RANDOMNESS AND EXTRAPOLATION

D. W. MÜLLER

UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction

This article proposes a theory of sequential observation as a basis for a definition of random sequences—which is more general than the approaches inspired by the intuitive situations of gambling and sequential testing. It investigates implications of the constructivist thesis which equates sequential observation and extrapolation, in the case of repeated independent random experiments.

As shown in Section 4 this leads to a definition of the concept of "infinite random sequence" considerably narrower than those proposed by Martin-Löf [10] and Schnorr [21] (a discussion of these approaches can be found in Section 3). There exist sequences random in the sense of Martin-Löf and generated by finite rules (of the class $\Sigma_2 \cap \Pi_2$), revealing the incompatibility of these notions and intuition.

The approach of this article will be guided by the intuitive notion of random phenomena as collections of finite samples which will, on the average, be ultimately observed in sequential experiments. The corresponding class of random sequences does not show pathologies of the type indicated.

In Section 5 "on the average" is interpreted as "with high probability" rather than "with probability 1" (as before), and distribution limit theorems (invariance principles) are stated yielding the probability levels of certain sequentially observable events related to almost sure convergence theorems.

Let $x_1 x_2 \cdots$ be a machine generated sequence subject to sequential observation, information about the computing mechanism not being available. After a large number of observations $x_1 x_2 \cdots x_n$ have been taken it is possible to reconstruct the generating rule from the data; this is equivalent to an extrapolation of $x_1 x_2 \cdots x_n$. The number n being unknown, however, all one can say is that an extrapolation is possible ultimately.

In contrast to the above, assume now that $x_1 x_2 \cdots$ is generated by a random experiment (say, coin tossing). Then, despite considerable regularities that might occur in the first outcomes, the observer will find himself unable to extrapolate, the complexity of a random sequence being unattained by any extrapolation.

However, functionals operating on the initial segments of a sequence may very well have a regular set of values, under the assumption of randomness. An example is given by

$$(1.1) \qquad \Phi(x_1 \cdots x_n) = \sup_{100 \le k \le n} \frac{1}{k} \left| \sum_{i=1}^{k} x_i \right|$$

with $P[x_i = 1] = P[x_i = -1] = \frac{1}{2}, i \ge 1$.

The collection of all functionals which have regular values with probability 1 should summarize the intuitive impression of randomness. This thesis may appear objectionable since intuition associates the recurrence of certain events with randomness as well (for example, the crossing of zero by partial sums of independent centered random variables). However, as a statement about a completed infinity, recurrence has a merely mathematical meaning unless supplemented by information about the frequency of recurrence times. In this case again, it can be described by a functional with regular values.

The notion of a functional with regular values ("observable") will be weakened to that of a constant partial recursive functional or "trace class," which is a set of finite sequences that, for almost every infinite sample, contains all but finitely many of its initial segments. How far the family of trace classes can be reduced will be studied in Section 4; the author conjectures that no single trace class suffices to describe the event of randomness. Intuitively this would mean that randomness fails to be a "random phenomenon." However, as will be shown, a surprisingly small family of trace classes turns out to be sufficient, namely those consisting of segments with high Δ_2 complexity.

Notation (Section 2 to Section 4). Let N be the set of positive integers. Call $\mathscr{X}$ the space of all infinite sequences $x = x_1 x_2 \cdots$ where the x_i are real numbers or where, as in Section 3 and Section 4, the x_i are 0 or 1. Let Ω be the corresponding space of all finite sequences $x = x_1 \cdots x_n$ of arbitrary length $\ell(x) = n \ge 0$. The sequence of length 0 will be written $\square$. If $x \in \mathscr{X}$ we write $x_{n]} = x_1 \cdots x_n$ for the nth initial segment of x. The same notation will be used for $x \in \Omega$, $x_{n]}$ being equal to the string of the first n terms of x if $\ell(x) \ge n$, and equal to x otherwise. The space Ω is ordered by $\ll$ (the ordering of continuation): $x \ll y$ if $\ell(x) \le \ell(y)$ and $y_{\ell(x)]} = x$. In the case of binary sequences $\le$ will be the lexicographical ordering in Ω: if $\ell(x) = \ell(y)$ then $x \le y$ if $y - x$ (coordinatewise) is of the form $0 \cdots 01 z_j \cdots z_{\ell(x)}$; otherwise if $\ell(x) < \ell(y)$ then also $x \le y$. The product xy of $x \in \Omega$ and a (finite or infinite) sequence y is $x_1 \cdots x_{\ell(x)} y_1 y_2 \cdots$. Throughout, unless otherwise stated, $\mathscr{X}$ is supposed to be endowed with the product topology and the sigma field of Borel sets $\mathscr{A}$: the sigma field generated by the first n coordinates is $\mathscr{A}_n$. For any subset $\Phi \subset \Omega$, let $\sup \Phi = \bigcup_{n>0} \{x : x_{n]} \in \Phi\}$ and $\inf \Phi = \bigcap_{n>0} \{x : x_{n]} \in \Phi\}$, $\limsup \Phi = \bigcap_{m>0} \bigcup_{n \ge m} \{x : x_{n]} \in \Phi\}$, $\liminf \Phi = \bigcup_{m>0} \bigcap_{n \ge m} \{x : x_{n]} \in \Phi\}$. "Function" means real valued or integer valued function, "partial function" partially defined function. The indicator function of the set A is I_A.

2. The logic of sequential experience

This section states definitions and simple consequences.

DEFINITION 2.1. *A function Φ on Ω is called trace function if for every $n > 0$ the restriction of Φ to sequences of length n is Borel measurable.*

Let lim* denote the limit operation in the discrete topology of the reals.

DEFINITION 2.2. *A partial function f on $\mathcal{X}$ is called observable if there is a trace function Φ such that*

(i) $\lim^*_{n \to \infty} \Phi(x_{n]}) = f(x)$ *whenever x is in the domain of f,*

(ii) *this limit does not exist otherwise.*

Every observable is an $\mathcal{A}$-measurable function on an $\mathcal{A}$-measurable domain, its range is countable. A constant observable will be called observable event (under an obvious identification); the observable events turn out to be precisely the domains of observables.

DEFINITION 2.3. *An event $E \subset \mathcal{X}$ is called observable if there is a constant observable with domain E.*

PROPOSITION 2.1. *The following statements are equivalent:*

(i) E *is an observable event;*

(ii) *there is an observable being equal to 1 exactly on E;*

(iii) E *is the domain of an observable;*

(iv) *there is a function $\Xi : \Omega \rightarrow \{0, 1, 2\}$ such that $E = \{x \in \mathcal{X} : n \rightarrow \Xi(x_{n]})$ is a recursive function$\}$.*

PROOF. For (i) $\Rightarrow$ (iii) the proof is immediate.

Now we prove (iii) $\Rightarrow$ (ii). Let f be an observable with domain E being defined by the trace function Φ. Define another trace function Φ' by

$$(2.1) \qquad \Phi'(x_1 \cdots x_n) := \begin{cases} 1 & \text{if} \quad \Phi(x_1 \cdots x_n) = \Phi(x_1 \cdots x_{n-1}), \\ 0 & \text{otherwise.} \end{cases}$$

Φ' defines an observable with values 0 and 1 only, equal to 1 precisely on E.

Next we prove (ii) $\Rightarrow$ (i). Let f be an observable being equal to 1 exactly on E, defined by the trace function Φ. Define another Φ' by

$$(2.2) \qquad \Phi'(x_1 \cdots x_n) := \begin{cases} 1 & \text{if} \quad \Phi(x_1 \cdots x_n) = 1, \\ n & \text{if} \quad \Phi(x_1 \cdots x_n) \neq 1. \end{cases}$$

Then Φ' defines E.

For (ii) $\Rightarrow$ (iv), let Φ be a $\{0, 1\}$ valued trace function defining E in the sense of (ii); define Ξ by

$$(2.3) \qquad \Xi(x_1 \cdots x_n) := \begin{cases} 1 & \text{if} \quad \Phi(x_1 \cdots x_n) = 1, \\ F\big(n - \sum_{j \leq n} \Phi(x_1 \cdots x_j)\big) & \text{otherwise,} \end{cases}$$

where F is a nonrecursive function with values in $\{0, 2\}$.

Finally we prove (iv) $\Rightarrow$ (ii). Define Φ by

$$(2.4) \qquad \Phi(x_1 \cdots x_n) := \begin{cases} 1 & \text{if} \quad K\big(\Xi(x_1)\cdots\Xi(x_1\cdots x_n)\big|n\big) \\ & \qquad \leq \max_{j<n} K\big(\Xi(x_1)\cdots\Xi(x_1\cdots x_j)\big|j\big), \\ \\ 0 & \text{otherwise,} \end{cases}$$

where K is a version of Kolmogorov's complexity measure; Φ defines E after Proposition 3.1.

PROPOSITION 2.2. *The observable events are the events of the form* $E = \bigcup_{i=1}^{\infty} \inf \Psi_i$, *where* $\Psi_i \subset \Omega$ *are such that*

$$(2.5) \qquad \{x : x \in \Omega, \ell(x) = n\} \cap \Psi_i \in \mathscr{A}_n,$$

$n, i \in N$. *In particular, F_σ subsets of $\mathscr{X}$ are observable.*

PROOF. Without restriction of generality one can assume $\inf \Psi_i \subset \inf \Psi_{i+1}$, $i \in N$; thus E^c is the intersection of the decreasing family $\{\sup \Psi_i^c : i \in N\}$. On Ω define

$$(2.6) \qquad m(x) := \max \{n : \sup \{x\} \subset \sup \Psi_n^c\}.$$

Then $x^1 \ll x^2$ implies $m(x^1) \leq m(x^2)$. Moreover $x \in \sup \Psi_n^c - \sup \Psi_{n+1}^c$ if and only if $\max_i m(x_{i]}) = \lim_{i \to \infty} m(x_{i]}) = n$. Hence $x \in E^c$ is equivalent to $m(x_{n]})$ unbounded as $n \to \infty$. Now for $x \in \Omega$ put

$$(2.7) \qquad \Phi(x) := \begin{cases} 1 & \text{if} \quad m(x) \leq \max \{m(y) : y \ll x, y \neq x\}, \\ 0 & \text{otherwise.} \end{cases}$$

The function Φ defines E.

COROLLARY 2.1. *The class of observable events is closed with respect to countable unions and finite intersections; it is not closed with respect to complementations.*

COROLLARY 2.2. *Let $\mathscr{X} = \{0, 1\}^N$; then the class of observable events coincides with the class of F_σ subsets of $\mathscr{X}$.*

COROLLARY 2.3. *Let P be a probability measure on $\mathscr{A}$, and E an event of P measure 1. Then there is an observable event $E' \subset E$ such that $PE' = 1$.*

COROLLARY 2.4. *For every observable f there exists an observable g, defined almost everywhere P (with P as above) and extending f.*

PROOF. Let E denote the domain of f; we can assume $PE < 1$. There is an observable event $E' \subset E^c$ such that $PE' = 1 - PE$. Let Φ_1, Φ_2 be $\{0, 1\}$ valued trace functions for E, E', respectively, and let Ψ be a trace function for f. A $\{1, 2\}$ valued trace function Ξ will be defined inductively as follows:

$$\Xi(x_1 \cdots x_{n+1})$$
$$:= 1 \text{ if} \begin{cases} \text{either} & \Phi_1(x_1 \cdots x_{n+1}) = 1 \ \& \ \Xi(x_1 \cdots x_n) = 1 \\ \text{or} & \Phi_1(x_1 \cdots x_{n+1}) > \Phi_2(x_1 \cdots x_{n+1}) \ \& \ \Xi(x_1 \cdots x_n) = 2 \\ \text{or} & \Phi_1(x_1 \cdots x_{n+1}) = \Phi_2(x_1 \cdots x_{n+1}) = 0 \end{cases}$$
$$(2.8)$$

$$\Xi(x_1 \cdots x_{n+1})$$
$$:= 2 \text{ if} \begin{cases} \text{either} & \Phi_2(x_1 \cdots x_{n+1}) = 1 \ \& \ \Xi(x_1 \cdots x_n) = 2 \\ \text{or} & \Phi_2(x_1 \cdots x_{n+1}) > \Phi_1(x_1 \cdots x_{n+1}) \ \& \ \Xi(x_1 \cdots x_n) = 1. \end{cases}$$

The trace function Φ defines g:

(2.9)
$$\Phi(x_1 \cdots x_n) := \begin{cases} \Phi_1(x_1 \cdots x_n) \vee \Phi_2(x_1 \cdots x_n) & \text{if } \Xi(x_1 \cdots x_n) = \Xi(x_1 \cdots x_{n-1}), \\ 0 & \text{otherwise.} \end{cases}$$

PROPOSITION 2.3. *The almost everywhere P defined observables elements of $\mathscr{L}^\infty(P)$ form a dense subalgebra and a sublattice of $\mathscr{L}^\infty(P)$.*

PROOF. For every simple function there is an observable function differing from it only on a nullset; the detailed proof of this uses the construction of the preceding proof and is omitted.

3. Martingales and randomness tests

This section is mainly a historical one. It reviews the approaches by J. Ville, P. Martin-Löf and C. P. Schnorr towards a definition of random sequence. The appendix contains a short exposition of the complexity theory of Kolmogorov and Martin-Löf. In the sequel (Sections 3, 4) $\mathscr{X}$ and Ω will always refer to the binary case; P will be the "coin tossing" measure (product of uniform distributions on $\{0, 1\}$).

3.1. According to Ville [27] the concept of random sequence refers to a preassumed gambling system ("martingale"): whenever a "random" sequence occurs the gambler's gain stays bounded throughout the whole infinite game. A gambling system in the sense of Ville can be characterized by two functions λ, μ on Ω. The value $\lambda(x)$ (respectively $\mu(x)$) is the proportion of the gambler's capital $s(x)$ he is willing to bet on "1" (respectively "0") in the $(\ell(x) + 1)$st trial to gain $2\lambda(x)s(x)$ or $2\mu(x)s(x)$, respectively. Clearly $\lambda(x) + \mu(x) \leqq 1$, and $s(\square)$ may be taken to be 1. The sequence s is a nonnegative martingale with respect to $(\mathscr{A}_n)$, with $s(\square) = 1$ (called simply martingale in the sequel); and each such martingale results from some gambling system λ, μ. Ville's definition will now be stated.

DEFINITION 3.1. *Let s be a martingale (in the sense indicated above). Then $x \in \mathscr{X}$ is called s random if $\sup_n s(x_{n]}) < \infty$.*

This concept is very flexible as the following theorem shows.

THEOREM 3.1 (Ville [27]).

(i) $P\{x : \sup_n s(x_{n]}) < \infty\} = 1$;

(ii) *for every event E, $PE = 1$, there is a martingale s such that $\{x : \sup_n s(x_{n]}) < \infty\} \subset E$.*

PROOF. Part (i) follows from the martingale inequality

(3.1)
$$P\{x : \sup_{i \leqq n} s(x_{i]}) \geqq \lambda\} \leqq \frac{Es(x_{n]})}{\lambda}.$$

As for part (ii) a martingale can be generated by a suitable function f: let $\{G_n : n \in N\}$ be a family of open sets such that $E^c \subset G_n$, $G_{n+1} \subset G_n$, $PG_n = 2^{-2n}$,

$n \in N$ and put $f(x) := \sum_{n \geq 1} 2^n I_{G_n}(x)$ such that $Ef = 1$. Define $s(x) := \int f(xy)P(dy)$, $x \in \Omega$.

Ville could not settle the question of which reference martingale should be used for the definition of random sequence. Further criteria are needed. The following is widely accepted.

POSTULATE Δ_1.   *Recursive sequences are not random.*

The set of all recursive sequences being countable there is a martingale s_0 unbounded on every recursive sequence, whence "s_0 random" implies "nonrecursive." However, every such martingale is noncomputable in the sense that its integer part is not a recursive function.

THEOREM 3.2.   *If the martingale s_0 is unbounded on every recursive sequence then* $[s_0(x)] = \max \{n \in N : n \leq s_0(x)\}$ *is not recursive as a function of* $x \in \Omega$.

This is a consequence of Theorem 3.3 below.

DEFINITION 3.2.   *An observable event E is called regular if there is a recursively enumerable set $\Phi \subset \Omega$ such that* $\lim \inf \Phi = E$ *and* $x0 \in \Phi$ *or* $x1 \in \Phi$ *whenever* $x \in \Phi$ *(see Definitions 4.1, 4.2, 4.4). A Φ of this type will also be called regular.*

THEOREM 3.3.   *Every nonempty regular event contains a recursive sequence.*

PROOF.   The proof is clear.

PROOF OF THEOREM 3.2.   For every martingale s the event $\{x : \sup_n s(x_n]) < \infty\}$ is regular. This follows from the martingale equation $s(x0) + s(x1) = 2s(x)$ which implies $[s(x0)] \wedge [s(x1)] \leq [s(x)]$.

In addition this proves that $\{x : \sup_n s(x_n]) < \infty\} = \lim \inf \Phi$ where Φ is recursive.

3.2.   The idea of taking the behavior of infinite sequences towards sequential tests as a defining property of randomness is due to Martin-Löf [10]. It leads to an observable event E which can be written in the form $\lim \inf \Phi$, where Φ^c is recursively enumerable. Indeed, this corresponds to intuition since for any test, the critical region has to be fixed in advance, and hence in some constructive way. Moreover, tests for infinite sequences should be sequential, a sequence being rejected (at some level) on the basis of only finitely many observations. Critical regions, therefore, are suitably represented by open sets. The following is Martin-Löf's definition of a sequential test.

DEFINITION 3.3.   *A sequential test is given by a recursively enumerable set $U \subset N \times \Omega$ such that*

(i) *if $U_n = \{x \in \Omega : (n, x) \in U\}$ then the regions $\sup U_n$ are nested*: $\sup U_1 \supset \sup U_2 \supset \cdots$;

(ii) $P(\sup U_n) \leq 2^{-n}$, $n \in N$.

DEFINITION 3.4.   *A sequential test $U \subset N \times \Omega$ is called universal if for every sequential test V there is a constant c such that* $\sup V_{n+c} \subset \sup U_n$, $n \in N$.

Sequences being rejected at every level $\alpha > 0$ by some sequential test will also be rejected by a universal test. Consequently the nullset R_M^c of those sequences rejected at every level $\alpha > 0$ by a universal test does not depend on the particular choice of this test.

THEOREM 3.4 (Martin-Löf [10]). *There exist universal sequential tests.*

The set R_M is an observable event; it satisfies Postulate Δ_1.

COROLLARY 3.1. *The set R_M does not contain recursive sequences.*

PROOF. Every recursive sequence defines a sequential test in an obvious way.

THEOREM 3.5. $R_M = \lim \inf \Phi$ *with* $\Phi \subset \Omega$ *recursively enumerable.*

PROOF. This is an immediate consequence of Theorem 4.1 below.

3.3. In this connection another class of events has been proposed in the literature (Schnorr [21]): namely the nullsets of sequential tests (Definition 3.3) which satisfy the additional requirement

(iii) $\{PU_m : m = 1, 2, \cdots\}$ forms a recursively enumerable sequence of computable real numbers.

These nullsets are called "totally recursive" in [21].

THEOREM 3.6. *The complements of totally recursive nullsets contain regular events of P measure 1.*

PROOF. Let U be a sequential test such that $E^c := \cap_{i \geq 1} \sup U_i$ is totally recursive. We are going to construct a recursively enumerable $\Phi \subset \Omega$ such that Φ is regular, $\lim \inf \Phi \subset E$ and $P(\lim \inf \Phi) = 1$. The essence of this procedure can be described as follows: let

$$(3.2) \qquad U_n^* := \{x \in \Omega : P(\sup \{x\} \cap (\sup U_n)^c) > 0\} \supset U_n,$$

$n \in N$; then U_n^{*c} is a recursively enumerable tree (that is, containing $y \ll x$ whenever $x \in U_n^{*c}$), $\inf (U_n^{*c})$ is regular and differs from $\inf (U_n^c)$ by a P nullset. Discarding $x \in U_n^{*c}$ with suitably small $P(\sup \{x\} \cap (\sup U_n^*)^c)$ one replaces U_n^{*c} by a subtree Ψ_n, $n \in N$, such that (i) $\inf \Psi_n$ is regular, (ii) $P(\inf \Psi_n) \geq 1 - 2^{-n+1}$, (iii) $\Psi_n \subset \Psi_{n+1}$ $(n \in N)$, (iv) $\{(n, x) : x \in \Psi_n\} \subset N \times \Omega$ is recursive. Then $E' := \cup_{n \geq 1} \inf \Psi_n$ is a regular subset of E of P measure 1. To verify regularity, let $m(x) := \min \{n : x \in \Psi_n\}$. This is a monotone increasing recursive function for $\ll$. The set

$$(3.3) \qquad \Phi := \{x \in \Omega : m(x) = m(x_{\ell(x)-1}]})\}$$

defines a regular $\lim \inf \Phi = E'$. Essentially the same ideas lead to

THEOREM 3.7 (Schnorr [21]). *R_M is contained in, but not identical to, the intersection of all events whose complements are totally recursive nullsets.*

3.4. Historically the theory of random sequences as reviewed above originated from the theory of complexity of finite sequences (Kolmogorov [8]; Chaitin [3], [4]), although it does not depend on this concept. However, a generalized complexity measure will play a role in our exposition of Section 4. We state Kolmogorov's and Martin-Löf's theorems without proof.

DEFINITION 3.5. *An algorithm is a partial recursive function $A : \Omega \times \Omega \rightsquigarrow \Omega$. The conditional complexity (with respect to A) of x given y $(x, y \in \Omega)$ is*

$$(3.4) \qquad K_A(x|y) := \min \{\ell(p) : A(p, y) = x\}.$$

It is $\leq \infty$, with equality if $A(p, x) \neq x$ for all p, y in the domain of A.

The conditional complexity $K_A(x|y)$ is the minimal length of a program needed to compute x from the "input" y on the "machine" A. In the sequel the second argument of $K_A(x|y)$ will always be (the binary expansion of) $\ell(x)$.

Since K_A depends on the choice of A and moreover assumes infinite values in general, the following theorem is of basic importance.

DEFINITION 3.6. *An algorithm U is called universal if for every algorithm A there is a constant c (depending only on A and U) such that*

$$(3.5) \qquad\qquad K_U(x|y) \leqq K_A(x|y) + c, \qquad\qquad x, y \in \Omega.$$

Obviously $K_U(x|\ell(x)) \leqq \ell(x) + c$ for some constant c.

THEOREM 3.8 (Kolmogorov [8]). *There exist universal algorithms.*

In the sequel, $K(x|y)$ will always denote $K_U(x|y)$ for some universal algorithm U, fixed once and forever. The following simple lemma is fundamental.

LEMMA 3.1. *For every $c, n \in N$*

$$(3.6) \qquad \text{card } \{x: \ell(x) = n; K(x|n) \geqq n - c\} \geqq (1 - 2^{-c})2^n.$$

Let us summarize the asymptotic theory of K.

PROPOSITION 3.1. *For $x \in \mathscr{X}$ $\sup_n K(x_{n]}|n) < \infty$ if and only if x is recursive.*

A proof of the "only if" part can be found in [9].

The following two results concerning large values of K are due to Martin-Löf [11], [13]. The first theorem reveals the surprising fact that $K(x_{n]}|n)$ cannot stay $\geqq n - c$ (c some constant) as $n \to \infty$, for any $x \in \mathscr{X}$.

THEOREM 3.9. (i) *If $F(n)$ is a recursive function having $\Sigma \, 2^{-F(n)} = \infty$, then for every $x \in \mathscr{X}$, $K(x_{n]}|n) < n - F(n)$ for infinitely many n.*

(ii) *If $F(n)$ is a recursive function such that $\Sigma \, 2^{-F(n)}$ is recursively convergent then for every $x \in R_M$, $K(x_{n]}|n) \geqq n - F(n)$ for all but finitely many n.*

THEOREM 3.10. (i) *If there is a constant c such that $K(x_{n]}|n) \geqq n - c$ for infinitely many n, then $x \in R_M$.*

(ii) *Case* (i) *occurs with probability* 1.

4. Randomness and the class Δ_2 of the arithmetical hierarchy

4.1. As indicated in the preceding chapter gambling and sequential testing presuppose effective descriptions of the complement of Φ, where lim inf Φ is the corresponding class of random sequences; moreover, at least the gambling approach implies that both Φ and Φ^c should be recursively enumerable. In contrast to this, Section 1 rather suggests considering those lim inf Φ for which Φ is recursively enumerable.

DEFINITION 4.1. *A trace class is a recursively enumerable subset of Ω.*

DEFINITION 4.2. *An event $E \subset \mathscr{X}$ is called observable if there is a trace class Φ such that $E = $ lim inf Φ.*

DEFINITION 4.3. *A sequence $x \in \mathscr{X}$ is random provided x is an element of every observable event of P measure 1. The set of all random sequences will be denoted by R.*

The author conjectures that R is not an observable event.

The following theorem is a consequence of Theorem 4.2 (which is used in the proof).

THEOREM 4.1. *R_M is an observable event.*

PROOF. Let U be a universal test (Definition 3.4); hence $R_M^c = \bigcap_{i \geq 1} (\sup U_i)$. There is a recursive function $F: N \twoheadrightarrow N \times \Omega$, with range U; the function F is called an enumeration of U. Using F we construct a recursive function $G : N \twoheadrightarrow U$ with the property: for every $n \in N$ and $x \in \mathscr{X}$, $\{i : (n, x_{i]}) \in G(N)\}$ is finite. This is simply achieved by omitting any pair (n, y) provided some (n, z), $z \ll y$, has been enumerated before. Thus $R_M = \lim \inf G(N)^c$. But $G(N)^c \in \Pi_1 \subset \Pi_2$. After Theorem 4.2 there is a trace class such that $\lim \inf \Phi = \lim \inf G(N)^c$.

As follows from Corollary 3.1 and Theorem 3.3, R_M is not regular.

4.2. In this section we summarize the elementary part of the theory of the arithmetical hierarchy which we apply later (a detailed presentation can be found for instance in [20]). The basic reference set will be N which, without further comment, will be identified with its recursive equivalents such as Ω and N^k, $k \in N$. The symbol Σ_0 represents the class of all recursive relations $(:=$ recursive subsets of N^k for some k). A relation is said to belong to the arithmetical hierarchy if it can be obtained from some $S \in \Sigma_0$ by a finite number of complementations and projections; that is, if it is of the form $(Q_1 x_1) \cdots (Q_{k-1} x_{k-1})$ $T(x_1 \cdots x_k)$ where the Q_i are quantifiers and $T \in \Sigma_0$. The relations of the form $(\exists x_1) \cdots T(x_1 \cdots x_k)$ $\big($respectively $(\forall x_1) \cdots T(x_1 \cdots x_k)\big)$, which have $n - 1$ alterations in the prefix (that is, $n - 1 = \mathrm{card}\{\ell : Q_{\ell+1} \neq Q_\ell)$ form the class Σ_n (respectively Π_n), $n > 0$. Thus Σ_1 are the recursively enumerable relations, Π_1 their complements, Σ_2 projections of Π_1 relations, and so on. The hierarchy theorem says: $\Sigma_n - \Pi_n \neq \varnothing$, $n > 0$. The symbol Δ_n denotes $\Sigma_n \cap \Pi_n$. Thus $\Delta_1 = \Sigma_0$. We are particularly interested in Δ_2. The importance of Σ_2, Π_2 and Δ_2 results from the following propositions (see Putnam [18]).

PROPOSITION 4.1. *The following statements are equivalent:*

  (i) *$E \in \Sigma_2$;*

  (ii) *there is a recursive function $F(n)$ such that*

$$(4.1) \qquad E = \big\{n : n = F(m) \text{ for an odd number of } m\big\};$$

  (iii) *there is a recursive sequence of finite sets E_n $\big($that is, a recursive set $S \subset N \times N$ such that $E_n = \{i : (n, i) \in S\}\big)$ with the property $I_E = \lim \inf_{n \to \infty} I_{E_n}$.*

PROOF. First we prove (i) $\Rightarrow$ (ii). According to the assumption there is a recursively enumerable set $T \subset N \times N$ such that

$$(4.2) \qquad E = \mathrm{proj}_2 T^c := \big\{n : (\exists j)(j, n) \in T^c\big\}.$$

The set T^c can be described by a recursive function $G : N \twoheadrightarrow N \times N$, the multiplicity card $G^{-1} G(n)$ of each $G(n)$ being ≤ 2, such that $T^c = \big\{G(n) : \mathrm{card}\, G^{-1} G(n) = 1\big\}$. This $G(n)$ "enumerates" T^c if one interprets the second

occurrence of an element in the range of G as cancellation. Adopting this interpretation for the time being we write

$$(4.3) \qquad \text{card}_n^G (A) :=$$

$$\text{card} \{ G(j) : G(j) \in A, \text{ card } [G^{-1} G(j) \cap \{1, \cdots, n\}] \equiv 1 \bmod 2 \}.$$

The function $F(n)$ is defined inductively: let $k_1 < k_2 < \cdots$ be the arguments for which

$$(4.4) \qquad \min_{t = k_j - 1, k_j} \text{card}_t^G \left(\text{proj}_2^{-1} \left(\text{proj}_2 \, G(k_j) \right) \right) = 0$$

(that is, those "times" at which $\text{proj}_2 \, G(k_j)$ appears as an element of E or is cancelled from E). Put $F(n) := \text{proj}_2 \, G(k_n)$.

Now we prove (ii) $\Rightarrow$ (iii). Let $E_n := \{F(j) : \text{card}_n^F (F(j)) = 1\}$, $n \in N$. This is a recursive sequence of finite sets having $\liminf_{n \to \infty} I_{E_n} = I_E$.

For (iii) inplies (i): $E = \{k : (\exists n)(\forall m)(m \geqq n \Rightarrow k \in E_m)\} \in \Sigma_2$.

COROLLARY 4.1. *The following statements are equivalent:*

(i) $E \in \Delta_2$;

(ii) *there is a recursive function $F(n)$ such that*

$$(4.5) \qquad E = \{n : n = F(m) \text{ for an odd number of } m\},$$

$$(4.6) \qquad \text{for every } m \text{ card } \{j : F(j) = F(m)\} \text{ is finite;}$$

(iii) *there is a recursive sequence of finite sets E_n such that $I_E = \lim_{n \to \infty} I_{E_n}$.*

PROOF. We prove first (i) $\Rightarrow$ (ii): $E \in \Sigma_2 \cap \Pi_2$ means that $E \in \Sigma_2$ and $E^c \in \Sigma_2$. Hence, according to the preceding proposition, there is a recursive function $G(n)$ such that $E = \{n : n = G(m) \text{ for an odd number of } m\}$ and a recursive function $G'(n)$ such that $E^c = \{n : n = G'(m) \text{ for an odd number of } m\}$. Let H be a recursive function that enumerates each n, $t + 1$ times whenever G' enumerates it t times. Then $E^c = \{n : n = H(m) \text{ for an even (positive) number of } m\}$. Now F may be chosen as a modification of G, only the multiplicity being restricted: let $\varphi(n)$ be the recursive function defined inductively by: $\varphi(1) := 1$,

$$(4.7) \qquad \varphi(n)$$

$$:= \min \{j : j \notin \{\varphi(1), \cdots, \varphi(n-1)\} \text{ and } \text{mult}_n^G (G(j)) \leqq \text{mult}_n^A (G(j))\},$$

$n > 1$; here $\text{mult}_n^G (m) := \text{card} \{k : k \leqq n, G(k) = m\}$. Put $F(n) := G(\varphi(n))$. Then $\text{mult}_\infty^F (n) \equiv \text{mult}_\infty^G (n) \bmod 2$ if $n \in E$ and $\equiv 0 \bmod 2$ if $n \in E^c$.

For (ii) $\Rightarrow$ (iii) E_n is defined as above.

For (iii) $\Rightarrow$ (ii) $E^c \in \Sigma_2$ since (ii) is symmetric in E and E^c.

Condition (ii) shows that sets of class Δ_2, although not necessarily being recursively enumerable, can still be generated by a computing machine, if cancellations (at most finitely many per element) are allowed. This motivates

POSTULATE Δ_2. *Sequences of class Δ_2 are not random.*

4.3. The importance of the class Δ_2 for our theory of sequential observation results from the following fact.

THEOREM 4.2.   *The set E is an observable event (Definition 4.2) if and only if there is a subset Φ of Ω, of class Π_2, such that $E = \lim \inf \Phi$.*

These Φ will be called "Π_2 trace classes" occasionally.

PROOF.   Because of the inclusion $\Sigma_1 \subset \Pi_2$ it suffices to show that for $\Phi \in \Pi_2$ there is a $\Psi \in \Sigma_1$ such that $\lim \inf \Phi = \lim \inf \Psi$. After Proposition 4.1 (ii) one can find a recursive function $F(n)$ such that

$$(4.8) \qquad \Phi = \{x : \text{card } \{n : F(n) = x\} \equiv 1 \bmod 2 \text{ or } = \infty\}.$$

For $x \in \Omega$, $n \in N$, let $B(n, x)$ be the finite set of minimal (with respect to $\ll$) $y \in \Omega$, $y \gg x$, $y \neq F(i)$, $i \leq n$, $y \notin B(i, z) - \{x\}$ for all $i < n$ and $z \leq x$. The $B(n, x)$ which are $\neq \{x\}$ are pairwise disjoint. We now start an inductive definition of $\Psi = \{x^1, x^2, \cdots\}$ by setting $x^1 := F(1)$. Suppose that $x^1, \cdots, x^k$ have been defined in the n first steps of this procedure, then the $(n + 1)$st step is described as follows:

> suppose $\text{mult}_{n+1}^F \big(F(n + 1)\big) \equiv 1 \bmod 2$; let $x^{k+1}, \cdots, x^{k+r}$ be the $F(i) \in \bigcup_{j \leq n} B\big(j, F(n + 1)\big)$, $i \leq n$; if $F(n + 1) \notin B\big(i, F(i)\big)$ for all those $i \leq n$ for which $\text{mult}_n^F \big(F(i)\big) = \text{mult}_i^F \big(F(i)\big) \equiv 0 \bmod 2$ then $x^{k+r+1} := F(n + 1)$. Otherwise proceed to the next step.

COROLLARY 4.2.   *For every subset Φ of Ω, of class Σ_2, $\lim \inf \Phi^c$ is an observable event; in particular: for every sequence $x \in \mathscr{X}$ of class Δ_2 there is an observable event of P probability 1 not containing it.*

COROLLARY 4.3.   *R does not contain sequences of class Δ_2.*

Corollary 4.3 can be considerably sharpened (see Corollary 4.5) by means of a generalized complexity theory. Before, let us show that R_M does not satisfy Postulate Δ_2. Thus, some random sequences in the sense of Martin-Löf can be generated by a computer.

THEOREM 4.3.   *There is an $x \in R_M$ of class Δ_2.*

This can be derived from the more general

THEOREM 4.4.   *To every P nullset of the form $E = \lim \sup \Phi$, $\Phi \subset \Omega$ recursively enumerable, there is an $x \in (\lim \sup \Phi)^c$ of class Δ_2.*

PROOF.   One can assume that $(\sup \Phi)^c \neq \varnothing$, which otherwise is obtained by omitting finitely many $x \in \Phi$. Furthermore, card $\Phi = \infty$ without restriction of generality. Let $F(n)$ be a recursive function enumerating Φ. The following describes the construction of a sequence $x \in E^c$ of class Δ_2. First step: let $x^1, x^2, \cdots, x^{\ell(F(1))}$ be the successive initial segments of the lexicographically lowest $x^{\ell(F(1))} \in \{0, 1\}^{\ell(F(1))} - \{F(1)\}$. For the nth step, with $x^1, x^2, \cdots, x^m$ already defined in the first $n - 1$ steps, let $x^{m+1}, \cdots, x^{m + \max\{\ell(F(i)):i \leq n\}}$ be the successive initial segments of the lexicographically lowest $z = x^{m + \max\{\ell(F(i)):i \leq n\}} \in \{0, 1\}^{\max\{\ell(F(i))\}}$ satisfying $F(1) \not< z, \cdots, F(n) \not< z$. Let $(\hat{x}^m : m = 1, 2, \cdots)$ be the sequence $(x^m : m = 1, 2, \cdots)$ without repetitions. Then the following is true:

(i) put $E_n := \{\hat{x}^n_{1]}, \cdots, \hat{x}^n_{n]}\}$, $n \in N$; there is a (unique) $x \in \mathscr{X}$ such that $I_{\{x_{1]}, x_{2]}, \cdots\}} = \lim_{n \to \infty} I_{E_n}$;

(ii) $x \notin E$.

For the proof introduce

$$(4.9) \qquad y^i : = \max_{\ll} \{x^j : x^j \text{ occurs in the } i\text{th step and sup } \{x^j\} \cap \text{sup } \Phi = \varnothing\};$$

one can assume that this maximum exists, otherwise nothing has to be proved. We have $y^i \ll y^{i+1}$, $i \in N$, since $y^{i+1}_{\ell(y^i)]} \nleq y^i$ would be contrary to construction. Because of $(\text{sup } \Phi)^c \neq \varnothing$, $\ell(y^i)$ cannot be bounded as $i \to \infty$ whence $x_{\ell(y^i)]} := y^i$, $i \in N$, defines an infinite sequence x.

Clearly x satisfies (i), and no other sequence does. Furthermore, if (ii) would be false then for some $F(n)$ there would be a $y^i \gg F(n)$, contrary to definition. Since the E_n form a recursive sequence, (i) is equivalent to $x \in \Delta_2$ (Corollary 4.2).

COROLLARY 4.4.   *To every P nullset of the form* lim sup Φ ($\Phi \subset \Omega$ *recursively enumerable*), *there is a nullset* lim sup Ψ ($\Psi \subset \Omega$ *recursively enumerable*) *such that* lim sup Φ *is properly contained in* lim sup Ψ.

PROOF.   This follows from the equation

$$(4.10) \qquad \text{lim sup } (\Phi \cup \{\hat{x}^m : m \in N\}) = \text{lim sup } \Phi \cup \{x\},$$

which is immediate from the preceding proof (notations are as above).

It would be desirable to have a theorem of this nature for $\Phi, \Psi \in \Pi_1$ (instead of Σ_1).

NOTE.   In view of Proposition 2.1 (iv), we have the following characterization of observable events.

THEOREM 4.5.   *The set $E \subset \mathcal{X}$ is an observable event if and only if there exists a function $\Xi : \Omega \twoheadrightarrow \{0, 1, 2\}$ of class Δ_2 (that is, its graph $\{(x, \Xi(x)) : x \in \Omega\}$ belongs to Δ_2) such that $E = \{x \in \mathcal{X} : n \twoheadrightarrow \Xi(x_n)\}$ is a recursive function}.*

The proof follows the lines of our proof of Proposition 2.1 and moreover uses Theorem 4.2.

4.4.   The following presents a generalized complexity theory of the class Δ_2. A Δ_2 complexity measure should be bounded on the initial segments of any infinite sequence $x \in \Delta_2$ (see Proposition 3.1). The asymptotic theory of such a measure is analogous to that of Kolmogorov's complexity measure (see Section 3.4).

DEFINITION 4.4.   *A Σ_2 algorithm A is a function with domain in $\Omega \times \Omega$ and range in Ω having a graph of class Σ_2. The conditional Δ_2 complexity (with respect to A) of x given y, with $x, y \in \Omega$, is*

$$(4.11) \qquad K_A^2(x|y) := \min \{\ell(p) : A(p, y) = x\}, \qquad \leqq \infty.$$

The function A can be interpreted as a machine computing approximations of increasing accuracy, $K_A^2(x|y)$ as the length of the shortest program for which the corresponding procedure used to compute x from y converges.

DEFINITION 4.5.   *A Σ_2 algorithm U^2 is called universal if for every Σ_2 algorithm A there is a constant c (depending only on A and U^2) such that*

$$(4.12) \qquad K_{U^2}^2(x|y) \leqq K_A^2(x|y) + c, \qquad x, y \in \Omega.$$

THEOREM 4.6.   *There exist universal Σ_2 algorithms.*

The proof uses a Σ_2 enumeration of all Σ_2 algorithms (the program p representing, in part, the Gödel number) and follows the lines of Kolmogorov's proof in [8]. Let K^2 be the Δ_2 complexity with respect to some universal Σ_2 algorithm U^2, fixed once and forever.

PROPOSITION 4.2. *There is a constant c such that*

$$(4.13) \qquad K^2\big(x\,\big|\,\ell(x)\big) \leqq K\big(x\,\big|\,\ell(x)\big) + c, \qquad\qquad x \in \Omega.$$

The functions K and K^2 essentially differ by their degree of computability: $\{(x, y, \ell) : K(x\,|\,y) \leqq \ell\}$ being recursively enumerable whereas $\{(x, y, \ell) : K^2(x\,|\,y) \leqq \ell\} \in \Sigma_2 - \Sigma_1$.

LEMMA 4.1. *For every* $c, n \in N$,

$$(4.14) \qquad \text{card } \{x : \ell(x) = n, K^2(x\,|\,n) \geqq n - c\} \geqq (1 - 2^{-c})2^n.$$

PROPOSITION 4.3. *The complexity* $K^2(x_{n]}\,|\,n)$ *stays bounded on every infinite sequence* $x \in \Delta_2$ *as* $n \to \infty$.

PROOF. The relation $A(p, n) := x_{n]}$, $p \in \Omega$, $n \in N$, defines a Σ_2 algorithm.

THEOREM 4.7. *Let* $F(n)$ *be a recursive function such that* $\Sigma 2^{-F(n)} = \infty$. *Then for every* $x \in \mathscr{X}$, $K^2(x_{n]}\,|\,n) < n - F(n)$ *for infinitely many* n.

PROOF. This is an immediate consequence of Theorem 3.9 (i) and Proposition 4.2.

As in Section 3 one gets a converse for random sequences x (now in the sense of Definition 4.3).

THEOREM 4.8. *Let* $F(n)$ *be a recursive function such that* $\Sigma 2^{-F(n)} < \infty$. *Then for every* $x \in R$, $K^2(x_{n]}\,|\,n) \geqq n - F(n)$ *for all but finitely many* n.

PROOF. For every n_0, we have

$$(4.15) \qquad P\{x \in \mathscr{X} : K^2(x_{n]}\,|\,n) < n - F(n) \text{ for infinitely many } n\}$$
$$\leqq \sum_{n > n_0} P\{x \in \mathscr{X} : K^2(x_{n]}\,|\,n) < n - F(n)\}$$
$$\leqq \sum_{n > n_0} 2^{-F(n)},$$

whence

$$(4.16) \qquad P\{x \in \mathscr{X} : K^2(x_{n]}\,|\,n) \geqq n - F(n) \text{ for all but finitely many } n\} = 1.$$

The set $\Phi := \big\{y \in \Omega : K^2\big(y\,\big|\,\ell(y)\big) \geqq \ell(y) - F\big(\ell(y)\big)\big\}$ is a Π_2 trace class for this event; $x \in \liminf \Phi$, since $x \in R$.

COROLLARY 4.5. *There is an observable event of* P *measure 1 containing no sequence of class* Δ_2.

PROOF. Put $F(n) := [(1 + \varepsilon)\,^2\log n]$, $\varepsilon > 0$, in the previous theorem and define the event by the corresponding Π_2 trace class Φ. The statement then follows from Proposition 4.3.

None of these conditions seems to imply randomness (see our conjecture). However, there is an interesting connection between randomness and Δ_2 complexity.

THEOREM 4.9. *If there is a c such that for infinitely many* n, $n - K^2(x_{n]}\,|\,n) \leqq c$ *then* $x \in R$; *the set of these* x *has* P *measure* 1.

PROOF. We have to show that $x \in \lim \inf \Phi$ for every trace class Φ (such that $P(\lim \inf \Phi) = 1$). Since $\Phi^c \in \Delta_2$ there is a Σ_2 algorithm A_Φ which, given n, successively enumerates the class of $x_1 \cdots x_n$ for which $x_1 \notin \Phi$, $x_1 x_2 \notin \Phi$, $\cdots$, $x_1 \cdots x_n \notin \Phi$, then the class of $x_1 \cdots x_n$ for which exactly $n - 1$ initial segments are $\notin \Phi$, and so on, lexicographically within each such class.

For $y \in \Omega$, $\ell(y) = n$, let $Z(y)$ be the number of initial segments of y in Φ^c. Then

$$(4.17) \qquad K_{A_\Phi}^2(y\,|\,n) \leqq {}^2\log \text{card } \{z : \ell(z) = n \quad \text{and} \quad Z(z) \geqq Z(y)\} + 1,$$

such that if a function $F(m, n)$ is defined by

$$(4.18) \qquad {}^2\log \text{card } \{z : \ell(z) = n, Z(z) \geqq m\} + 1 = n - F(m, n), \qquad m \leqq n,$$

then

$$(4.19) \qquad K_{A_\Phi}^2(y\,|\,n) \leqq n - F(Z(y), n).$$

Hence,

$$(4.20) \qquad K^2(y\,|\,n) \leqq n - F(Z(y), n) + c_0$$

for some constant c_0 depending on the choice of the particular enumeration of Φ. Consequently, after assumption,

$$(4.21) \qquad F(Z(x_{n]}), n) \leqq n - K^2(x_{n]}\,|\,n) + c_0 \leqq c + c_0$$

for infinitely many n. This means that there is an $\varepsilon > 0$ such that for infinitely many n

$$(4.22) \qquad P\{y \in \mathcal{X} : Z(y_{n]}) \geqq Z(x_{n]})\} \geqq \varepsilon.$$

Now it is easy to deduce a contradiction from the assumption that $x \notin \lim \inf \Phi$, which also can be stated in the form $Z(x_{n]}) \to \infty$ as $n \to \infty$:

$$(4.23) \qquad P\{y \in \mathcal{X} : Z(y_{n]}) \geqq M\} \leqq P\{y \in \mathcal{X} : \lim_{n \to \infty} Z(y_{n]}) \geqq M\} < \varepsilon,$$

for some sufficiently large M, x being element of the first event for all but finitely many n. Together with (4.22) this yields $Z(x_{n]}) < M$ for infinitely many n, contrary to the above assumption. Hence $x \in \lim \inf \Phi$. The second part of the theorem is a simple consequence of the lemma.

5. Observables related to almost sure convergence

In this section we consider observable events which occur with high probability, but in general not with probability 1. Obviously our systematic framework does not provide reason for studying any particular event of this type. However, there are observables traditionally attracting the interest of probabilists and

statisticians. Consequently, the question of how to determine the probability of related observable events deserves some interest. In the sequel we restrict our attention to those observables which are naturally related to almost sure convergence theorems like the strong law of large numbers and the Glivenko-Cantelli theorem. The appendix will briefly outline a "large deviation" approach due to V. Strassen concerning observables not belonging to this class.

The observables considered now are those having a regular range whenever a particular almost sure convergence statement holds. For example, in the case of the strong law, we study observables $\lim^* \Phi(x_{n]})$ (Φ a trace function), whose domain contains all $x = x_1 x_2 \cdots$ for which $\lim_{n \to \infty} \Sigma_{i \leq n} x_i/n = Ex_1$, and the corresponding events of the form $\{x : \lim^* \Phi(x_{n]}) \leq \alpha\}$. We are going to evaluate probabilities of such events by means of certain distribution invariance principles. This idea is suggested by the following proposition: let $X_1, X_2, \cdots$ be a sequence of random variables with values in a separable metric vector space $\mathscr{E}$ (the metric being $\rho_\mathscr{E}$); then $\lim_{n \to \infty} X_n = 0$ almost surely if and only if the distribution of the (infinite dimensional) vector $(X_n, X_{n+1}, \cdots)$ converges to the unit mass at $(0, 0, \cdots) \in \mathscr{E}^N$ pointwise on the space of bounded continuous measurable functions on $\mathscr{E}^N$ with respect to uniform topology. From the following reformulation it becomes apparent in which way this proposition can be sharpened: to each vector $x = (x_1, x_2 \cdots) \in \mathscr{E}^N$ let $t \rightsquigarrow x(t)$ be its linear interpolation defined by

$$(5.1) \qquad x(t) = \big(1 - \big(t - [t]\big)\big) x_{[t]} + \big(t - [t]\big) x_{[t]+1}, \qquad t \geq 1.$$

Let $\mathscr{C}_\mathscr{E}[1, \infty)$ be the space of all bounded continuous functions with values in $\mathscr{E}$ and domain $[1, \infty)$, endowed with the uniform topology. Then $\lim_{n \to \infty} X_n = 0$ almost surely if and only if the distribution of $t \rightsquigarrow X(tn)$ converges to the unit mass at $0 \in \mathscr{C}_\mathscr{E}[1, \infty)$ pointwise on all bounded continuous measurable functions on $\mathscr{C}_\mathscr{E}[1, \infty)$, as $n \to \infty$.

Once almost sure convergence is known one can restrict oneself to the subspace $\mathscr{C}_\mathscr{E}^0[1, \infty)$ of $\mathscr{C}_\mathscr{E}[1, \infty)$ of those x which satisfy $\lim_{t \to \infty} x(t) = 0$. On this (separable) space the uniform topology can be described by the metric ρ

$$(5.2) \qquad \rho(x, y) := \sup_{t \geq 1} \rho_\mathscr{E}\big(x(t), y(t)\big), \qquad x, y \in \mathscr{C}_\mathscr{E}^0[1, \infty).$$

The above proposition suggests to look for the limit points, if any, of the distribution of $t \rightsquigarrow N(n)X(tn)$ as $n \to \infty$, where $\{N(n) : n \in N\}$ is a sequence of norming constants, $N(n) \to \infty$. In all interesting cases such norming constants will turn out to exist (that is, yield nondegenerate limit points) and, moreover, there will be convergence to a limit distribution.

5.1.   Let $\mathscr{C}_*[0, \infty)$ be the (separable) Banach space of continuous functions x on $[0, \infty)$ such that (i) $x(0) = 0$, (ii) $\lim_{t \to \infty} x(t)/t = 0$, endowed with the norm

$$(5.3) \qquad \|x\|_* := \sup_{t \geq 0} \frac{|x(t)|}{t \vee 1}.$$

The space $\mathscr{C}_*[0, \infty)$ will play the role of the path space of a process ξ obtained from a sequence of partial sums of independent random variables by linear interpolation; more precisely: let $X_1, X_2, \cdots$ be a sequence of independent random variables having $EX_i = 0$ and $0 < EX_i^2 < \infty$, $i \in N$, satisfying Lindeberg's condition and the strong law of large numbers in the form

$$(5.4) \qquad P\left[\lim_{n \to \infty} \frac{1}{s_n^2} \sum_{i \le n} X_i = 0\right] = 1, \qquad s_n^2 := \sum_{i \le n} EX_i^2.$$

We consider stochastic processes $t \rightsquigarrow \xi(t)$ with the properties: (i) $\xi(0) = 0$, (ii) $\xi(s_n^2) = \Sigma_{i \le n} X_i$, (iii) $t \rightsquigarrow \xi(t)$ is monotone and continuous in each of the intervals $[s_{n-1}^2, s_n^2]$, $n \in N$.

THEOREM 5.1 (Müller [15]). *With probability* 1, $\xi \in \mathscr{C}_*[0, \infty)$; *the sequence of distributions of* $t \rightsquigarrow \xi(s_n^2 t)/s_n$, $0 \le t < \infty$, *converges (pointwise on bounded continuous functions on* $\mathscr{C}_*[0, \infty)$) *to the distribution of a Wiener process* $t \rightsquigarrow \zeta(t)$, $0 \le t < \infty$, $\zeta(0) = 0$.

Before proceeding to the applications of this theorem let us mention an estimate of the speed of convergence, extending a result of Skorohod [22], [23].

Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables having $EX_1 = 0$, $EX_1^2 = 1$ and $\sup |X_1| = C < \infty$. Let $\xi(k)$ be the kth partial sum of this sequence as above. We are interested in comparing the probabilities of the following events for $\alpha \ge 0$:

$$(5.5) \quad \begin{aligned} \Xi_n &:= \left[g_1\left(\frac{k}{n}\right) < n^{-1/2}\xi(k) < g_2\left(\frac{k}{n}\right) \quad \text{for all} \quad k \ge \alpha n\right], \quad n \in N, \\ Z &= \left[g_1(t) < \zeta(t) < g_2(t) \quad \text{for} \quad \text{all} \quad t \ge \alpha\right]; \end{aligned}$$

here g_i are continuous functions on $[0, \infty)$ having continuous derivatives subject to the following conditions:

$$g_1(0) < 0 < g_2(0),$$

$$-\infty < \lim_{t \to \infty} \frac{g_1(t)}{t} < 0 < \lim_{t \to \infty} \frac{g_2(t)}{t} < +\infty,$$

$$(5.6)$$

$$\limsup_{t \to \infty} \left|g_i(t) - t\frac{d}{dt}g_i(t)\right| < \infty, \qquad i = 1, 2.$$

Only the following six bounds for the above quantities (and C) will enter the error term:

$$0 < G_1 \le |g_i(0)| \le G_2,$$

$$0 < G_3 \le \lim_{t \to \infty} \left|\frac{g_i(t)}{t}\right| \le G_4,$$

$$(5.7)$$

$$\sup_{0 \le t \le 1} \left|\frac{d}{dt}g_i(t)\right| \le K_1,$$

$$\sup_{t \ge 1} \left|g_i(t) - t\frac{d}{dt}g_i(t)\right| \le K_2, \qquad i = 1, 2.$$

THEOREM 5.2 (Müller [15]). *There is a constant L, depending only on C,* $K_1, K_2, G_1, \cdots . G_4$ *such that*

$$(5.8) \qquad \qquad |P\,\Xi_n \,-\, PZ| \leq L\,\frac{\log n}{\sqrt{n}}, \qquad\qquad n \geq 2.$$

The theorem remains true if one puts $g_1(t) = -\infty$, $t \geq 0$, disregarding the above assumptions about this function.

The following corollaries are true under the assumptions of Theorem 5.1 and Theorem 5.2 respectively. Moreover, in order to simplify the formulas, it will be assumed that $EX_i = 1$, $i \in N$.

COROLLARY 5.1.

$$(5.9) \qquad \lim_{\varepsilon \downarrow 0} P\left[\varepsilon^2 \max\left\{n : \frac{1}{n}\sum_{i=1}^{n} X_i \geq \varepsilon\right\} < \alpha\right] = \int_0^\alpha (2\pi u)^{-1/2}\, e^{-u^2/2}\, du$$

uniformly in α. There is a constant L depending only on C and $\alpha_1 > 0$ such that for all $\alpha \geq \alpha_1$, we have

$$(5.10) \qquad \left|P\left[\varepsilon^2 \max\left\{n : \frac{1}{n}\sum_{i=1}^{n} X_i \geq \varepsilon\right\} < \alpha\right] - \int_0^\alpha (2\pi u)^{-1/2}\, e^{-u^2/2}\, du\right|$$

$$\leq\, -\, L\,\varepsilon \log \varepsilon.$$

COROLLARY 5.2.

$$(5.11) \qquad \lim_{\varepsilon \downarrow 0} P\left[\varepsilon^2 \max\left\{n : \frac{1}{n}\left|\sum_{i=1}^{n} X_i\right| \geq \varepsilon\right\} < \alpha\right]$$

$$= \frac{4}{\pi}\sum_{n=0}^\infty \frac{(-1)^n}{2n+1} \exp\left\{-\frac{(2n+1)^2\pi^2}{8\alpha}\right\}$$

uniformly in α. There is a constant L depending only on C and $\alpha_1 > 0$ such that for $\alpha \geq \alpha_1$, we have

$$(5.12) \qquad \left|P\left[\varepsilon^2 \max\left\{n : \frac{1}{n}\left|\sum_{i=1}^{n} X_i\right| \geq \varepsilon\right\} < \alpha\right] - \frac{4}{\pi}\sum_{n=0}^\infty \frac{(-1)^n}{2n+1}\right.$$

$$\left.\exp\left\{-\frac{(2n+1)^2\pi^2}{8\alpha}\right\}\right| \leq\, -\, L\varepsilon \log \varepsilon.$$

COROLLARY 5.3.

$$(5.13) \qquad \lim_{n \to \infty} P\left[\sqrt{n} \max_{k \geq n}\left(\frac{1}{k}\sum_{i=1}^{k} X_i\right) < \alpha\right] = \left(\frac{2}{\pi}\right)^{1/2} \int_0^\alpha e^{-u^2/2}\, du$$

uniformly in α. There is a constant L depending only on C, α_1, and α_2 such that for $0 < \alpha_1 \leq \alpha \leq \alpha_2$,

$$(5.14) \qquad \left|P\left[\sqrt{n} \max_{k \geq n}\left(\frac{1}{k}\sum_{i=1}^{k} X_i\right) < \alpha\right] - \left(\frac{2}{\pi}\right)^{1/2} \int_0^\alpha e^{-u^2/2}\, du\right| \leq L\,\frac{\log n}{\sqrt{n}}.$$

COROLLARY 5.4.

$$(5.15) \quad \lim_{n \to \infty} P\left[\sqrt{n} \max_{k \geq n} \frac{1}{k} \left| \sum_{i=1}^{k} X_i \right| < \alpha \right]$$

$$= \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left\{ -\frac{(2n+1)^2 \pi^2}{8\alpha^2} \right\}$$

uniformly in α. *There is a constant* L *depending only on* C, α_1, *and* α_2 *such that for* $0 < \alpha_1 \leq \alpha \leq \alpha_2$,

$$(5.16) \quad \left| P\left[\sqrt{n} \max_{k \geq n} \left(\frac{1}{k} \left| \sum_{i=1}^{k} X_i \right| \right) < \alpha \right] - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left\{ -\frac{(2n+1)^2 \pi^2}{8\alpha^2} \right\} \right|$$

$$\leq L \frac{\log n}{\sqrt{n}}.$$

COROLLARY 5.5.   *For* $\alpha \geq \beta \vee 0$, *we have*

$$(5.17) \quad \lim_{n \to \infty} P\left[\sqrt{n} \max_{k \geq n} \left(\frac{1}{k} \sum_{i=1}^{k} X_i \right) > \alpha \, \bigg| \, \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i = \beta \right] = e^{-2\alpha(\alpha - \beta)}$$

uniformly in α. *There is a constant* L *depending only on* C, α_i, β_i, $i = 1, 2$, *such that for* $0 \leq \alpha_1 \leq \alpha < \alpha_2$, $0 < \beta_1 \leq \beta \leq \beta_2$,

$$(5.18) \quad \left| P\left[\sqrt{n} \max_{k \geq n} \left(\frac{1}{k} \sum_{i=1}^{k} X_i \right) > \alpha \, \bigg| \, \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i = \beta \right] - e^{-2\alpha(\alpha - \beta)} \right| \leq L \frac{\log n}{\sqrt{n}}.$$

COROLLARY 5.6 (Müller [17]).   *Let* N_ε^+ *be the number of* n *such that* $\sum_{i=1}^{n} X_i / n \geq \varepsilon$. *Then*

$$(5.19) \qquad \lim_{\varepsilon \downarrow 0} P[\varepsilon^2 N_\varepsilon^+ < \alpha] = \int_0^\alpha f^+(t) \, dt,$$

where

$$(5.20) \qquad \begin{aligned} f^+(t) &= \left(\frac{2}{\pi t} \right)^{1/2} e^{-t/2} - \mathrm{erfc}\left(\sqrt{t} \right), \\ \mathrm{erfc}\,(t) &= \left(\frac{2}{\pi} \right)^{1/2} \int_t^\infty \exp\left\{ \frac{-u^2}{2} \right\} du. \end{aligned}$$

PROOF.   This follows as in the proof of Corollary 5.7.

COROLLARY 5.7.   *Let* N_ε *be the number of* n *such that* $\left| \sum_{i=1}^{n} x_i \right| / n \geq \varepsilon$. *Then*

$$(5.21) \qquad \lim_{\varepsilon \downarrow 0} P[\varepsilon^2 N_\varepsilon < \alpha] = \int_0^\alpha f(t) \, dt,$$

where $f(t)$ is by definition equal to

$$(5.22) \qquad f^+(t) - \sum_{n > 0} \left[(2n + 1)!!\right]^2 (2n + 1)^2 \left[f^+\left[(2n + 1)^2 t\right] + 2\right]^*$$

$$\overset{n}{\underset{j=0}{\Pi^*}} \left[f^+((2j + 1)^2 t) - \delta\right].$$

Here δ denotes the Dirac delta "function" which is only used for convenience of notation: the convolution products which actually occur are functions. This limit distribution will be derived from an identity for Markov processes to be stated next.

PROPOSITION 5.1. *Let ζ be a strong symmetric $\left(\text{that is, } P^a[\zeta \in S] = P^{-a}[-\zeta \in S]\right)$ for all real a and all measurable subsets S of the path space$\big)$ Markov process having continuous paths; moreover we assume that it fulfills the additional requirement*

$$(5.23) \qquad A := \sup_{a \geq 0} E^a \int_0^\infty I[\zeta(t) > t + \zeta(0)] \, dt < \infty.$$

Denote by $T(t)$ the (random) amount of time $|\zeta(u)|$ spends above $\zeta(0) + u$, $u \leq t$; in other words

$$(5.24) \qquad T(t) = \int_0^t I\big[|\zeta(u)| > \zeta(0) + u\big] \, du.$$

The random time T has the following addition property for $s < t$,

$$(5.25) \qquad T(t) = T(s) + T_s(t - \tau(s) \wedge t) + (\tau(s) \wedge t - s)I[|\zeta(s)| > s + \zeta(0)];$$

here

$$(5.26) \qquad \tau(s) = \inf\{u \geq s : |\zeta(u)| = \zeta(0) + u\}, \qquad\qquad \leq + \infty,$$

and

$$(5.27) \qquad T_s(t) = \int_0^t I\big[|\zeta(\tau(s) + u)| > |\zeta(\tau(s))| + u\big] \, du.$$

Put $L(a, \lambda) = E^a e^{-\lambda T(\infty)}, \big(T(\infty) \text{ being finite because of } (5.23)\big)$. Then the following identity holds for $\alpha \geq 0$:

$$(5.28) \qquad L(a, \lambda) + \lambda \int_0^\infty E^a I\big[|\zeta(s)| > \zeta(0) + s\big] \{e^{-\lambda(\tau(s) - s)} L\big(\tau(s) + a, \lambda\big)\} \, ds = 1.$$

This is a Volterra integral equation of the second kind, having $L(\cdot, \lambda)$ as its bounded solution, $0 < \lambda < 1/2A$.

PROOF. For the proof put $L(a, \kappa, \lambda) = \int_0^\infty e^{-\kappa t} E^a e^{-\lambda T(t)} \, dt$, $\kappa > 0$. Using (5.25) we get the following chain of equalities (the interchange of integrals always being justified by (5.23)):

$$(5.29) \qquad L(a, \kappa, \lambda) - \frac{1}{\kappa} = E^a \int_0^\infty e^{-\kappa t}[e^{-\lambda T(t)} - 1] \, dt$$

$$= -\lambda E^a \int_{t=0}^\infty e^{-\kappa t - \lambda T(t)} \, dt \int_{s=0}^t e^{\lambda T(s)} I[|\zeta(s)| > \zeta(0) + s] \, ds$$

$$= -\lambda \int_{s=0}^\infty ds \, E^a I[|\zeta(s)| > \zeta(0) + s] \int_{t=s}^\infty \exp\{-\kappa t - \lambda[T(t) - T(s)]\} \, dt$$

$$= -\lambda \int_{s=0}^\infty ds \, E^a I[|\zeta(s)| > \zeta(0) + s]$$

$$\times E^a \left\{ \int_{t=s}^\infty \exp\{-\kappa t - \lambda T_s(t - \tau(s) \wedge t) - \lambda[\tau(s) \wedge t - s]\} \, dt \, | \, \tau(s) \right\}$$

$$= -\lambda \int_{s=0}^\infty \int_{t=s}^\infty \exp\{-(\kappa + \lambda)t + \lambda s\} P^a[|\zeta(s)| > \zeta(0) + s \, \& \, \tau(s) > t] \, dt \, ds$$

$$-\lambda \int_{s=0}^\infty \int_{t=s}^\infty e^{-\kappa t} E^a I[|\zeta(s)| > \zeta(0) + s] \, \& \, \tau(s) \leq t]$$

$$\times (\exp\{-\lambda[\tau(s) - s]\} E^a[\exp\{-\lambda T_s(t - \tau(s))\} | \tau(s)]) \, dt \, ds$$

$$= O(1)[\kappa \to 0] - \lambda \int_{s=0}^\infty E^a I[|\zeta(s)| > \zeta(0) + s]$$

$$\times \left(\int_{t=\tau(s)}^\infty \exp\{-\kappa t - \lambda[\tau(s) - s]\} E^a[\exp\{-\lambda T_s(t - \tau(s))\} | \tau(s)] \, dt \right) ds$$

$$= O(1) - \lambda \int_{s=0}^\infty E^a I[|\zeta(s)| > \zeta(0) + s]$$

$$\times \left(\exp\{-(\kappa + \lambda)\tau(s) + \lambda s\} E^a \left[\int_{v=0}^\infty \exp\{-\kappa v - \lambda T_s(v)\} \, dv \, | \, \tau(s) \right] \right) ds$$

$$= O(1) - \lambda \int_{s=0}^\infty E^a I[|\zeta(s)| > \zeta(0) + s]$$

$$\times \left(\exp\{-(\kappa + \lambda)\tau(s) + \lambda s\} E^{\zeta(\tau(s))} \left[\int_{v=0}^\infty \exp\{-\kappa v - \lambda T(v)\} \, dv \right] \right) ds.$$

Now according to a Tauberian theorem (see Feller [5], p. 423), $\lim_{\kappa \to 0} \kappa L(a, \kappa, \lambda) = E^a e^{-\lambda T(\infty)} = L(a, \lambda)$, so that after application of Lebesgue's bounded convergence theorem our assertion will follow.

PROOF OF COROLLARY 5.7.   In the case of a Brownian motion ζ we solve the integral equation (5.28),

$$(5.30) \qquad\qquad L(\cdot, \lambda) + \lambda K_\lambda L(\cdot, \lambda) = 1,$$

where

$$(5.31) \qquad (K_\lambda \Psi)(a) =$$

$$\int_0^\infty ds \, E^a I[|\zeta(s)| > \zeta(0) + s][\exp\{-\lambda[\tau(s) - s]\} \Psi(a + \tau(s))].$$

This may appear to lead to insurmountable difficulties since not even its kernel K_λ can be expressed by elementary functions. The following lemma, however, reveals its unexpected simplicity.

LEMMA 5.1. *We can write*

$$(5.32) \quad K_\lambda(e^{-\mu t})(a) = c(\lambda, \mu)e^{-\mu a} + c(\lambda, \mu) \exp\left\{-[\mu + 2(1 + (1 + 2\mu)^{1/2})]a\right\}$$

with

$$(5.33) \quad c(\lambda, \mu)^{-1} = (1 + 2\mu)^{1/2}[(1 + 2\mu + 2\lambda)^{1/2} + (1 + 2\mu)^{1/2}], \quad \mu \geqq 0.$$

PROOF. Let $\xi(t) = \zeta(t) - t$. First we remark that with $\sigma_a(s) = \inf\{t \geqq s : \xi(t) = a\}$, and for $x \geqq a > 0$ and $\kappa > 0$,

$$(5.34) \quad E^x e^{-\kappa\sigma_a(0)} = \int_0^\infty e^{-\kappa t} P^{x-a}[\xi(t) \leqq 0]\, dt \left[\int_0^\infty e^{-\kappa t} P^0[\xi(t) \leqq 0]\, dt\right]^{-1}$$

(for a proof in a simpler case see Itô and McKean [6], p. 25), which will be used in the following evaluation:

$$(5.35) \quad K_\lambda(e^{-\mu t})(a) = e^{-\mu a} \int_0^\infty E^0 I[\xi(s) > 0][\exp\{-(\lambda + \mu)\sigma_0(s) + \lambda s\}]\, ds$$

$$+ e^{-\mu a} \int_0^\infty E^0 I[\xi(s) > 2a][\exp\{-(\lambda + \mu)\sigma_{2a}(s) + \lambda s\}]\, ds$$

$$= I + II.$$

We confine ourselves to computing II.

The following formula will be used twice (see [6], p. 17):

$$(5.36) \quad E^x \int_0^\infty e^{-\kappa t} f(\xi(t))\, dt = \int_y \frac{\exp\{-(1 + 2\kappa)^{1/2}|y - x|\}}{(1 + 2\kappa)^{1/2}} e^{-(y-x)} f(y)\, dy.$$

Thus,

$$(5.37) \quad II = e^{-\mu a} \int_{s=0}^\infty e^{\lambda s}\, ds \int_{x=2a}^\infty P^0[\xi(s) \in dx]$$

$$\times E^x\left(\exp\{-(\lambda + \mu)(\sigma_{2a}(0) + s)\}\right)$$

$$= e^{-\mu a} \int_{s=0}^\infty e^{-\mu s}\, ds \int_{x=0}^\infty P^0[\xi(s) - 2a \in dx]$$

$$\times E^{x+2a}\left(\exp\{-(\lambda + \mu)\sigma_{2a}(0)\}\right)$$

$$= e^{-\mu a} \int_{s=0}^\infty e^{-\mu s}\, ds \int_{x=0}^\infty P^0[\xi(s) - 2a \in dx]$$

$$\times \int_{t=0}^\infty e^{-(\lambda+\mu)t} P^x[\xi(t) \leqq 0]\, dt \left[\int_0^\infty e^{-(\lambda+\mu)t} P^0[\xi(t) \leqq 0]\, dt\right]^{-1}$$

$$= e^{-\mu a} \int_{x=0}^\infty E^0 \int_{s=0}^\infty e^{-\mu s} I[\xi(s) \in dx + 2a]\, ds$$

$$\times E^x \int_{t=0}^\infty e^{-(\lambda+\mu)t} I[\xi(t) \leqq 0]\, dt \left[E^0 \int_0^\infty e^{-(\lambda+\mu)t} I[\xi(t) \leqq 0]\, dt\right]^{-1}$$

$$= e^{-\mu a} \int_{x=0}^{\infty} \int_{y} \frac{\exp\{-(1 + 2\mu)^{1/2}|y|\}}{(1 + 2\mu)^{1/2}} e^{-y} I[y \in 2a + dx] \, dy$$

$$\times \int_{z=-\infty}^{0} \frac{\exp\{-(1 + 2\lambda + 2\mu)^{1/2}|z - x|\}}{(1 + 2\lambda + 2\mu)^{1/2}} e^{-(z-x)} \, dz$$

$$\times \left[\int_{z=-\infty}^{0} \frac{\exp\{[(1 + 2\lambda + 2\mu)^{1/2} - 1]z\}}{(1 + 2\lambda + 2\mu)^{1/2}} \, dz\right]^{-1}$$

$$= e^{-\mu a} \int_{x=0}^{\infty} \frac{\exp\{[(1 + 2\mu)^{1/2} + 1](2a + x)\}}{(1 + 2\mu)^{1/2}} \, dx$$

$$\times \exp\{-x[(1 + 2\lambda + 2\mu)^{1/2} - 1]\}$$

$$= \frac{\exp\{-\mu a - 2a[1 + (1 + 2\mu)^{1/2}]\}}{(1 + 2\mu)^{1/2}[(1 + 2\lambda + 2\mu)^{1/2} + (1 + 2\mu)^{1/2}]}.$$

LEMMA 5.2. *Let E be the space of all functions of the form $\Sigma_{n \geq 0} a_n z^{2n(n+1)}$ with $z = e^{-t}$ and $\Sigma_{n \geq 0} |a_n| < \infty$. Then $K_\lambda E \subset E$.*

PROOF. It suffices to check that the recurrence relation

(5.38)
$$\mu_{n+1} = \mu_n + 2[1 + (1 + 2\mu_n)^{1/2}],$$
$$\mu_0 = 0$$

has the solution $\mu_n = 2n(n + 1)$, $n \geq 0$.

On E our integral equation (5.28) has the form

(5.39) $$\sum_{n \geq 0} a_n(\lambda) z^{\mu_n} + \lambda \sum_{n \geq 0} a_n(\lambda) c(\lambda, \mu_n)(z^{\mu_n} + z^{\mu_{n+1}}) = 1,$$

such that

(5.40) $$a_0(\lambda) = \frac{1}{1 + \lambda c(\lambda, \mu_0)},$$

(5.41) $$a_n(\lambda) = -\frac{\lambda c(\lambda, \mu_{n-1})}{1 + \lambda c(\lambda, \mu_n)} a_{n-1}(\lambda), \qquad\qquad n > 0.$$

Thus,

(5.42) $$a_n(\lambda) = a_0(\lambda)(-1)^n \lambda^n \prod_{j=1}^{n} \frac{c(\lambda, \mu_{j-1})}{1 + \lambda c(\lambda, \mu_j)}$$

which, by $(1 + 2\mu_n)^{1/2} = 2n + 1$, becomes

(5.43) $$-\prod_{j=0}^{n} \varphi((2j + 1)^{-2}\lambda)\chi((2n + 1)^{-2}\lambda),$$

where

(5.44) $$\varphi(\lambda) = \frac{\lambda}{1 + (1 + 2\lambda)^{1/2} + \lambda}, \qquad \chi(\lambda) = \lambda^{-1}[1 + (1 + 2\lambda)^{1/2}].$$

Thus

$$(5.45) \quad L(a, \lambda)$$

$$= \frac{(1 + 2\lambda)^{1/2} - 1}{\lambda} - \sum_{n>0} \prod_{j=0}^{n} \varphi\big((2j + 1)^{-2}\lambda\big)\chi\big((2n + 1)^{-2}\lambda\big) e^{-2n(n+1)a}$$

is the bounded solution of (5.28).

As to the inversion of the Laplace transform it suffices to observe that

$$(5.46) \qquad\qquad \varphi(\lambda) = \frac{(1 + 2\lambda)^{1/2} - 1}{\lambda} - 1$$

is the transform of $f^{+}(t) - \delta$.

Let us now consider observables related to relative occupation and crossing times. For simplicity we assume $X_1, X_2, \cdots$ to be a sequence of independent identically distributed variables having $EX_1 = 0$ and $EX_1^2 = 1$. Let X_1 be a centered lattice or centered nonlattice variable (see Breiman [2]). We denote by ξ the process of partial sums of the X_i (linearly interpolated) and by v the occupation time of the compact interval $I = [a, b]$, that is, $v(t) = \text{card}\ \{n \in N : n \le t, \xi(n) \in I\}$ (linearly interpolated). The process $(v, |\xi|)$ will be studied in the space $\mathscr{C}_+^0[0, \infty)$ of all continuous functions $x : [0, \infty) \rightsquigarrow [0, \infty)^2$ having $\lim_{t \to \infty} x(t)/t = 0$. A topology on this space will be induced by the norm

$$(5.47) \qquad\qquad \|x\| := \sup_{t \geq 0} \frac{|x(t)|_2}{t \vee 1},$$

where $|\cdot|_2$ denotes the two dimensional Euclidean norm.

The following invariance principle extends a result of Stone [24].

THEOREM 5.3. *With probability* 1, $(v, |\xi|) \in \mathscr{C}_+^0[0, \infty)$; *the sequence of distributions of*

$$(5.48) \qquad\qquad t \rightsquigarrow n^{-1/2}\big(v(nt), |\xi(nt)|\big), \qquad\qquad t \geq 0$$

converges (*pointwise on bounded continuous functions on* $\mathscr{C}_+^0[0, \infty)$) *to the distribution of*

$$(5.49) \qquad\qquad t \rightsquigarrow \big(|I| \max_{0 \le s \le t} \zeta(s), \max_{0 \le s \le t} \zeta(s) - \zeta(t)\big),$$

where $|I|$ *is the number of lattice points in* I *or* $b - a$, *respectively*.

PROOF. The proof will be divided into five parts.

Part 1. The following result due to Kallianpur and Robbins [7] will be used:

$$(5.50) \qquad\qquad \lim_{n \to \infty} P\big[n^{-1/2} v(n) < \alpha\big] = P\big[|I| \, |\zeta(1)| < \alpha\big]$$

uniformly in α; moreover, there is convergence of all moments.

Part 2. The limit distribution stated in the theorem coincides with the distribution of $\big(|I| L(t), |\zeta(t)|\big)$, where L is the local time of ζ at 0 (see [1], [6], [14]). The process $\big(|I| L(t), |\zeta(t)|\big)$ is Markovian (so is $n \rightsquigarrow \big(v(n), \xi(n)\big)$, but in general not $n \rightsquigarrow \big(v(n), |\xi(n)|\big)$).

Part 3. According to Kallianpur and Robbins there is a constant c such that $E\left(n^{-1/2}t^{-1}v(nt)\right) \leqq ct^{-1/2}$, $n \in N$, $t > 0$; and from Kolmogorov's inequality applied to ξ (see [15]) it follows that it suffices to prove the theorem for the processes restricted to $[0, T]$, for every $T > 0$. In the sequel, we carry out the argument for $T = 1$.

Part 4. In a preliminary step we show that coordinatewise convergence holds, that is, that the distribution of $t \rightsquigarrow |I| \max_{0 \leqq s \leqq t} \zeta(s)$ is the limit distribution of $t \rightsquigarrow n^{-1/2}v(nt)$. To this end we first prove an invariance principle about the inverse processes $\alpha \rightsquigarrow n^{-1}v^{-1}(\sqrt{n}\alpha)$. These processes have almost independent increments as follows, in the nonlattice case, for $n\alpha_i \in N$, B open, from the following relations

$$(5.51) \qquad P\left[n^{-1}v^{-1}\left(\sqrt{n}\alpha_1\right) \in B \ \& \ n^{-1}v^{-1}\left(\sqrt{n}(\alpha_1 + \alpha_2)\right) \leqq t_0\right]$$

$$= \int_{t \in B, y \in I} P\left[n^{-1}v^{-1}\left(\sqrt{n}\alpha_2\right) \leqq t_0 \big| \xi(0) = y\right] P\left[n^{-1}v^{-1}\left(\sqrt{n}\alpha_1\right) \in dt\right.$$

$$\left. \& \ \xi(t) \in dy\right]$$

$$= P\left[n^{-1}v^{-1}\left(\sqrt{n}\alpha_2\right) \leqq t_0\right] P\left[n^{-1}v^{-1}\left(\sqrt{n}\alpha_1\right) \in B\right] + R_n,$$

Also

$$(5.52) \qquad R_n \leqq \max_{y \in I} \left| P\left[n^{-1/2}v(nt_0)\right.\right.$$

$$\left.\left. \leqq \alpha_2 \big| \xi(0) = y\right] - P\left[n^{-1/2}v(nt_0) \leqq \alpha_2 \big| \xi(0) = 0\right]\right|.$$

By the following lemma this tends to zero as $n \to \infty$.

LEMMA 5.3. *Let* X_1 *be centered nonlattice. For every* $K > 0$, $\alpha > 0$, *the sequence of functions*

$$(5.53) \qquad y \rightsquigarrow P\left[n^{-1/2}v(n) < \alpha \big| \xi(0) = y\right], \qquad y \in [-K, K]$$

converges uniformly to a constant.

PROOF. For every $\varepsilon > 0$ and $y_0 \in [-K, K]$ there is a neighborhood U of y_0 and a pair of intervals $\underline{I}, \overline{I}$ such that

$$\underline{I} \subset I - y \subset \overline{I}, \qquad\qquad y \in U,$$

$$(5.54)$$

$$\limsup_{n \to \infty} \left| P\left[n^{-1/2}v_{\underline{I}}(n) < \alpha\right] - P\left[n^{-1/2}v_{\overline{I}}(n) < \alpha\right]\right| < \varepsilon,$$

where $v_{\underline{I}}, v_{\overline{I}}$ are the occupation times of $\underline{I}, \overline{I}$, respectively (this is a consequence of the Kallianpur-Robbins theorem). But for $y \in U$, $n \in N$,

$$(5.55) \qquad P\left[n^{-1/2}v_{\overline{I}}(n) < \alpha\right] \leqq P\left[n^{-1/2}v(n) < \alpha \big| \xi(0) = y\right]$$

$$\leqq P\left[n^{-1/2}v_{\underline{I}}(n) < \alpha\right].$$

Convergence of the one dimensional marginals of the inverse processes can be concluded from the following formula: for $\lambda > 0$, $\alpha \geqq 0$,

$$(5.56) \qquad E \int_0^\infty e^{-\lambda t} I\left[n^{-1/2}v(nt) > \alpha\right] dt = E \int_{n^{-1}v^{-1}(\sqrt{n}\alpha)}^\infty e^{-\lambda t} dt$$

$$= \frac{1}{\lambda} E \exp\left\{-\lambda n^{-1} v^{-1}(\sqrt{n}\alpha)\right\},$$

which implies

$$(5.57) \qquad \lim_{n \to \infty} \frac{1}{\lambda} E \exp\left\{-\lambda n^{-1} v^{-1}(\sqrt{n}\alpha)\right\} = \int_0^\infty e^{-\lambda t} P[|I| L(t) \geqq \alpha]\, dt$$

$$= \frac{1}{\lambda} E \exp\left\{-\lambda L^{-1}\left(\frac{\alpha}{|I|}\right)\right\}.$$

We are going to show that the distributions of $\alpha \rightsquigarrow n^{-1} v^{-1}(\sqrt{n}\alpha)$ converge (to the distribution of $\alpha \rightsquigarrow L^{-1}(\alpha/|I|)$) in an appropriate sense. To this end let us introduce the function space $\mathcal{M}[0, \infty)$; this is the space of all functions x such that

(i) $x : [0, \infty) \rightsquigarrow [0, \infty)$,
(ii) $x(t_1) \leqq x(t_2)$ whenever $t_1 \leqq t_2$,
(iii) $x(t) \to \infty$ as $t \to \infty$,
(iv) x is right continuous.

The function space $\mathcal{M}[0, \infty)$ will be endowed with a topology by means of its isomorphic relationship to the space $\check{\mathcal{M}}[0, \infty)$ of "graphs" of functions

$$(5.58) \quad x \in \mathcal{M}[\dot{0}, \infty) - \check{\mathcal{M}}[0, \infty) := \{\bigcup_{t \geqq 0} \{t\} \times [x^-(t), x^+(t)] : x \in \mathcal{M}[0, \infty)\},$$

which carries the family of pseudometrics

$$(5.59) \qquad \rho_n(z_1, z_2) = \mathrm{dist}\left(z_1 \cap [0, n]^2, z_2 \cap [0, n]^2\right), \qquad z_1, z_2 \in \check{\mathcal{M}}[0, \infty);$$

here dist denotes the Euclidean Hausdorff distance of compacts. Let us point out the following properties of $\mathcal{M}[0, \infty)$:

(a) $x \rightsquigarrow x^{-1}$ is a homeomorphism of $\mathcal{M}[0, \infty)$;

(b) diffuse measures with compact support are continuous on $\mathcal{M}[0, \infty)$;

(c) a fundamental family of compacts of $\mathcal{M}[0, \infty)$ is given by the order intervals $[x_1, x_2]$ (with respect to coordinatewise ordering, $x_1, x_2 \in \mathcal{M}[0, \infty)$);

(d) a sequence of stochastic processes $\xi_n \in \mathcal{M}[0, \infty)$ has a tight family of distributions if and only if the following two conditions are satisfied:

$$\lim_{b \to \infty} \limsup_{n \to \infty} P[\xi_n(t) > b] = 0, \qquad\qquad t \geqq 0,$$

$$(5.60)$$

$$\lim_{b \to \infty} \limsup_{n \to \infty} P[\xi_n^{-1}(\alpha) > b] = 0, \qquad\qquad \alpha \geqq 0;$$

The tightness conditions are easily verified for the processes $\alpha \rightsquigarrow n^{-1} v^{-1} (\sqrt{n}\alpha)$ since

$$\lim_{n \to \infty} P[n^{-1/2} v(nt) > b] = P[|I| L(t) > b],$$

$$(5.61)$$

$$\lim_{n \to \infty} P[n^{-1} v^{-1}(\sqrt{n}\alpha) > b] = P\left[L^{-1}\left(\frac{\alpha}{|I|}\right) > b\right].$$

Hence the distributions of $\alpha \leadsto n^{-1}v^{-1}(\sqrt{n}\alpha)$ converge weakly to the distribution of $\alpha \leadsto L^{-1}(\alpha/|I|)$ in $\mathcal{M}[0, \infty)$. By (a) above, this implies the analogous statement about the inverse processes $t \leadsto n^{-1/2}v(nt)$ and $t \leadsto |I|L(t)$, in $\mathcal{M}[0, \infty)$; for these processes, however, convergence takes place even in the space $\mathscr{C}[0, 1]$ of continuous functions with respect to uniform topology. This follows from the fact that for each $\varepsilon > 0$,

$$(5.62) \qquad \lim_{\delta \to \infty} \lim_{n \to \infty} \sup P[\max_{0 \leq k\delta \leq 1} (n^{-1/2}v(n(k+1)\delta) - n^{-1/2}v(nk\delta)) > \varepsilon] = 0$$

which can be proved with the help of (b) above.

Part 5.   Proceeding to the general case we only have to establish the convergence of finite dimensional distributions for the Markov chains $t \leadsto n^{-1/2}(v(nt), \xi(nt))$, $nt \in N$ (tightness is an immediate consequence of the previous section and Donsker's invariance principle). The method of proof will be outlined in the case of one and two dimensional marginals.

Convergence of one dimensional distributions follows from the formula

$$(5.63) \qquad \lim_{n \to \infty} E \int_0^\infty e^{-\lambda t} I[n^{-1/2}v(nt) > \alpha] \exp \{i\mu n^{-1/2}\xi(nt)\} \, dt$$

$$= E \exp \left\{-\lambda L^{-1}\left(\frac{\alpha}{|I|}\right)\right\} E \int_0^\infty \exp \{-\lambda t + i\mu \zeta(t)\} \, dt,$$

μ real, $\lambda > 0$. Since, on each compact,

$$(5.64) \qquad t \leadsto \varphi_n(t) := EI[n^{-1/2}v(nt) > \alpha] \exp \{i\mu n^{-1/2}\xi(nt)\}, n \in N$$

has a uniformly equicontinuous family of uniform limit points, α, μ fixed, this implies

$$(5.65) \qquad \lim_{n \to \infty} \varphi_n(t) = EI[|I|L(t) \geq \alpha] \, e^{i\mu \zeta(t)},$$

for each t, α, μ. Uniform equicontinuity follows from Hölder's inequality together with a previously stated limit theorem (see (5.57)) for $nt_2 > nt_1$, integers, explicitly

$$(5.66) \qquad |\varphi_n(t_2) - \varphi_n(t_1)|$$

$$\leq E|I[n^{-1/2}v(nt_2) > \alpha] - I[n^{-1/2}v(nt_1) > \alpha]|$$

$$+ (E|\exp\{i\mu n^{-1/2}\xi(nt_2)\} - \exp\{i\mu n^{-1/2}\xi(nt_1)\}|^2)^{1/2}$$

$$\leq P[n^{-1/2}v(nt_2) > \alpha \ \& \ n^{-1/2}v(nt_1) \leq \alpha] + |\mu|(t_2 - t_1)^{1/2}$$

$$= P[n^{-1}v^{-1}(\sqrt{n}\alpha) \in [(t_1, t_2)] + |\mu|(t_2 - t_1)^{1/2}.$$

This tends to $P[L^{-1}(\alpha/|I|) \in (t_1, t_2)] + |\mu|(t_2 - t_1)^{1/2}$ uniformly in t_1, t_2 as $n \to \infty$. The limit is equal to

$$(5.67) \qquad \int_{t_1}^{t_2} \alpha[|I|(2\pi s^3)^{1/2}]^{-1} \exp \left\{\frac{-\alpha^2}{2s|I|^2}\right\} ds + |\mu|(t_2 - t_1)^{1/2}$$

which is less than or equal to $c(t_2 - t_1)^{1/2}$, the constant c depending only on α and $|\mu|$. Equation (5.65) is equivalent to the convergence of one dimensional distributions.

For the two dimensional distributions the argument will again be outlined only in the nonlattice case; in general linear interpolations of the occurring lattice functions have to be considered. For any two dimensional open set B_1 and $t_2 > t_1$, nt_1, $nt_2 \in N$, we have

$$(5.68) \quad P[n^{-1/2}(v(nt_1), \xi(nt_1)) \in B_1 \ \& \ n^{-1/2}(v(nt_2), \xi(nt_2)) \geq (\ell_2, u_2)]$$

$$= \int_{B_1} P[n^{-1/2}v(n(t_2 - t_1)) \geq \ell_2 - \ell_1$$

$$\& \ n^{-1/2}\xi(n(t_2 - t_1)) \geq u_2 - u_1 | n^{-1/2}\xi(0) = u_1]$$

$$\times \ P[n^{-1/2}(v(nt_1), \xi(nt_1)) \in d\ell_1 \times du_1].$$

This expression converges to the corresponding quantity of Brownian motion provided that the integrand converges to a continuous function uniformly on compacts. For the purpose of proving this let us put

$$(5.69) \quad w_n(u, u_0, \ell) := P[n^{-1/2}v(n(t_2 - t_1)) \geq \ell_2 - \ell$$

$$\& \ n^{-1/2}\xi(n(t_2 - t_1)) \geq u_2 - u | n^{-1/2}\xi(0) = u_0].$$

LEMMA 5.4. *In (ℓ_1, u_1) and uniformly on compacts, we have*

$$(5.70) \quad \lim_{n \to \infty} w_n(u_1, u_1, \ell_1)$$

$$= P[(|I|L(t_2), \zeta(t_2)) \geq (\ell_2, u_2) | (|I|L(t_1), \zeta(t_1)) = (\ell_1, u_1)].$$

PROOF. Application of the triangle inequality to

$$(5.71) \quad w_n(\bar{u}, \bar{u}, \bar{\ell}) - w_n(u, u, \ell)$$

$$= w_n(\bar{u}, \bar{u}, \bar{\ell}) - w_n(u, \bar{u}, \bar{\ell}) + w_n(u, \bar{u}, \bar{\ell}) - w_n(u, u, \bar{\ell})$$

$$+ w_n(u, u, \bar{\ell}) - w_n(u, u, \ell), \qquad \bar{u} > u, \bar{\ell} > \ell$$

yields

$$(5.72) \quad |w_n(\bar{u}, \bar{u}, \bar{\ell}) - w_n(u, u, \ell)|$$

$$\leq P[n^{-1/2}\xi(n(t_2 - t_1)) \in [u_2 - \bar{u}, u_2 - u) | n^{-1/2}\xi(0) = \bar{u}]$$

$$+ |P[n^{-1/2}v(n(t_2 - t_1)) \geq \ell_2 - \bar{\ell} | n^{-1/2}\xi(0) = \bar{u}]$$

$$- P[n^{-1/2}v(n(t_2 - t_1)) \geq \ell_2 - \bar{\ell} | n^{-1/2}\xi(0) = u]|$$

$$+ P[n^{-1/2}v(n(t_2 - t_1)) \in [\ell_2 - \bar{\ell}, \ell_2 - \ell) | n^{-1/2}\xi(0) = u]$$

$$=: I + II + III.$$

The smallness of II and III as t_2 approaches t_1 and $n \to \infty$ is a consequence of the following fundamental lemma (see Blumenthal and Getoor [1], p. 87 4.14)).

LEMMA 5.5. *Let $\{Y_n : n \in N\}$ be a sequence of functionals on the paths of ξ such that $0 \leqq Y_n \leqq 1$, $n \in N$, and let $y_n(u)$ denote $E(Y_n | n^{-1/2}\xi(0) = u)$. Then $\{y_n : n \in N\}$ (their domain being restricted to a compact K) has a uniformly equicontinuous family of uniform limit points provided that the following two conditions are satisfied:*

(i) $\displaystyle \lim_{n \to \infty} \sup_{|t| \leq A, u \in K} |y_n u + t/\sqrt{n} - y_n(u)| = 0$, $A > 0$;

(ii) $\displaystyle \lim_{\delta \to 0} \limsup_{n \to \infty} P[\sup_{0 \leqq s \leqq \delta} |Y_n - Y_n \circ \theta_s| > \varepsilon | n^{-1/2}\xi(0) = u] = 0$,

$u \in K$, $\varepsilon > 0$; θ_s *denotes the translation operator on the space of all continuous functions x defined by $\theta_s x(t) = x(s + t)$.*

PROOF.   It suffices to show that given $\varepsilon > 0$ for every u_0, there is a $\delta > 0$ such that for all but finitely many n

(5.73) $\{u : |u - u_0| < \delta\} \subset \{u : |y_n(u) - y_n(u_0)| < \varepsilon\}$.

The set $\{u : |y_n(u) - y_n(u_0)| < \varepsilon\}$ can be written as $A(u_0, \varepsilon, n) \cup B(u_0, \varepsilon, n)$, where

(5.74)
$$A(u_0, \varepsilon, n) = \{u : y_n(u) - y_n(u_0) \geqq \varepsilon\},$$
$$B(u_0, \varepsilon, n) = \{u : y_n(u_0) - y_n(u) \geqq \varepsilon\}.$$

Assume that there does not exist a $\delta > 0$ with the above property; then there exists a sequence $\{u_n\}$ converging to u_0, $u_n \in A(u_0, \varepsilon, n)$, say, $n \in N$; and $D_n := \{u : |u - u_n| \leqq An^{-1/2}\}$ will be contained in $A(u_0, \varepsilon/2, n)$ for all but finitely many n and every $A > 0$, according to (i). Moreover, if $T_n := \inf\{t : nt \in N \ \& \ n^{-1/2}\xi(nt) \in D_n\}$ (almost surely finite, if $|[0, A]| \geqq 2$), we have

(5.75) $E(Y_n \circ \theta_{T_n} | n^{-1/2}\xi(0) = u_0) = E(y_n(n^{-1/2}\xi(nT_n)) | n^{-1/2}\xi(0) = u_0)$

$$\geqq y_n(u_0) + \frac{\varepsilon}{2}$$

$$= E(Y_n | n^{-1/2}\xi(0) = u_0) + \frac{\varepsilon}{2} \quad \text{as } n \to \infty.$$

On the other hand

(5.76) $E(|Y_n \circ \theta_{T_n} - Y_n| \,|\, n^{-1/2}\xi(0) = u_0)$
$$\leqq P[|Y_n \circ \theta_{T_n} - Y_n| > \varepsilon^2 | n^{-1/2}\xi(0) = u_0] + \varepsilon^2$$
$$\leqq P[T_n > \delta | n^{-1/2}\xi(0) = u_0] + P[\sup_{0 \leqq s \leqq \delta} |Y_n \circ \theta_s - Y_n| > \varepsilon^2] + \varepsilon^2$$

leading to a contradiction since this quantity will be smaller than $2\varepsilon^2$ when δ is chosen properly and $n \to \infty$.

The preceding lemma applied to functionals Y_n of the form $Y_n = I[n^{-1/2}v(n) \geqq \lambda]$ yields Lemma 5.4.

5.2.   Let $X_1, X_2, \cdots$ be a sequence of independent random variables having uniform distribution over $[0, 1]$ and let

$$(5.77) \qquad s \leadsto F_n(s) := \frac{1}{n} \operatorname{card} \{i \leq n : X_i \leq s\}, \qquad\qquad s \in [0, 1], F_0 \equiv 0$$

be their nth initial empirical distribution function. We consider the processes

$$(5.78) \qquad\qquad\qquad t \leadsto \sqrt{n}D(nt), \qquad\qquad\qquad n \in N, t \geq 0,$$

where

$$(5.79) \qquad D(t)_s = \begin{cases} F_t(s) - s & t \in N, s \in [0, 1], \\ (1 - (t - [t]))F_{[t]}(s) \\ \qquad + (t - [t])F_{[t]+1}(s) - s, & t \notin N. \end{cases}$$

Let $\mathscr{D}[0, 1]$ denote the space of functions on the unit interval without discontinuities of the second kind, endowed with the Skorohod topology, and let $\rho_{\mathscr{D}}$ be a metric generating this topology such that

$$(5.80) \qquad\qquad \rho_{\mathscr{D}}(x_1, x_2) \leq \sup_{s \in [0, 1]} |x_1(s) - x_2(s)|, \qquad\qquad x_1, x_2 \in \mathscr{D}[0, 1].$$

THEOREM 5.4 (Müller [16]).   *There exists a process* $t \leadsto \Delta(t) \in \mathscr{D}[0, 1]$ *determined by the following two conditions:*

(i) *Δ has independent increments* $\Delta(t + h) - \Delta(t), h > 0$;

(ii) *the distribution of* $\Delta(t + h) - \Delta(t)$ *coincides with the distribution of* $s \leadsto \sqrt{h}[\zeta(s) - s\zeta(1)]$ *("Brownian bridge"). The process* $(t, s) \leadsto \Delta(t)_s (= \Delta(t)$ *evaluated at* $s)$ *is a Gaussian process over* $[0, \infty) \times [0, 1]$, *continuous with probability* 1, *having covariance function*

$$(5.81) \qquad\qquad E\Delta(t)_s\Delta(t')_{s'} = ts(1 - s'), \qquad\qquad s \leq s', t \leq t'.$$

The process

$$(5.82) \qquad\qquad t \leadsto \hat{\Delta}(t) = \begin{cases} t\Delta(1/t), & t > 0 \\ 0, & t = 0 \end{cases}$$

has the distribution of Δ.

As $n \to \infty$ *the distributions of* $t \leadsto \sqrt{n}D(nt)$ *converge pointwise on continuous bounded functions on* $\mathscr{C}_{\mathscr{D}[0,1]}^0[1, \infty)$ *to the distribution of* $t \leadsto \Delta(t)/t$, *which is the distribution of* $t \leadsto \Delta(1/t)$, *too.*

Moreover, if d_n *denotes the Prohorov distance of the distributions of* $\sqrt{n}D(n \cdot)$ *and* Δ/t *on the space* $\mathscr{C}_{\mathscr{D}[0,1]}^0[1, \infty)$, *then* $d_n = o(n^{-1/6+\varepsilon})$ *for every* $\varepsilon > 0$.

Furthermore, if

$$(5.83) \quad Z_n = \left[g_1\left(\frac{k}{n}, s\right) < \sqrt{n}D(k)_s < g_2\left(\frac{k}{n}, s\right) \quad \text{for all } k \geq n, s \in [0, 1] \right]$$

$$(5.84) \quad Z = \left[g_1(t, s) < \frac{\Delta(t)}{t} < g_2(t, s) \quad \text{for all} \quad (t, s) \in [1, \infty) \times [0, 1] \right],$$

where $g_i, i = 1, 2$, are continuous functions on $[1, \infty) \times [0, 1]$ subject to the conditions

$$(5.85) \qquad \lim_{t^{-1}s(1-s)\to 0} \sup g_1(t, s) < 0, \qquad \lim_{t^{-1}s(1-s)\to 0} \sup g_2(t, s) > 0,$$

then

$$(5.86) \qquad |PZ_n - PZ| = o(n^{-1/6+\varepsilon})$$

for every $\varepsilon > 0$.

COROLLARY 5.8. *Let $X_i, i \in N$, be as in Theorem 5.4 and let μ be the linear interpolation of the sequence median $(X_1, \cdots, X_n)$, $n \in N$. The distributions of $t \rightsquigarrow \sqrt{n}[\mu(nt) - \frac{1}{2}]$ converge to the distribution of $t \rightsquigarrow \zeta(t)/2t$ as $n \to \infty$, pointwise on bounded continuous functions on $\mathscr{C}^0[1, \infty)$.*

PROOF. It suffices to show that for each $\varepsilon > 0$,

$$(5.87) \qquad \lim_{n \to \infty} P[\sup_{t \geq 1} |\sqrt{n}[\mu(nt) - \frac{1}{2}] + \sqrt{n}D(nt)_{1/2}| > \varepsilon] = 0.$$

This follows from

$$(5.88) \qquad \sqrt{n}\big(\inf\{s : F_{n,t}(s) \geq \frac{1}{2}\} - \frac{1}{2}\big)$$

$$= \inf\{\sqrt{n}(s - \frac{1}{2}): \sqrt{n}D(nt)_s \geq \sqrt{n}(\frac{1}{2} - s)\}$$

$$= \inf\{\tau: \sqrt{n}D(nt)_{\tau/\sqrt{n}+1/2} \geq -\tau\}$$

which is close to $-\sqrt{n}D(nt)_{1/2}$ (in probability, as $n \to \infty$).

$$\diamond \qquad \diamond \qquad \diamond \qquad \diamond \qquad \diamond$$

APPENDIX

The following theorem on the large deviations of last entrance times is due to Strassen [26].

Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables having $EX_1 = 0$, $EX_1^2 = 1$ and $Ee^{\lambda X_1} < \infty$ for λ in a neighborhood of 0. Moreover let φ be a function on the positive line satisfying the following conditions:

  (i) $\varphi(t) > 0, t > 0$;
  (ii) $\varphi(t)/\sqrt{t}$ is nondecreasing in t;
  (iii) φ has a continuous derivative φ';
  (iv) $\lim_{t/s \to 1, t \to \infty} \varphi'(t)/\varphi'(s) = 1$;
  (v) $\varphi \leq t^h$ for some $h < \frac{3}{5}$.

THEOREM A.1 (Strassen [26]). *Provided that the left side tends to zero,*

$$(A.1) \qquad P\left[\max\left\{k: \sum_{j=1}^{k} X_j \geq \varphi(k)\right\} > n\right]$$

$$\sim \int_n^\infty (2\pi t)^{-1/2}\varphi'(t) \exp\left\{-\frac{\varphi(t)^2}{2t}\right\} dt, \qquad n \to \infty.$$

REFERENCES

[1] R. M. Blumenthal and R. K. Getoor, *Markov Processes and Potential Theory*, New York, Academic Press, 1968.

[2] L. Breiman, *Probability*, Reading, Addison-Wesley, 1968.

[3] G. J. Chaitin, "On the length of programs for computing finite binary sequences," *J. Assoc. Comput. Mach.*, Vol. 13 (1966), pp. 547–569.

[4] ———, "On the length of programs for computing finite binary sequences: statistical considerations," *J. Assoc. Comput. Mach.*, Vol. 16 (1969), pp. 145–159.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2, New York, Wiley, 1966.

[6] K. Itô and H. P. McKean, *Diffusion Processes and Their Sample Paths*, Berlin–Heidelberg–New York, Springer-Verlag, 1965.

[7] G. Kallianpur and H. Robbins, "The sequence of sums of independent random variables," *Duke Math. J.*, Vol. 21 (1954), pp. 285–307.

[8] A. N. Kolmogorov, "Three approaches to the definition of the concept 'amount of information,' " *Problemy Peredači Informacii*, Vol. 1 (1965), pp. 3–11.

[9] D. W. Loveland, "A variant of the Kolmogorov concept of complexity," *Information and Control*, Vol. 15 (1969), pp. 510–526.

[10] P. Martin-Löf, "The definition of random sequences," *Information and Control*, Vol. 9 (1966), pp. 602–619.

[11] ———, "On the oscillations of the complexity in infinite binary sequences," unpublished, 1965. (In Russian.)

[12] ———, "Algorithms and randomness," *Rev. Inst. Internat. Statist.*, Vol. 37 (1969), pp. 265–272.

[13] ———, "Complexity oscillations in infinite binary sequences," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 19 (1971), pp. 225–230.

[14] P. A. Meyer, "Sur les lois certaines fonctionelles additives: applications aux temps locaux," *Publ. Inst. Statist. Univ. Paris*, Vol. 15 (1966), pp. 295–310.

[15] D. W. Müller, "Verteilungs-Invarianzprinzipien für das starke Gesetz der grossen Zahl," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 10 (1968), pp. 173–192.

[16] ———, "On Glivenko-Cantelli convergence," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 16 (1970), pp. 195–210.

[17] ———, "On the number of averages deviating from the mean," unpublished. (Abstract in *Ann. Math. Statist.*, Vol. 41 (1970), p. 329.)

[18] H. Putnam, "Trial and error predicates and the solution to a problem of Mostowski's," *J. Symbolic Logic*, Vol. 30 (1965), pp. 49–57.

[19] R. Pyke, "The weak convergence of the empirical process with random sample size," *Proc. Cambridge Philos. Soc.*, Vol. 64 (1968), 155–160.

[20] H. Rogers, *Theory of Recursive Functions and Effective Computability*, New York, McGraw-Hill, 1967.

[21] C. P. Schnorr, "Eine Bemerkung zum Begriff der zufälligen Folge," *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, Vol. 14 (1969), pp. 27–35.

[22] A. V. Skorohod, "A limit theorem for sums of independent random variables," *Soviet Math. Dokl.*, Vol. 1 (1960), pp. 810–811.

[23] ———, *Studies in the Theory of Random Processes*, Reading, Addison-Wesley, 1965.

[24] C. Stone, "Limit theorems for random walks, birth and death processes, and diffusion processes," *Illinois J. Math.*, Vol. 7 (1963), pp. 638–660.

[25] ———, "Weak convergence of stochastic processes defined on semi-infinite time intervals," *Proc. Amer. Math. Soc.*, Vol. 14 (1963), pp. 694–696.

[26] V. Strassen, "Almost sure behavior of sums of independent random variables and martingales," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1967, Vol. 2, Part 1, pp. 315–343.

[27] J. Ville, *Etude Critique de la Notion de Collectif*, Paris, Gauthiers-Villars, 1939.