Z. Wahrscheinlichkeitstheorie 3, 211—226 (1964)

An Invariance Principle for the Law of the Iterated Logarithm*

By

V. STRASSEN**

Summary

Let S_n be the sum of the first n of a sequence of independent identically distributed r. v. s. having mean 0 and variance 1. One version of the law of the iterated logarithm asserts that with probability one the set of limit points of the sequence

$$((2\,n\log\log n)^{-1/2} S_n)_{n\geqq 3}$$

coincides with $\langle -1, 1\rangle = \{x : x \text{ real and } |x| \leqq 1\}$ (see HARTMAN-WINTNER [6]). Now consider the continuous function η_n on $\langle 0, 1\rangle$ obtained by linearly interpolating $(2n\log\log n)^{-1/2} S_i$ at i/n. Then we prove (theorem 3) that with probability one the set of limit points of the sequence $(\eta_n)_{n\geqq 3}$ with respect to the uniform topology coincides with the set of absolutely continuous functions x on $\langle 0, 1\rangle$ such that

$$x(0) = 0$$

and

$$\int \dot{x}^2\,dt \leqq 1\,.$$

As applications we obtain, e. g.,

$$Pr\left\{\limsup_{n\to\infty} n^{-1-(a/2)}(2\log\log n)^{-(a/2)}\sum_{i=1}^{n}|S_i|^a = \frac{2(a+2)^{(a/2)-1}}{\left(\int\limits_0^1 \frac{dt}{\sqrt{1-t^a}}\right)^a a^{a/2}}\right\} = 1$$

for any $a \geqq 1$, and

$$Pr\left\{\limsup_{n\to\infty} \nu_n = 1 - \exp\left\{-4\left(\frac{1}{c^2}-1\right)\right\}\right\} = 1\,,$$

where ν_n is the frequency of the events

$$S_i > c\,(2\,i\log\log i)^{1/2}$$

among the first n integers i $(0 \leqq c \leqq 1)$.

To prove theorem 3, we first derive an analogous result for the (k-dimensional) Brownian Motion using well-known ideas of KOLMOGOROV [7] and of ERDÖS and KAC [4]. Also CHUNG's profound paper [1] is to be mentioned here. Then we prove

* This paper was prepared with the partial support of the Office of Naval Research, Contract Nonr-222-43. This paper in whole or in part may be reproduced for any purpose of the U.S. Government.

** Present address: Inst. f. Math. Stat. der Universität, 34 Göttingen, Bürgerstr. 32.

an invariance principle for the law of the iterated logarithm by a powerful device of Skorokhod's [*11*], designed by him to yield improvements of the ordinary invariance principle. The paper assumes no knowledge of the classical law of the iterated logarithm.

I would like to thank Professor M. Loève for directing my attention to the book of Skorokhod and Professor D. Freedman for pointing out an error in the original manuscript.

1. Brownian Motion

Let ζ be the Brownian Motion in R^k. Define

$$\zeta_n(t) = (2\,n \log\log n)^{-1/2}\,\zeta(n\,t)$$

for $t \in \langle 0, 1\rangle$ and $n \geqq 3$. Let C be the Banach space of continuous maps from $\langle 0, 1\rangle$ to R^k endowed with the supremum norm $\|\ \|$, using the euclidean norm in R^k. ζ_n is then a r. v. with values in C. Let K be the set of absolutely continuous $x \in C$ such that

$$x(0) = 0$$

and

$$\int_0^1 (\dot{x}(t))^2\, dt \leqq 1$$

(where $\dot{x}$ denotes the derivative of x determined almost everywhere with respect to Lebesgue measure and the square is meant as inner product). K is a norm-compact subset of C: In fact for $a \leqq b$

$$(1) \qquad |x(b) - x(a)| = \Big|\int_a^b \dot{x}\, dt\Big| \leqq \leqq \left(\int_a^b dt \int_a^b \dot{x}^2\, dt\right)^{1/2} \leqq \leqq \sqrt{b-a},$$

so that K is relatively norm-compact. That K is closed follows immediately from a result of F. Riesz (see [*10*], p. 68, lemma).

If one considers an $x \in C$ as the motion of a mass point with mass 2 from time 0 to time 1, then K consists of those motions for which the mean kinetic energy is $\leqq 1$. This interpretation has a nice feature: The obscure factor $2^{-1/2}$ in the definition of ζ_n which has been kept for historical reasons cancels with the factor 2 of the mass of our masspoint, if one connects them by theorem 1. A similar remark applies to theorem 3.

Theorem 1. *With probability one the sequence* $(\zeta_n)_{n \geqq 3}$ *is relatively norm-compact and the set of its (norm-)limit points coincides with* K.

Proof. Let $\varepsilon > 0$, K_ε be the set of all points in C which have a distance $< \varepsilon$ from K. Then if m is any positive integer and $r > 1$ real

$$\Pr\{\zeta_n \notin K_\varepsilon\} \leqq \Pr\left\{2\,m \sum_{i=1}^{2m}\left(\zeta_n\left(\frac{i}{2m}\right) - \zeta_n\left(\frac{i-1}{2m}\right)\right)^2 > r^2\right\} +$$

$$+ \Pr\left\{2\,m \sum_{i=1}^{2m}\left(\zeta_n\left(\frac{i}{2m}\right) - \zeta_n\left(\frac{i-1}{2m}\right)\right)^2 \leqq r^2 \quad\text{and}\quad \zeta_n \notin K_\varepsilon\right\} = I + II \qquad \text{(say)}.$$

Now

$$\mathrm{I} = \Pr\left\{\chi^2_{2mk} > 2\, r^2 \log\log n\right\} = \frac{1}{\Gamma(mk)} \int\limits_{r^2 \log\log n}^{\infty} t^{mk-1} e^{-t}\, dt$$

$$\sim \frac{(r^2 \log\log n)^{\,mk-1}\, e^{-r^2 \log\log n}}{\Gamma(mk)}$$

for $n \to \infty$ (recall that k is the dimension of the Brownian Motion ζ). Moreover, let η_n be the r. v. with values in C obtained by linearly interpolating the points $\zeta_n\left(\frac{i}{2m}\right)$ at $\frac{i}{2m}$ $(i = 1, \ldots, 2m)$. Then

$$2\, m \sum_{i=1}^{2m} \left(\zeta_n\left(\frac{i}{2m}\right) - \zeta_n\left(\frac{i-1}{2m}\right)\right)^2 \leqq r^2$$

means just $\frac{1}{r}\,\eta_n \in K$ and we get

$$\mathrm{II} = \Pr\left\{\frac{1}{r}\,\eta_n \in K,\ \zeta_n \notin K_\varepsilon\right\} \leqq \Pr\left\{\frac{1}{r}\,\eta_n \in K, \left\|\frac{1}{r}\,\eta_n - \zeta_n\right\| \geqq \varepsilon\right\}.$$

Define the r. v. T by

$$T = \begin{cases} \min\left\{t : t \in \langle 0, 1\rangle, \left|\frac{1}{r}\,\eta_n(t) - \zeta_n(t)\right| \geqq \varepsilon\right\} & \text{if this set is nonempty} \\ 2 & \text{otherwise,} \end{cases}$$

and let F be its d. f., so that

$$\mathrm{II} \leqq \int_0^1 \Pr\left\{\frac{1}{r}\,\eta_n \in K \,\middle|\, T = t\right\} dF(t)$$

$$= \int_0^1 \Pr\left\{\frac{1}{r}\,\eta_n \in K, \left|\frac{1}{r}\,\eta_n(t) - \zeta_n(t)\right| = \varepsilon \,\middle|\, T = t\right\} dF(t).$$

If $i(t)$ is the smallest i with $\frac{i}{2m} \geqq t$, the statement

$$\frac{1}{r}\,\eta_n \in K$$

implies

$$\left|\eta_n\left(\frac{i(t)}{2m}\right) - \eta_n(t)\right| \leqq r \int_t^{i(t)/2m} \left|\frac{1}{r}\,\dot{\eta}_n\right| ds \leqq r \int_t^{i(t)/2m} \left(\frac{1}{r}\,\dot{\eta}_n\right)^2 ds\, \frac{1}{\sqrt{2m}} \leqq \frac{r}{\sqrt{2m}}.$$

The two statements

$$\frac{1}{r}\,\eta_n \in K$$

and

$$\left|\frac{1}{r}\,\eta_n(t) - \zeta_n(t)\right| = \varepsilon$$

together therefore imply

$$\left|\zeta_n\left(\frac{i(t)}{2m}\right) - \zeta_n(t)\right| \geqq \left|\eta_n(t) - \zeta_n(t)\right| - \left|\eta_n\left(\frac{i(t)}{2m}\right) - \eta_n(t)\right|$$

$\left(\text{remember that } \eta_n\left(\frac{i(t)}{2m}\right) = \zeta_n\left(\frac{i(t)}{2m}\right)\right)$

$$\geqq r\,\varepsilon - (r-1) - \frac{r}{\sqrt{2m}} \geqq \frac{\varepsilon}{2}$$

if r is chosen close enough to 1 and m is sufficiently large. Then

$$\mathrm{II} \leqq \int_0^1 \Pr\left\{\left|\zeta_n\left(\frac{i(t)}{2m}\right)-\zeta_n(t)\right| \geqq \frac{\varepsilon}{2} \,\middle|\, T=t\right\} dF(t) \leqq \Pr\left\{\left|\zeta_n\left(\frac{1}{2m}\right)\right| \geqq \frac{\varepsilon}{2}\right\} \int_0^1 dF(t)$$

$$\leqq \Pr\left\{\left|\zeta\left(\frac{n}{2m}\right)\right| \geqq \frac{\varepsilon}{2}\sqrt{2n\log\log n}\right\} \sim \Gamma\left(\frac{k}{2}\right)^{-1}((\varepsilon^2 m \log\log n)/2)^{k/2-1} \times$$
$$\times \exp\{(-\varepsilon^2 m \log\log n)/2\}$$

(compare the estimation of I). By choosing m and r appropriately and using the above estimates for I and II, it is easily seen that for some $r > 1$ and sufficiently large n

$$\Pr\{\zeta_n \notin K_\varepsilon\} \leqq e^{-r^2 \log\log n}.$$

If $n_j = [c^j] + 1$, where $c > 1$, then

$$\sum_j \Pr\{\zeta_{n_j} \notin K_\varepsilon\} \leqq (\log c)^{-r^2} \sum_j j^{-r^2} < \infty$$

so that eventually $\zeta_{n_j} \in K_\varepsilon$ with probability 1. For c sufficiently close to 1 this implies that eventually $\zeta_n \in K_{2\varepsilon}$ with probability one.

This shows that almost surely at most the points of K are limit points of $(\zeta_n)_{n \geqq 3}$ and also that almost surely this sequence is relatively compact (for $\{\zeta_n : n \geqq 3\}$ is totally bounded). To prove the theorem it is therefore sufficient (because of the compactness of K) to show the following: given $x \in K$ and $\varepsilon > 0$, the probability that ζ_n is infinitely often in the open ε-sphere $\{x\}_\varepsilon$ around x equals one. Let $m \geqq 1$ be an integer, $0 < \delta < 1$ and $x^\varkappa$, $\zeta_n^\varkappa$ be the $\varkappa$-th coordinate of x and ζ_n respectively ($1 \leqq \varkappa \leqq k$). We denote the event

$$\left\{\left|\zeta_n^\varkappa\left(\frac{i}{m}\right)-\zeta_n^\varkappa\left(\frac{i-1}{m}\right)-\left(x^\varkappa\left(\frac{i}{m}\right)-x^\varkappa\left(\frac{i-1}{m}\right)\right)\right|<\delta \quad \text{for all } i \text{ with } 2 \leqq i \leqq m \text{ and all } \varkappa\right\}$$

by A_n. Then

$$\Pr(A_n) \geqq \prod_{i=2}^{m}\prod_{\varkappa=1}^{k} \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2m\log\log n}\left|x^\varkappa\left(\frac{i}{m}\right)-x^\varkappa\left(\frac{i-1}{m}\right)\right|}^{\sqrt{2m\log\log n}\left(\left|x^\varkappa\left(\frac{i}{m}\right)-x^\varkappa\left(\frac{i-1}{m}\right)\right|+\delta\right)} e^{-(s^2/2)}\,ds$$

$$\geqq \text{const.} \prod_{i=2}^{m}\prod_{\varkappa=1}^{k} \frac{\exp\left\{-m\left(x^\varkappa\left(\frac{i}{m}\right)-x^\varkappa\left(\frac{i-1}{m}\right)\right)^2 \log\log n\right\}}{\sqrt{m\log\log n}}$$

for n sufficiently large (having used $\int_a^b \frac{e^{-(s^2/2)}}{\sqrt{2\pi}} \geqq \frac{1}{b\sqrt{2\pi}} e^{-(a^2/2)}\left(1-e^{-1/2(b^2-a^2)}\right)$ for $0 \leqq a < b$). So by summing up the exponents and using Schwarz's inequality we get

$$\Pr(A_n) \geqq \frac{\text{const.}}{\log n \sqrt{m\log\log n}}$$

for large n ("large" depending on m and δ). We now put $n_j = m^j$. Then the A_{n_j}'s are mutually independent and

$$\sum_{j=1}^{\infty} \Pr(A_{n_j}) = \infty$$

because $\sum_j \frac{1}{j\sqrt{\log j}}$ diverges. Hence by Borel-Cantelli's lemma infinitely many events A_n happen almost surely. By what we previously have proved ζ_n is eventually close to K, and therefore almost surely we have eventually

$$|\zeta_n(t) - \zeta_n(s)| \leqq \sqrt{|t-s|} + \delta \tag{2}$$

for any $s, t \in \langle 0, 1\rangle$. Now if $y \in C$ the two statements

$$|y(t) - y(s)| \leqq \sqrt{|t-s|} + \delta$$

for all $s, t \in \langle 0, 1\rangle$ and

$$\left| y^{\varkappa}\left(\frac{i}{m}\right) - y^{\varkappa}\left(\frac{i-1}{m}\right) - \left(x^{\varkappa}\left(\frac{i}{m}\right) - x^{\varkappa}\left(\frac{i-1}{m}\right)\right)\right| < \delta$$

for all i and $\varkappa$ with $2 \leqq i \leqq m$ and $1 \leqq \varkappa \leqq k$ together imply

$$\|y - x\| < \varepsilon,$$

provided m is sufficiently large and δ is sufficiently small ("small" depending also on the choice of m).

Looking at the definition of A_n, at the fact that A_n happens infinitely often a. s. and at (2) we conclude that

$$\Pr\{\|\zeta_n - x\| < \varepsilon \text{ infinitely often}\} = 1.$$

This proves the theorem.

The discreteness of n is inessential for the previous considerations. So if $u > e$ (base for the natural logarithm) is real and we put

$$\zeta_u(t) = (2u \log\log u)^{-1/2} \zeta(ut)$$

for $t \in \langle 0, 1\rangle$, we have the following

Corollary 1. *With probability one the net* $(\zeta_u)_{u>e}$ *is relatively norm-compact and the set of its norm-limit points as* u *tends to* ∞ *coincides with* K.

2. The invariance principle

Let $Y_1, Y_2, \ldots$ be a sequence of independent and identically distributed r. v. s. Assume

$$E\,Y_1 = 0$$

and

$$E(Y_1^2) = 1$$

(no further moments are needed). Put

$$S_n = \sum_{i=1}^{n} Y_i, \quad S_0 = 0$$

and

$$\eta(t) = ([t] + 1 - t)\, S_{[t]} + (t - [t])\, S_{[t]+1},$$

i. e., the function η is obtained by linearly interpolating S_n at n.

Theorem 2. *The 1-dimensional Brownian Motion ζ and the above η can be redefined on a common probability space without changing their respective laws (the Wiener measure in the case of ζ), in such a way that*

$$\Pr\{\lim_{t\to\infty}(2t\log\log t)^{-1/2}\sup_{\tau\leq t}|\zeta(\tau)-\eta(\tau)|=0\}=1.$$

Proof. This follows easily from an important result of SKOROKHOD [11], p. 180:

"If a sequence $Y_1, Y_2, \ldots$ of independent identically distributed real r. v.'s satisfying

$$E\,Y_1=0$$

and

$$E(Y_1^2)<\infty$$

is defined together with a 1-dimensional Brownian Motion ζ on a probability space such that

$$Y_1, Y_2, \ldots$$

and

$$\zeta$$

are mutually independent, then there is a sequence

$$\tau_1, \tau_2, \ldots$$

of independent identically distributed nonnegative r. v.'s defined on the same space such that

$$E\,\tau_1=E(Y_1^2)$$

and such that the process

$$\zeta(\tau_1), \zeta(\tau_1+\tau_2)-\zeta(\tau_1), \ldots, \zeta(\sum_{i\leq n}\tau_i)-\zeta(\sum_{i\leq n-1}\tau_i), \ldots$$

and the process

$$Y_1, Y_2, \ldots, Y_n, \ldots$$

have the same distribution."

To apply this result, let us assume that our original sequence $Y_1, Y_2, \ldots$ and a Brownian Motion ζ are already defined on the same probability space and are mutually independent (this of course can always be done). Put $\tau_0=0$ and

$$\tilde{Y}_j=\zeta(\sum_{i=0}^{j}\tau_i)-\zeta(\sum_{i=0}^{j-1}\tau_i)$$

for $j\geq 1$, also

$$\tilde{S}_n=\sum_{i=1}^{n}\tilde{Y}_i=\zeta(\sum_{i=1}^{n}\tau_i), \qquad \tilde{S}_0=0$$

and

$$\tilde{\eta}(t)=([t]+1-t)\,\tilde{S}_{[t]}+(t-[t])\,\tilde{S}_{[t]+1}$$

for any $t\geq 0$. Then $\tilde{\eta}$ has the same distribution as η. Moreover

$$|\tilde{\eta}(t)-\zeta(t)|\leq\max\{|\zeta(\sum_{1\leq i\leq t}\tau_i)-\zeta(t)|, |\zeta(\sum_{1\leq i\leq t+1}\tau_i)-\zeta(t)|\}.$$

Let $\varepsilon > 0$. By KOLMOGOROFF's law of large numbers we have for sufficiently large T

$$\Pr\{|\sum_{1\leqq i\leqq t}\tau_i - t| > t\,\varepsilon \text{ for some } t > T\} < \frac{\varepsilon}{2}$$

and

$$\Pr\{|\sum_{1\leqq i\leqq t+1}\tau_i - t| > t\,\varepsilon \text{ for some } t > T\} < \frac{\varepsilon}{2},$$

so that with probability $> 1-\varepsilon$ for all $t > T$

$$\begin{aligned}&\max\{|\zeta(\sum_{1\leqq i\leqq t}\tau_i) - \zeta(t)|, |\zeta(\sum_{1\leqq i\leqq t+1}\tau_i) - \zeta(t)|\}\\ &\leqq \sup\{|\zeta(s) - \zeta(t)| : s\in\langle t(1-\varepsilon), t(1+\varepsilon)\rangle\}\\ &\leqq 2\sup\{|\zeta(s) - \zeta(t(1+\varepsilon))| : s\in\langle t(1-\varepsilon), t(1+\varepsilon)\rangle\}.\end{aligned}$$

Therefore with probability $> 1-\varepsilon$ we have for all $t > T$

$$\begin{aligned}&(2\,t\log\log t)^{-1/2}|\tilde{\eta}(t) - \zeta(t)|\\ &\leqq 2\sup\{(2\,t\log\log t)^{-1/2}|\zeta(s) - \zeta(t(1+\varepsilon))| : s\in\langle t(1-\varepsilon), t(1+\varepsilon)\rangle\}\\ &\leqq 2\sup\left\{\left((1+\varepsilon)\frac{\log\log(t(1+\varepsilon))}{\log\log t}\right)^{1/2}(2\,u\log\log u)^{-1/2}|\zeta(s'u) - \zeta(u)| : \frac{1-\varepsilon}{1+\varepsilon}\leqq s'\leqq 1\right\}\end{aligned}$$

where we put $u = t(1+\varepsilon)$. For large t this may be continued

$$\leqq 4\sup\{|\zeta_{t(1+\varepsilon)}(s') - \zeta_{t(1+\varepsilon)}(1)| : 1 - 2\,\varepsilon\leqq s'\leqq 1\}.$$

Applying corollary 1, we see that if we restrict attention to a suitable event of somewhat smaller but still large probability (say $> 1 - 2\,\varepsilon$), then for $t > T'$ (nonrandom) we have $\zeta_{t(1+\varepsilon)}\in K_\varepsilon$, so that

$$4\sup\{|\zeta_{t(1+\varepsilon)}(s') - \zeta_{t(1+\varepsilon)}(1)| : 1 - 2\,\varepsilon\leqq s'\leqq 1\}\leqq 4(\sqrt{2\,\varepsilon} + 2\,\varepsilon)$$

(recall (1)). So given $\varepsilon > 0$ there is a T' such that

$$\Pr\{(2\,t\log\log t)^{-1/2}|\tilde{\eta}(t) - \zeta(t)|\leqq 4(\sqrt{2\,\varepsilon} + 2\,\varepsilon) \text{ for all } t > T'\}\geqq 1 - 2\,\varepsilon.$$

This implies

$$\Pr\{\lim_{t\to\infty}(2\,t\log\log t)^{-1/2}|\tilde{\eta}(t) - \zeta(t)| = 0\} = 1$$

and this implies the theorem.

Put

$$\eta_n(t) = (2\,n\log\log n)^{-1/2}\eta(n\,t)$$

for $t\in\langle 0, 1\rangle$ and $n\geqq 3$, i. e., η_n is the r. v. with values in C (where $k = 1$) which is obtained by interpolating linearly

$$(2\,n\log\log n)^{-1/2}S_i$$

at $\frac{i}{n}$. Then we have

Theorem 3. *With probability one the sequence* $(\eta_n)_{n\geqq 3}$ *is relatively norm-compact and the set of its norm-limit points coincides with* K.

Proof. Replacing t by n in theorem 2 we get

$$\Pr\{\lim_{n\to\infty} \| \zeta_n - \eta_n \| = 0\} = 1 .$$

Theorem 3 then follows from theorem 1.

3. Some applications and comments

Let again

$$Y_1, Y_2, \ldots$$

be a sequence of independent identically distributed real r. v.'s with

$$E\, Y_1 = 0$$

and

$$E(Y_1^2) = 1 .$$

(i) The ordinary law of the iterated logarithm for the sequence $Y_1, Y_2, \ldots$ follows from theorem 3. In fact (1) with $a = 0$ and $b = 1$ yields

$$\sup_{x\in K} x(1) = 1 ,$$

where the supremum is attained for and only for $x = t$. In view of theorem 3 this means

$$\Pr\{\limsup_{n\to\infty} (2\, n \log\log n)^{-1/2} S_n = 1\} = 1 .$$

Moreover for small $\varepsilon > 0$ and sufficiently large n ("large" depending on ε and on chance)

$$(2\, n \log\log n)^{-1/2} S_n > 1 - \varepsilon$$

happens only if the sequence

$$S_1, S_2, \ldots, S_n$$

has an approximately linear shape (here we make use of the compactness of K).

(ii) Let f be any Riemann integrable real function on $\langle 0, 1\rangle$,

$$F(t) = \int_t^1 f(s)\, ds$$

for $t \in \langle 0, 1\rangle$. Then

$$\Pr\left\{\limsup_{n\to\infty} (2\, n^3 \log\log n)^{-1/2} \sum_{i=1}^{n} f\left(\frac{i}{n}\right) S_i = \left(\int_0^1 F(t)^2\, dt\right)^{1/2}\right\} = 1 . \tag{3}$$

In particular (putting $f(t) = t^\alpha$)

$$\Pr\left\{\limsup_{n\to\infty} (2\, n^{2\alpha+3} \log\log n)^{-1/2} \sum_{i=1}^{n} i^\alpha S_i = \frac{1}{\sqrt{\left(\alpha + \frac{3}{2}\right)(\alpha + 2)}}\right\} = 1$$

for any $\alpha > -1$.

To prove (3) we make use of the following

Corollary of Theorem 3. *If φ is a continuous map from C (= Banach space of continuous real functions on $\langle 0, 1\rangle$) to some Hausdorff space H (in all following applications H will be the set of real numbers), then with probability* 1 *the sequence $(\varphi(\eta_n))_{n \geq 3}$ is relatively compact and the set of its limit points coincides with $\varphi(K)$.*

Proof. In general if a relatively compact sequence $(y_n)_{n \geqq 1}$ of points in C has some (compact) K as the set of its limit points, then $\varphi(K)$ is the set of limit points of $(\varphi(y_n))_{n \geqq 1}$. The corollary follows if we assume the basic probability space to be complete (so as to be sure that the event of coincidence between the set of limit points of $(\varphi(\eta_n))_{n \geqq 3}$ and $\varphi(K)$ is measurable).

We apply this corollary to the function φ defined by

$$\varphi(x) = \int_0^1 x(t) f(t)\, dt, \qquad (x \in C)$$

to get

$$\Pr\{\limsup_{n\to\infty} \int_0^1 \eta_n(t) f(t)\, dt = \sup \varphi(K)\} = 1 .$$

But

$$\begin{aligned}\sup \varphi(K) &= \sup_{x \in K} \int_0^1 x(t) f(t)\, dt \\ &= \sup_{x \in K} \int_0^1 F(t) \dot{x}(t)\, dt \\ &= \sup\{ \int_0^1 F(t) y(t)\, dt : \int_0^1 y^2(t)\, dt \leqq 1\} \\ &= (\int_0^1 (F(t))^2 dt)^{1/2}\end{aligned}$$

(evaluation of the supremum of a linear functional on the unit sphere of a Hilbert space).

An elementary consideration yields

$$\Pr\left\{\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f\left(\frac{i}{n}\right) (2\, n \log\log n)^{-1/2} S_i = \limsup_{n\to\infty} \int_0^1 f(t)\, \eta_n(t)\, dt\right\} = 1 ,$$

so that (3) is proved.

(iii) Let $a \geqq 1$ be real. Then

$$(4) \qquad \Pr\left\{\limsup_{n\to\infty} n^{-1-(a/2)} (2 \log\log n)^{-a/2} \sum_{i=1}^{n} |S_i|^a = \frac{2(a+2)^{(a/2)-1}}{\left(\int_0^1 \frac{dt}{\sqrt{1-t^a}}\right)^a a^{a/2}}\right\} = 1 ,$$

in particular

$$\Pr\left\{\limsup_{n\to\infty} n^{-3/2} (2 \log\log n)^{-1/2} \sum_{i=1}^{n} |S_i| = \frac{1}{\sqrt{3}}\right\} = 1$$

and

$$\Pr\left\{\limsup_{n\to\infty} n^{-2} (2 \log\log)^{-1} \sum_{i=1}^{n} S_i^2 = \frac{4}{\pi^2}\right\} = 1 .$$

One reduces the proof of (4) in an entirely analogous way as above to that of the following assertion:

$$\sup_{x\in K}\int_0^1 |x(t)|^a\,dt = \frac{2(a+2)^{(a/2)-1}}{\left(\int_0^1 \frac{dt}{\sqrt{1-t^a}}\right)^a a^{a/2}}. \tag{5}$$

Now if Λ^2 is the Hilbert space of all absolutely continuous $x\in C$ such that

$$x(0) = 0$$

and

$$\int_0^1 \dot{x}^2\,dt < \infty$$

endowed with the inner product

$$((x, y)) = \int_0^1 \dot{x}\dot{y}\,dt,$$

we have

$$\sup_{x\in K}\int_0^1 |x|^a\,dt = \sup\left\{\int_0^1 |x|^a\,dt : x\in\Lambda^2 \text{ and } ((x,x)) = 1\right\}.$$

The right-hand side can be evaluated by classical methods of the calculus of variations. We know that the supremum is obtained by some x (because K is a norm compact subset of C and any $x\in K$ which maximizes $\int|x|^a dt$ satisfies $((x,x)) = 1$). Without loss of generality we may assume $x \geqq 0$. A necessary condition for x is the existence of a Lagrange multiplier β such that for all $y\in\Lambda^2$

$$\int_0^1 a\,x^{a-1}\,dt = \beta\, 2\int_0^1 \dot{x}\dot{y}\,dt$$

(the left-hand side is the derivative of the functional $\int|x|^a dt$ at x applied to y, i. e.,

$$\left(\frac{\partial}{\partial\varepsilon}\int_0^1 |x+\varepsilon y|^a\,dt\right)_{\varepsilon=0}$$

(see [2]), the right-hand side is β times the derivative of the functional $((x,x))$ at x applied to y). Partial integration of the left-hand side yields

$$\int_0^1\int_t^1 a(x(s))^{a-1}\,ds\,\dot{y}(t)\,dt = 2\beta\int_0^1 \dot{x}\dot{y}\,dt,$$

therefore

$$\int_t^1 a\,x^{a-1}\,ds = 2\beta\,\dot{x}(t) \tag{6}$$

which shows that $\dot{x}$ has a continuous derivative (of course $\beta \neq 0$) and also that

$$\dot{x}(1) = 0.$$

Differentiating, multiplying with $\dot{x}(t)$ and integrating again yields

$$x^a + \beta\dot{x}^2 = x(1)^a + \beta\dot{x}(1)^2 = x(1)^a.$$

From the significance of x and from our assumption $x \geqq 0$ it follows that x is nondecreasing, $x(1) > 0$, so from the above equation

$$\dot{x}(0) > 0,$$

$$\beta = \frac{x(1)^a}{\dot{x}(0)^2} > 0.$$

Using (6) we see also that

$$\dot{x}(t) > 0$$

for all t. The last equation becomes

$$x^a + \frac{x(1)^a}{\dot{x}(0)^2}\dot{x}^2 = x(1)^a. \tag{7}$$

Separation of variables and integration yields

$$t = \int_0^{x(t)} \frac{du}{\dot{x}(0)\sqrt{1 - \frac{u^a}{x(1)^a}}} \tag{8}$$

so that

$$1 = \frac{x(1)}{\dot{x}(0)} \int_0^1 \frac{dv}{\dot{x}(0)\sqrt{1 - v^a}} = \frac{x(1)}{\dot{x}(0)}\gamma \qquad \text{(say)}.$$

Now, on the one hand, using (7) and $\int \dot{x}^2 dt = 1$

$$\int_0^1 x^a\,dt = x(1)^a\left(1 - \frac{1}{\dot{x}(0)^2}\right),$$

on the other hand, using (8)

$$\int_0^1 x^a\,dt = \int_0^{x(1)} \frac{u^a\,du}{\dot{x}(0)\sqrt{1 - (u^a/x(1)^a)}} = \frac{2x(1)^{a+1}}{(a+2)\dot{x}(0)}\gamma.$$

Eliminating $x(1)$ and $\dot{x}(0)$ from the last 3 equations we get the desired conclusion (5).

Remarks: If a is an integer, $|S_i|^a$ in (4) can of course be replaced by S_i^a. Also: For sufficiently large n ("large" being random) whenever

$$n^{-1-(a/2)}(2\log\log n)^{-a/2}\sum_{i=1}^{n}|S_i|^a$$

is close to its limes superior

$$\frac{2(a+2)^{(a/2)-1}}{\gamma^a a^{a/2}}$$

then η_n is close to either ψ or $-\psi$, where

$$t = \frac{1}{\gamma}\int_0^{\gamma(a/a+2)^{1/2}\psi(t)} \frac{dv}{\sqrt{1-v^a}}$$

(e. g., if $a = 2$, we have $\psi(t) = (\sqrt{8}/\pi)\sin(\pi/2)t$).

On might inquire about

$$\limsup_{i\to\infty} n_i^{-2} (2 \log\log n_i)^{-1} \sum_{j=1}^{n_i} S_j^2 ,$$

where n_i runs through all n such that

$$S_n S_{n-1} \leqq 0 .$$

It is easy to see that the above quantity equals almost surely

$$\sup\{ \int_0^1 x^2 dt : x \in K \text{ and } x(1) = 0\} .$$

This supremum equals $1/\pi^2$ and is obtained by the two functions $\pm \sqrt{2}/\pi \sin \pi t$, as can be derived from (8) using a symmetry argument, or directly from WIRTINGER's lemma.

(iv) $$\Pr\left\{ \limsup (2 n \log\log n)^{-1/2} \frac{\sum_{i=1}^{n} S_i^2}{\sum_{i=1}^{n} |S_i|} = 2p \right\} = 1$$

where p is the largest solution of

$$\sqrt{1-p} \sin \frac{\sqrt{1-p}}{p} + \cos \frac{\sqrt{1-p}}{p} = 0$$

for $0 < p < 1$.

Again one has only to prove

$$\sup \left\{ \frac{\int x^2 dt}{\int |x| dt} : x \in \Lambda^2 \text{ and } ((x, x)) = 1 \right\} = 2p .$$

The supremum is attained, say at x. We may assume $x \geqq 0$, in fact $x(t) > 0$ for all $t > 0$. Taking derivatives we get with a Lagrange multiplier β

$$(\int x\,dt \int 2\,x\,y\,dt - \int x^2 dt \int y\,dt)\,(\int x\,dt)^{-2} = \beta \int 2\,\dot{x}\,\dot{y}\,dt$$

for all $y \in \Lambda^2$ as a necessary condition on x. Substituting for $\dot{y}$ functions of the form

$$\dot{y}_s(t) = \begin{cases} \frac{1}{s} & \text{if } t \in \langle t_0, s \rangle \\ 0 & \text{otherwise} \end{cases}$$

we see that $\dot{x}$ has a continuous derivative. Putting

$$\alpha = (\int x\,dt)^2 \beta$$

we obtain by a partial integration from the above equation

$$\int x\,dt \int 2\,x\,y\,dt - \int x^2 dt \int y\,dt = 2\,\alpha\,\dot{x}(1)\,y(1) - 2\,\alpha \int \ddot{x}\,y\,dt$$

and therefore

(9) $$2\,x \int x\,dt - \int x^2 dt + 2\,\alpha\,\ddot{x} = 0$$

and

$$\dot{x}(1) = 0.$$

Multiplying (9) by $\dot{x}$ and integrating we get

$$x^2 \int x\,dt - x \int x^2\,dt + \alpha(\dot{x}^2 - \dot{x}(0)^2) = 0.$$

Integrating from 0 to 1

$$\dot{x}(0) = 1.$$

Any solution of (9) has the form

$$z(t) = p + q\cos\lambda t + r\sin\lambda t.$$

The conditions $z(0) = 0$, $\dot{z}(0) = 1$, $\dot{z}(1) = 0$ and $\int \dot{z}^2 dt = 1$ specify p, q, r, λ:

$$z(t) = p - p\cos\frac{\sqrt{1-p}}{p}\,t + \frac{p}{\sqrt{1-p}}\sin\frac{\sqrt{1-p}}{p}\,t$$

where $-\infty < p < 1$, $p \neq 0$ and

(10) $$\sqrt{1-p}\,\sin\frac{\sqrt{1-p}}{p} + \cos\frac{\sqrt{1-p}}{p} = 0.$$

For any such p the corresponding z actually solves (9) and satisfies the side conditions. Moreover

$$\int z^2\,dt = \frac{2p^2}{1-p},$$

$$\int z\,dt = \frac{p}{1-p},$$

so that

$$\frac{\int z^2 dt}{\int z\,dt} = 2p.$$

One easily checks that the z corresponding to the largest value of $p < 1$ satisfying (10) is positive in $(0, 1\rangle$, so that it must coincide with x.

(v) In view of the ordinary law of the iterated logarithm it seems natural to ask about the relative frequency of events

$$S_n > (1-\varepsilon)(2n\log\log n)^{1/2}.$$

Let $0 \leqq c \leqq 1$ and

$$c_i = \begin{cases} 1 & \text{if} \quad S_i > c(2i\log\log i)^{1/2} \\ 0 & \text{otherwise.} \end{cases}$$

Then

(11) $$\Pr\left\{\limsup_{n\to\infty} \frac{1}{n}\sum_{i=3}^{n} c_i = 1 - \exp\left\{-4\left(\frac{1}{c^2} - 1\right)\right\}\right\} = 1.$$

For $c = \frac{1}{2}$ as an example we get the somewhat surprising result that with probability one for infinitely many n the percentage of times $i \leqq n$ when

$$S_i > \tfrac{1}{2}(2i\log\log i)^{1/2}$$

exceeds 99.999, but only for finitely many n exceeds 99.9999.

It suffices to prove (11) for $0 < c < 1$. One easily shows, using the fact that for any $\alpha \in (0, 1)$

$$\log\log\alpha n \sim \log\log n,$$

that with probability one

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=[\alpha n]}^{n} c_i \leqq \sup_{x\in K} m\{t : \alpha \leqq t \leqq 1 \text{ and } x(t) \geqq c'\sqrt{t}\}$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=[\alpha n]}^{n} c_i \geqq \sup_{x\in K} m\{t : \alpha \leqq t \leqq 1 \text{ and } x(t) \geqq c''\sqrt{t}\},$$

where $c' < c < c''$ arbitrary and m denotes Lebesgue measure (note that

$$m\{t : \alpha \leqq t \leqq 1 \text{ and } x(t) \geqq c\sqrt{t}\}$$

is not a continuous function of x, so that one has to go back to theorem 3 instead of using its corollary). For continuity reasons it is therefore sufficient to prove

$$\sup_{x\in K} m\{t : x(t) \geqq c\sqrt{t}\} = 1 - \exp\left\{-4\left(\frac{1}{c^2} - 1\right)\right\}. \tag{12}$$

Now $m\{t : x(t) \geqq c\sqrt{t}\}$ is upper semicontinuous in x, so that the supremum is attained, say by $x_0 \in K$. It is easy to see that all functions y_0 for which the supremum is attained have to satisfy

$$\int \dot{y}_0^2\, dt = 1 \tag{13}$$

(one uses the fact that for $c > 0$ the set $\{t : y_0(t) < c\sqrt{t}\}$ is not empty). Generally if $0 \leqq t_0 < t_1 \leqq 1$ and $x \in K$ (only $x \in \Lambda^2$ is needed), then

$$\int_{t_0}^{t_1} \dot{x}^2\, dt \geqq \frac{(x(t_1) - x(t_0))^2}{t_1 - t_0}, \tag{14}$$

where equality holds if x is linear. This follows from Jensen's inequality (see [3]) and is known in point mechanics. The linear connection of $x(t_0)$ with $x(t_1)$ will be called the (t_0, t_1)-secant of x.

The set $\{t : x_0(t) > c\sqrt{t}\}$ is empty. Otherwise there would exist t_0, t_1 such that $0 \leqq t_0 < t_1 \leqq 1$, x_0 is not linear in $\langle t_0, t_1\rangle$ and the (t_0, t_1)-secant of x_0 would still be greater than $c\sqrt{t}$ in $\langle t_0, t_1\rangle$. Replacing x_0 by its (t_0, t_1)-secant in $\langle t_0, t_1\rangle$ we would get a y_0 contradicting (13).

The point 0 is not an accumulation point of $\{t : x_0(t) = c\sqrt{t}\}$. Otherwise there would be $1 \geqq t_1 > t_2 > \cdots$ such that for all i

$$\sqrt{t_{i+1}} < \tfrac{1}{2}\sqrt{t_i}$$

and

$$x(t_i) = c\sqrt{t_i}.$$

But then using (14)

$$1 \geqq \int_0^{t_1} \dot{x}_0^2\, dt \geqq \sum_{i=1}^{\infty} \int_{t_{i+1}}^{t_i} \dot{x}_0^2\, dt \geqq \sum_{i\geqq 1} \frac{c^2\left(\sqrt{t_i} - \sqrt{t_{i+1}}\right)^2}{t_i - t_{i+1}} > \sum_{i\geqq 1} \frac{c^2}{4} = \infty.$$

Therefore the open set

$$\{t : x_0(t) < c\sqrt{t}\},$$

which as any open set is the disjoint union of open intervals, contains among its components an interval of the form $(0, s_0)$. By (13) and (14) x_0 is linear in $\langle 0, s_0\rangle$ and

$$\int_0^{s_0} \dot{x}_0^2\, dt = c^2$$

(note that this value does not depend on s_0). If $\{t : x_0(t) < c\sqrt{t}\} \neq (0, s_0)$, there is another nonempty open interval (s_1, s_2) such that $x_0(t) < c\sqrt{t}$ on (s_1, s_2) and $x_0(s_1) = c\sqrt{s_1}$, $x_0(s_2) = c\sqrt{s_2}$. But then consider the function y_0 defined by

$$y_0(t) = \begin{cases} \dfrac{tc}{\sqrt{s_0 + s_2 - s_1}} & \text{if} \quad 0 \leqq t \leqq s_0 + s_2 - s_1 \\ c\sqrt{s_0 + s_2 - s_1} + x_0(t - s_2 + s_1) - x_0(s_0) & \text{if} \quad s_0 + s_2 - s_1 \leqq t \leqq s_2 \\ c\sqrt{s_0 + s_2 - s_1} + x_0(s_1) - x_0(s_0) + x_0(t) - x_0(s_2) & \text{if} \quad s_2 \leqq t \leqq 1 . \end{cases}$$

We have

$$y_0 \in \Lambda^2 ,$$

$$m\{t : y_0(t) \geqq c\sqrt{t}\} \geqq m\{t : x_0(t) \geqq c\sqrt{t}\}$$

$$\int_0^1 \dot{y}_0^2\, dt = c^2 + \int_{s_0}^{s_1} \dot{x}_0^2\, dt + \int_{s_2}^{1} \dot{x}_0^2\, dt = 1 - \int_{s_1}^{s_2} \dot{x}_0^2\, dt < 1 ,$$

contradicting (13). So x_0 is linear in $\langle 0, s_0\rangle$ and coincides with $c\sqrt{t}$ in $\langle s_0, 1\rangle$. $\int \dot{x}_0^2 dt = 1$ determines s_0 as $\exp\{-4((1/c^2) - 1)\}$, which proves (11).

(vi) Though $K \subseteqq \Lambda^2$ and $\eta_n \in \Lambda^2$ for all $n \geqq 3$, the sequence $(\eta_n)_{n \geqq 3}$ with probability one has no limit points in Λ^2 with respect to the Hilbert space norm. In fact

$$\Pr\{\lim_{n\to\infty} \int (\dot{\eta}_n)^2\, dt = \infty\} = 1 .$$

For any strictly increasing sequence of integers $a_n \geqq 0$ such that $a_0 = 0$ and

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1$$

let $\tilde{\eta}_n \in C$ be obtained by linearly interpolating

$$(2\, a_n \log\log a_n)^{-1/2}\, S_{a_i}$$

at a_i/a_n for $0 \leqq i \leqq n$. It would be interesting to know for which sequences $(a_n)_{n \geqq 0}$ with probability one the set of limit points of $(\tilde{\eta}_n)_{n \geqq 3}$ with respect to the Hilbert space norm in Λ^2 coincides with the unit sphere K in Λ^2. In this connection see also LAMPERTI [8].

It would also be very interesting to find the strong form (in the sense of FELLER [5]) of our law of the iterated logarithm, at least for Brownian Motion.

References

[1] CHUNG, K. L.: On the maximum partial sum of sequences of independent random variables. Trans. Amer. Math. Soc., **64**, 205—233 (1948).

[2] DIEDONNÉ, I.: Foundations of Modern Analysis, Pure and Applied Mathematics. New York-London: 1960.

[3] DOOB, J. L.: Stochastic Processes, Wiley-Publications in Statistics. New York-London: Wiley 1959.

[4] ERDÖS, P., and M. KAC: On certain limit theorems of the theory of probability. Bull. Amer. Math. Soc. **52**, 292–302 (1946).
[5] FELLER, W.: The general form of the so-called law of the iterated logarithm. Trans. Amer. Math. Soc., **54**, 373–402 (1943).
[6] HARTMAN, P., and A. WINTNER: On the law of the iterated logarithm. Amer. J. Math., **63**, 169–176 (1941).
[7] KOLMOGOROV, A.: Das Gesetz des iterierten Logarithmus. Math. Annalen **101**, 126–135 (1929).
[8] LAMPERTI, J.: On convergence of stochastic processes, Trans. Amer. Math. Soc. **104**, 430–435 (1962).
[9] LOÈVE, M.: Probability Theory, The University series in higher Mathem., Princeton (1960).
[10] RIESZ, F., and B. SZ. NAGY: Vorlesungen über Funktionalanalysis. Hochschulbücher für Mathematik, Berlin (1956).
[11] SKOROKHOD, A.B.: Research on the Theory of Random Processes. Kiew (1961) (in Russian).

University of California, Berkeley
and Universität Göttingen

(Received July 20, 1964)