat{x}$, where $\hat{x}$ in $\mathscr{D}(U_1)$ is defined by

$$\hat{x}(s) = \Gamma(x_s)$$

($s \in U_1$) (cf. (2.41) and (2.42)).

First suppose that b is in $K_\Gamma{}^\mu$. Choose (using the compactness of W_μ) a w in W_μ such that $\Gamma(w) = c_\Gamma(\mu)$. Define x in $\mathscr{D}(U_1 \times U_q)$ by

$$x(s, u) = (b(s)/c_\Gamma(\mu))w(u) .$$

Then x is in K^μ, and $\Phi(x) = b$.

Next suppose that x is in K^μ; we have to show that $\Phi(x)$ is in $K_\Gamma{}^\mu$. Let $s_1 < s_2$ in U_1. Then

$$|\Gamma(x_{s_2}) - \Gamma(x_{s_1})| \leqq \Gamma(x_{s_2} - x_{s_1}) = \Gamma(\textstyle\int_{(s_1,s_2]} f_\xi(\cdot)\, d\xi) ,$$

where $f_\xi(u) = \int_{\eta\leq u} \dot{x}_\xi(\eta)\mu(d\eta)$. We will show below that

$$\Gamma(\textstyle\int_{(s_1,s_2]} f_\xi\, d\xi) \leqq \int_{(s_1,s_2]} \Gamma(f_\xi)\, d\xi . \tag{6.3}$$

Accepting this for the moment, we get

$$\begin{aligned}|\Gamma(x_{s_2}) - \Gamma(x_{s_1})| &\leqq \textstyle\int_{(s_1,s_2]} ||\dot{x}_\xi||_\mu \Gamma(f_\xi/||\dot{x}_\xi||_\mu)\, d\xi \\ &\leqq c_\Gamma(\mu) \textstyle\int_{(s_1,s_2]} ||\dot{x}_\xi||_\mu\, d\xi \\ &\leqq c_\Gamma(u)(s_2 - s_1)^{\frac{1}{2}}(\textstyle\int_{(s_1,s_2]} ||\dot{x}||_\mu^2\, d\xi)^{\frac{1}{2}} .\end{aligned}$$

A martingale argument like that used in Section 3 to prove that K is closed now implies that $\Phi(x)$ is in K_{Γ^μ}.

It remains to establish (6.3). Let $\mathscr{A}_k$, $k \geqq 1$, be a nested sequence of rectangular partitions of $U_1 \times U_q$ such that the mesh size of $\mathscr{A}_k$ tends to 0 as $k \to \infty$. Let ${}_k\dot{x}$ be the conditional expectation of $\dot{x}$ relative to the partition $\mathscr{A}_k$ and the measure ν, and let ${}_kf_\xi$ be defined in terms of ${}_k\dot{x}$ in the same way f_ξ is defined in terms of $\dot{x}$. Using the sublinearity and positive homogeneity of Γ, it is easy to see that

$$\Gamma(\textstyle\int_{(s_1,s_2]} {}_kf_\xi \, d\xi) \leqq \int_{(s_1,s_2]} \Gamma({}_kf_\xi)\, d\xi \; . \tag{6.4}$$

Martingale theory implies that ${}_k\dot{x}$ converges to $\dot{x}$ in $L_2(\nu)$, therefore also in $L_1(\nu)$, and so $\int_{(s_1,s_2]} {}_kf_\xi \, d\xi$ converges, with respect to the uniform metric on $\mathscr{D}(U_q)$, to $\int_{(s_1,s_2]} f_\xi \, d\xi$. Since Γ is continuous, the left-hand side of (6.4) tends to the left-hand side of (6.3). On the other hand, by the assumption that Γ is majorized by the uniform norm $||\cdot||$, we have $|\Gamma(f_\xi) - \Gamma({}_kf_\xi)| \leqq \Gamma(f_\xi - {}_kf_\xi) \leqq ||f_\xi - {}_kf_\xi|| \leqq ||\dot{x}_\xi - {}_k\dot{x}_\xi||_{L_2(\mu)}$. Thus $\Gamma({}_kf_\cdot) \to \Gamma(f_\cdot)$ in $L_2(U_1, d\xi)$, and so the right-hand side of (6.4) converges to the right-hand side of (6.3).

Acknowledgment. I would like to thank Ron Pyke, Jack Kiefer, and the referee for some helpful suggestions.

REFERENCES

[1] Bahadur, R. (1971). Some limit theorems in statistics. *NSF Regional Conference Series in Applied Mathematics*. Monograph 4, *SIAM*.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[3] Cabaña, E. M. (1972). On barrier problems for the vibrating string. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **22** 13–24.

[4] Cassels, J. W. S. (1951). An extension of the law of the iterated logarithm. *Proc. Cambridge Philos. Soc.* **47** 55–64.

[5] Choyer, J. (1966). On Strassen's version of the log log law. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **8** 83–90.

[6] Chung, K. L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.* **67** 36–50.

[7] Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[8] Feller, W. (1943). Generalization of a probability limit theorem of Cramer. *Trans. Amer. Math. Soc.* **54** 361–372.

[9] Feller, W. (1968). An extension of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* **18** 343–355.

[10] Feller, W. (1969). Limit theorems for probabilities of large deviations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **14** 1–20.

[11] Finkelstein, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42** 607–615.

[12] Gál, I. S. (1951). Sur la majorization des suites des fonctions. *Nederl. Akad. Wetensch. Proc. Ser. A* **54** 243–251.

[13] Gaposhkin, V. P. (1965). The law of the iterated logarithm for Cesaro's and Abel's methods of summation. *Theor. Probability Appl.* **10** 411–420.

[14] Gikhman, I. I. and Skorokhod, A. V. (1969). *Introduction to the Theory of Random Processes*. Saunder, Philadelphia.

[15] Hartman, P. and Wintner, A. (1941). On the law of the iterated logarithm. *Amer. J. Math.* **63** 169–176.

[16] KIEFER, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649-660.

[17] KIEFER, J. (1971). Skorokhod embedding of multivariate rv's, and the sample df. To appear in *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.*

[18] KOLMOGOROV, A. (1929). Das Gesetz des iterierten Logarithmus. *Math. Ann.* **101** 126-135.

[19] MÜLLER, D. W. (1970). On Glivenko-Cantelli convergence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16** 195-210.

[20] OREY, S. and PRUITT, W. E. (1973). Sample functions of the N-parameter Wiener process. *Ann. Probability.* To appear.

[21] PHILLIP, W. (1969). Das Gesetz vom iterierten Logarithmus mit Anwendungen auf die Zahlentheorie. *Math. Ann.* **180** 75-94.

[22] PHILLIP, W. (1971). Mixing sequences of random variables and probabilistic number theory. *Mem. Amer. Math. Soc.* **114**.

[23] PYKE, R. (1971). Empirical processes. Technical Report No. 28, Univ. of Washington.

[24] PYKE, R. (1972). Partial sums of matrix arrays and Brownian sheets. To appear in *Stochastic Geometry and Stochastic Analysis*, eds. E. F. Harding and D. G. Kendall. Wiley, New York.

[25] ROYDEN, H. L. (1963). *Real Analysis.* Macmillan, New York.

[26] SMIRNOV, N. (1944). Approximate laws of distribution of random variables from empirical data (in Russian). *Uspehi. Mat. Nauk.* **10** 179-206.

[27] SMYTHE, R. (1973). Strong law of large numbers for r-dimensional arrays of random variables. *Ann. Probability.* **1** 164-170.

[28] STACKELBERG, O. (1964). On the law of the iterated logarithm. *Nederl. Akad. Wetensch. Proc. Ser. A* **67** 48-67.

[29] STATULEVICIUS, V. A. (1966). On large deviations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **6** 133-144.

[30] STEIGER, W. L., and ZAREMBA, S. K. (1972). The converse of the Hartman-Wintner theorem. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **22** 193-194.

[31] STOUT, W. F. (1970). The Hartman-Wintner law of the iterated logarithm for martingales. *Ann. Math. Statist.* **41** 2158-2160.

[32] STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** 211-226.

[33] STRASSEN, V. (1966). A converse to the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4** 265-268.

[34] TOMPKINS, R. J. (1971a). An iterated logarithm theorem for some weighted averages of independent random variables. *Ann. Math. Statist.* **42** 760-763.

[35] TOMPKINS, R. J. (1971b). Some iterated logarithm results related to the central limit theorem. *Trans. Amer. Math. Soc.* **156** 185-192.

[36] WICHURA, M. J. (1968). Inequalities with applications to the weak convergence of random processes with multidimensional time parameters. *Ann. Math. Statist.* **40** 681-687.

[37] ZAREMBA, S. K. (1971). Sur la discrépance des suites altéatoires. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **20** 236-248.

[38] ZIMMERMAN, G. (1972). Some sample function properties of the two-parameter Gaussian process. *Ann. Math. Statist.* **43** 1235-1246.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
1118 EAST 58TH STREET
CHICAGO, ILLINOIS 60637