Studia Scientiarum Mathematicarum Hungarica

VOLUME 33
NUMBERS 1–3
1997

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian. It is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
H-1117 Budapest, Prielle Kornélia u. 19-35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H-1364 Budapest

Tel.: +36 1 118 2875 Fax: +36 1 117 7166
e-mail: merza @ math-inst.hu

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H-1519 Budapest

For 1997 volume 33 is scheduled for publication. The subscription price is $ 120.00, air delivery plus $ 20.00.

Coden: SSMHAX August, 1997
Vol: 33 **Pages**: 1-343
Numbers: 1-3 **Whole**: 64

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 1–24

ALMOST SURE INVARIANCE PRINCIPLES FOR LOGARITHMIC AVERAGES

I. BERKES and L. HORVÁTH

Dedicated to Endre Csáki on his sixtieth birthday

1. Introduction

Let $X_1, X_2, \dots$ be independent, identically distributed random variables with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and let $S_n = X_1 + \dots + X_n$. By the a.s. central limit theorem (Brosamler [4], Schatte [15], Lacey and Philipp [11], Fisher [9])

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k \le N} \frac{1}{k} I\left\{\frac{S_k}{\sqrt{k}} < x\right\} = \Phi(x) \quad \text{a.s. for all } x, \tag{1.1}$$

where I denotes indicator function and Φ stands for the standard normal distribution function. Several papers dealt with 'logarithmic' limit theorems of the type (1.1) and many generalizations of (1.1) have been obtained. In particular, the following theorem extends (1.1) for a large class of independent sequences:

THEOREM A (Berkes and Dehling [2]). *Let $X_1, X_2, \dots$ be independent random variables and (a_n) a positive numerical sequence such that setting $S_n = X_1 + \dots + X_n$ we have*

$$\mathbf{E}\left(\log\log\left|\frac{S_n}{a_n}\right|\right)^{1+\delta} \le K \qquad (n = 1, 2, \dots) \tag{1.2}$$

$$a_l/a_k \ge C(l/k)^\gamma \qquad (1 \le k \le l) \tag{1.3}$$

for some positive constants C, K, δ and γ. Then for any bounded Lipschitz 1 function f on R we have

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k \le N} \frac{1}{k} \left(f\left(\frac{S_k}{a_k}\right) - \mathbf{E}f\left(\frac{S_k}{a_k}\right)\right) = 0 \qquad a.s.. \tag{1.4}$$

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F17.
Key words and phrases. Pointwise central limit theorem, logarithmic averages, a.s. invariance principle.

The first author's research has been supported by Hungarian National Foundation for Scientific Research Grants No. T 16384 and T 19346.

0081–6906/97/$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

Theorem A and standard properties of weak convergence (see e.g. Dudley [7], Theorem 8.3) imply that under (1.2), (1.3) the relations

$$\lim_{N\to\infty}\frac{1}{\log N}\sum_{k\leqq N}\frac{1}{k}I\left\{\frac{S_k}{a_k}<x\right\}=\Phi(x)\quad\text{a.s. for all } x \tag{1.5}$$

and

$$\lim_{N\to\infty}\frac{1}{\log N}\sum_{k\leqq N}\frac{1}{k}\mathbf{P}\left\{\frac{S_k}{a_k}<x\right\}=\Phi(x)\quad\text{for all } x$$

are equivalent. In particular, a sufficient condition for the a.s. central limit theorem (1.5) is

$$S_n/a_n\xrightarrow{\mathcal{D}}N(0,1).$$

Condition (1.3) is satisfied, e.g., if $n^{-\gamma}a_n$ is nondecreasing or if $a_n=n^\rho L(n)$ where $\rho>\gamma$ and L is a slowly varying function.

The purpose of the present paper is to prove a.s. invariance principles corresponding to relation (1.4). Our first result is the following

THEOREM 1. *Let $X_1,X_2,\dots$ be independent random variables, $f:R\to R$ a bounded measurable function and (a_n) a positive numerical sequence such that*

$$\mathbf{E}\left(\log\left|\frac{S_n}{a_n}\right|\right)^\alpha\leqq K\qquad(n=1,2,\dots) \tag{1.6}$$

$$\mathbf{E}\sup_{|t|\leqq h}\left|f\left(\frac{S_n}{a_n}+t\right)-f\left(\frac{S_n}{a_n}\right)\right|\leqq K\left(\log\frac{1}{h}\right)^{-\beta}\quad\textit{for}\quad Ca_n^{-1/2}\leqq h<1 \tag{1.7}$$

$$a_l/a_k\geqq C(l/k)^\gamma\qquad(1\leqq k\leqq l) \tag{1.8}$$

$$\lambda_N:=\mathbf{Var}\sum_{k\leqq N}\frac{1}{k}f\left(\frac{S_k}{a_k}\right)\geqq C(\log N)^\delta \tag{1.9}$$

for some positive constants K, C, α, β, γ, δ satisfying

$$\alpha>8,\qquad\beta>8,\qquad\delta>5/6. \tag{1.10}$$

Then there exists a Wiener process W such that

$$\sum_{k\leqq N}\frac{1}{k}\left(f\left(\frac{S_k}{a_k}\right)-\mathbf{E}f\left(\frac{S_k}{a_k}\right)\right)=W(\lambda_N)+O(\lambda_N^{\frac{1}{2}-\eta})\quad a.s. \tag{1.11}$$

for some positive constant η.

COROLLARY 1. *Under the conditions of Theorem 1 we have*

$$\lambda_N^{-1/2} \sum_{k\le N} \frac{1}{k}\left(f\left(\frac{S_k}{a_k}\right) - \mathbf{E}f\left(\frac{S_k}{a_k}\right)\right) \xrightarrow{\mathcal{D}} N(0,1) \tag{1.12}$$

$$\limsup_{N\to\infty} (2\lambda_N \log\log \lambda_N)^{-1/2} \sum_{k\le N} \frac{1}{k}\left(f\left(\frac{S_k}{a_k}\right) - \mathbf{E}f\left(\frac{S_k}{a_k}\right)\right) = 1 \quad a.s.. \tag{1.13}$$

The surprising feature of Theorem 1 is that the Wiener approximation (1.11) (and thus the CLT (1.12) and LIL (1.13)) hold regardless the limiting behavior of S_k/a_k. For example, the conditions of the theorem are satisfied if X_n are i.i.d. r.v.'s with symmetric stable distribution with parameter $0<p<2$ and $a_n = n^{1/p}$. (This special case is treated in Corollary 4 and in [3].)

Conditions (1.6) and (1.7) are very mild and are satisfied in most situations of interest. The bounds (1.10) for α, β, δ can be weakened but we made no effort to find the minimal values. (Actually, the proof of the theorem will show that the result is valid for any triple (α,β,δ) such that $\min(\alpha,\beta) > 4/(3\delta-2)$ and thus choosing δ closer to 1 leads to weaker bounds for α,β.) Condition (1.6) is trivially satisfied if S_n/a_n has bounded p-th moments for some $p>0$ and (1.7) is valid if f is a Lipschitz function or even if it is logarithmic Lipschitz, i.e.,

$$|f(x+h) - f(x)| \leqq \text{const} \cdot \left(\log \frac{1}{h}\right)^{-\beta} \qquad (0<h<1)$$

with some $\beta>0$. Moreover, (1.7) is satisfied if f is the indicator function of an interval and (X_n) obeys

$$\sup_a \mathbf{P}\left(a < \frac{S_n}{a_n} < a+h\right) \leqq K\left(\log\frac{1}{h}\right)^{-\beta} \qquad \text{for } Ca_n^{-1/2} \leqq h < 1. \tag{1.14}$$

In applications (1.14) can be verified by using standard concentration function inequalities (see the proof of Corollaries 2, 4). For example, (1.14) holds (even with Kh^β on the right-hand side) if X_n are i.i.d. r.v.'s with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and $a_n = \sqrt{n}$ or if X_n are i.i.d. r.v.'s in the domain of normal attraction of a symmetric stable law with parameter $0<p<2$ and $a_n = n^{1/p}$. Finally, (1.9) can also be shown to hold in a number of standard situations. For example, in Section 2 we shall see that (1.9) holds with $a_k = \sqrt{k}$ if $\mathbf{E}X_n = 0$, $\mathbf{E}X_n^2 = 1$ $(n=1,2,\dots)$ and X_n^2 is uniformly integrable, i.e.,

$$\sup_n \mathbf{E}X_n^2 I(|X_n| \geqq t) \to 0 \quad \text{as} \quad t\to\infty. \tag{1.15}$$

This covers the i.i.d. case and leads to Corollaries 2, 3 improving and extending several earlier results in the field, in particular the CLT's and a.s. invariance principles in [5], [10], [18]. It is also worth noting that if we strengthen (1.6) and (1.7) then the rate $(\log N)^\delta$ in (1.9) can be substantially weakened. In fact, we have the following

THEOREM 2. *Let $X_1, X_2, \ldots$ be independent random variables, $f\colon R \to R$ a bounded measurable function and (a_n) a positive numerical sequence such that setting $S_n = X_1 + \cdots + X_n$ we have*

$$\mathbf{E}\left|\frac{S_n}{a_n}\right|^p \leqq K \qquad (n=1,2,\ldots), \tag{1.16}$$

$$\mathbf{E}\sup_{|t|\leqq h}\left|f\left(\frac{S_n}{a_n}+t\right)-f\left(\frac{S_n}{a_n}\right)\right| \leqq Kh^\beta \qquad \text{for } h \geqq Ca_n^{-1/2}, \tag{1.17}$$

$$a_l/a_k \geqq C(l/k)^\gamma \qquad (1 \leqq k \leqq l), \tag{1.18}$$

$$\mathbf{Var}\sum_{k=2^M+1}^{2^{M+N}} \frac{1}{k} f\left(\frac{S_k}{a_k}\right) \geqq \omega(N) \qquad \text{for all } M \geqq 0,\ N \geqq 1, \tag{1.19}$$

where K, C, p, β, γ are positive constants and ω is a positive function with $\omega(N) \to +\infty$. Let

$$\lambda_N = \mathbf{Var}\sum_{k \leqq N} \frac{1}{k} f\left(\frac{S_k}{a_k}\right). \tag{1.20}$$

Then there exists a Wiener process W and a positive constant η such that (1.11) *holds.*

We mention now a few consequences of our theorems. To simplify the formulas, let $\mathcal{L}$ denote the class of nonconstant Lipschitz functions, i.e., the set of nonconstant, bounded functions $f\colon R \to R$ satisfying

$$|f(x+h)-f(x)| \leqq Kh^\beta \qquad (x \in R,\ h > 0)$$

for some $\beta > 0$; let further $\mathcal{I}$ denote the set of indicator functions of intervals (finite or infinite, but $\neq R$).

COROLLARY 2. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and let $f \in \mathcal{L} \cup \mathcal{I}$. Then there exists a constant $\sigma_f > 0$, a sequence*

$\lambda_N \sim \sigma_f \log N$ and a Wiener process W such that (1.11) *holds with $a_k = \sqrt{k}$. In particular, we have*

$$(1.21)\qquad (\log N)^{-1/2} \sum_{k \leq N} \frac{1}{k}\left(f\left(\frac{S_k}{\sqrt{k}}\right) - \mathbf{E}f\left(\frac{S_k}{\sqrt{k}}\right)\right) \xrightarrow{\mathcal{D}} N(0,\sigma_f)$$

and

$$\limsup_{N\to\infty}(2\log N \log\log\log N)^{-1/2}\sum_{k\leq N}\frac{1}{k}\left(f\left(\frac{S_k}{\sqrt{k}}\right)-\mathbf{E}f\left(\frac{S_k}{\sqrt{k}}\right)\right) = \sigma_f^{1/2}\ a.s..$$

This extends earlier results of Csörgő and Horváth [5] and Horváth and Khoshnevisan [10] who proved (1.11) under the additional assumption $\mathbf{E}X_1^2(\log(|X_1|+1))^{1+\delta} < +\infty$ for some $\delta > 0$. We note also that in the case when $X_1, X_2, \dots$ are independent random variables with $\mathbf{P}(X_n{=}1) = \mathbf{P}(X_n = -1) = 1/2$ $(n = 1, 2, \dots)$ and f is an indicator function of an interval, (1.21) was proved by Weigl [18].

The following result relaxes the assumptions of Corollary 2.

Corollary 3. *Let $X_1, X_2, \dots$ be independent random variables with mean zero and variance* 1 *and set $S_n = X_1 + \dots + X_n$. Let $f \in \mathcal{L}$ and assume that (X_n) satisfies one of the following conditions:*

$$\text{(a)}\qquad \sup_n \mathbf{E}X_n^2 I(|X_n| \geqq t) \to 0 \qquad \text{as } t \to \infty,$$

$$\text{(b)}\qquad (S_l - S_k)/\sqrt{l-k} \xrightarrow{\mathcal{D}} N(0,1) \quad \text{as } l-k \to \infty,$$

$$\text{(c)}\qquad (S_n - W(n))/\sqrt{n} \xrightarrow{\mathbf{P}} 0 \quad \text{for some Wiener process } W.$$

Then the conclusion of Corollary 2 holds.

As a comparison, note that the proofs of the a.s. approximation theorems in [5], [10] use invariance techniques and require the existence of a Wiener process W such that

$$(1.22)\qquad S_n - W(n) = O(\sqrt{n}(\log n)^{-1-\varrho}) \quad \text{a.s.}$$

for some $\varrho > 0$. Condition (c) of Corollary 3 is similar to (1.22) but it is much weaker and, unlike (1.22), does not require the existence of any moments of X_n beyond the second.

We note that condition (a) implies condition (b) by an extension of the Berry–Esseen theorem (see [13], formula (5.26)).

COROLLARY 4. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F satisfying*

$$1 - F(x) \sim cx^{-\alpha} \quad \text{and} \quad F(-x) \sim cx^{-\alpha} \quad \text{as} \quad x \to \infty \tag{1.23}$$

where $c > 0$, $0 < \alpha < 2$. Let $f \in \mathcal{L} \cup \mathcal{I}$ and let λ_N be defined by (1.20) *with $a_k = k^{1/\alpha}$. Then $\sigma_f = \lim_{N\to\infty} (\log N)^{-1} \lambda_N$ exists and if $\sigma_f \neq 0$ then there exists a Wiener process W and a positive constant η such that* (1.11) *holds with $a_k = k^{1/\alpha}$.*

For additional information on σ_f (including criteria for $\sigma_f \neq 0$) we refer to [3].

2. Proofs

We begin with the proof of Theorem 1. Let (X_n), (a_n), f satisfy the assumptions of the theorem; without loss of generality we can assume that $|f| \leqq 1$.

LEMMA 1. *Let $r < p < q$ be positive integers and*

$$X = \sum_{i=2^p+1}^{2^q} \frac{1}{i} f\left(\frac{S_i}{a_i}\right), \qquad X' = \sum_{i=2^p+1}^{2^q} \frac{1}{i} f\left(\frac{S_i - S_{2^r}}{a_i}\right).$$

Then for any $d \geqq 1$ we have

$$\mathbf{E}|X - X'|^d \leqq A_1 \frac{1}{(p-r)^{\gamma_1}} (q-p)^d$$

where A_1 is a positive constant and $\gamma_1 = \min(\alpha, \beta)$.*

PROOF. Set

$$Q(i) = \left| f\left(\frac{S_i}{a_i}\right) - f\left(\frac{S_i - S_{2^r}}{a_i}\right) \right| \qquad \text{for } 2^p + 1 \leqq i \leqq 2^q.$$

Relation (1.8) implies that for any $2^p \leqq i \leqq 2^q$ we have $a_i / a_{2^r} \geqq C 2^{\gamma(p-r)}$ and thus by (1.6) we obtain

$$\mathbf{P}\left(\left| \frac{S_{2^r}}{a_i} \right| \geqq 2^{-\frac{\gamma}{2}(p-r)} \right) \leqq \mathbf{P}\left(\left| \frac{S_{2^r}}{a_{2^r}} \right| \geqq 2^{\frac{\gamma}{2}(p-r)} \right) \leqq \text{const} \cdot \frac{1}{(p-r)^{\alpha}}.$$

* Here, and in the sequel, constants may depend on f and the distribution of the sequence (X_n).

Hence by (1.7) and $|f| \leqq 1$ we get, setting $h = \text{const} \cdot 2^{-\frac{\gamma}{2}(p-r)}$ and observing that $h \geqq \text{const}/\sqrt{a_i}$ by the estimates above,

$$\mathbf{E}|Q(i)| \leqq \mathbf{E} \sup_{|t| \leqq h} \left| f\left(\frac{S_i}{a_i}\right) - f\left(\frac{S_i}{a_i} - t\right)\right| + \text{const} \cdot \frac{1}{(p-r)^{\alpha}} \leqq \text{const} \cdot \frac{1}{(p-r)^{\gamma_1}}.$$

Thus observing that by $|f| \leqq 1$ we have

$$|X - X'| \leqq 2 \sum_{i=2^p+1}^{2^q} \frac{1}{i} \leqq 2(q-p)$$

we obtain

$$\mathbf{E}|X - X'|^d \leqq \text{const} \cdot (q-p)^{d-1} \mathbf{E}|X - X'| \leqq$$

$$\leqq \text{const} \cdot (q-p)^{d-1} \frac{1}{(p-r)^{\gamma_1}} \sum_{i=2^p+1}^{2^q} \frac{1}{i} \leqq \text{const} \cdot \frac{1}{(p-r)^{\gamma_1}} (q-p)^d.$$

LEMMA 2. *Let*

$$\delta_k = \sum_{i=2^k+1}^{2^{k+1}} \frac{1}{i} \left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E} f\left(\frac{S_i}{a_i}\right)\right).$$

Then for any $M \geqq 1$, $N \geqq 1$ *we have*

$$\mathbf{E}\left(\sum_{k=M+1}^{M+N} \delta_k\right)^2 \leqq C_1 N, \tag{2.1}$$

$$\mathbf{E}\left(\sum_{k=M+1}^{M+N} \delta_k\right)^4 \leqq C_1 N^3 \tag{2.2}$$

for some positive constant C_1.

PROOF. We prove (2.2); the proof of (2.1) is similar (in fact simpler). Clearly

$$\mathbf{E}\left(\sum_{k=M+1}^{M+N} \delta_k\right)^4 = \sum_{k=M+1}^{M+N} \mathbf{E}\delta_k^4 + 6 \sum_{M+1 \leqq i < j \leqq M+N} \mathbf{E}\delta_i^2 \delta_j^2$$

$$+ 4 \sum_{\substack{M+1 \leqq i,j \leqq M+N \\ i \neq j}} \mathbf{E}\delta_i^3 \delta_j + 12 \sum_{M+1 \leqq i \neq j \neq k \leqq M+N} \mathbf{E}\delta_i^2 \delta_j \delta_k$$

$$+ 24 \sum_{M+1 \leqq i < j < k < l \leqq M+N} \mathbf{E}\delta_i \delta_j \delta_k \delta_l =: S^{(1)} + \cdots + S^{(5)}.$$

By $|f| \leqq 1$ we have $|\delta_\nu| \leqq 2$ and thus

$$S^{(1)} + S^{(2)} + S^{(3)} \leqq AN^2,$$

where A is a positive constant. Next we show

$$S^{(5)} \leqq A_1 N^3 \tag{2.3}$$

where A_1 is a positive constant. To this end we first prove that if $M+1 \leqq i < j < k < l \leqq M+N$ and at least one of $j-i$ and $l-k$ is $\geqq \sqrt{N}$, then

$$|\mathbf{E}(\delta_i \delta_j \delta_k \delta_l)| \leqq \mathrm{const} \cdot N^{-2}. \tag{2.4}$$

Assume, e.g., that $j - i \geqq \sqrt{N}$ and set

$$\delta^*_{j,i} = \sum_{\nu=2^j+1}^{2^{j+1}} \frac{1}{\nu} \left(f\left(\frac{S_\nu - S_{2^{i+1}}}{a_\nu} \right) - \mathbf{E} f\left(\frac{S_\nu - S_{2^{i+1}}}{a_\nu} \right) \right).$$

Using Lemma 1 we get

$$\mathbf{E}|\delta_j - \delta^*_{j,i}| \leqq \mathrm{const} \cdot \frac{1}{(j-i-1)^{\gamma_1}} \leqq \mathrm{const} \cdot N^{-\gamma_1/2} \leqq \mathrm{const} \cdot N^{-2}$$

since $\gamma_1 > 4$ by (1.10). Similar estimates hold for $\mathbf{E}|\delta_k - \delta^*_{k,i}|$, $\mathbf{E}|\delta_l - \delta^*_{l,i}|$ and thus

$$|\mathbf{E}(\delta_i \delta_j \delta_k \delta_l) - \mathbf{E}(\delta_i \delta^*_{j,i} \delta^*_{k,i} \delta^*_{l,i})| \leqq \mathrm{const} \cdot N^{-2} \tag{2.5}$$

for all $M+1 \leqq i < j < k < l \leqq M+N$ such that $j - i \geqq \sqrt{N}$. Observing that δ_i and $\delta^*_{j,i}\delta^*_{k,i}\delta^*_{l,i}$ are independent and $\mathbf{E}\delta_i = 0$, we see that the second expectation in (2.5) equals 0 and thus (2.4) is valid in the case $j - i \geqq \sqrt{N}$. The case $l - k \geqq \sqrt{N}$ can be treated similarly.

Relation (2.4) implies that the contribution of those terms in $S^{(5)}$ where at least one of $j-i$ and $l-k$ is greater than $\sqrt{N}$ is at most $\mathrm{const} \cdot N^2$. On the other hand, the contribution of the remaining terms is less than $\mathrm{const} \cdot N^3$ since the number of 4-tuples (i,j,k,l) satisfying $M+1 \leqq i < j < k < l \leqq M+N$, $j - i \leqq \sqrt{N}$, $l - k \leqq \sqrt{N}$ is clearly at most N^3. Hence we proved (2.3); a similar argument applies for $S^{(4)}$ and thus Lemma 2 is proved.

Let us divide $[1, \infty)$ into consecutive intervals $\Delta_1 = [p_1, q_1]$, $\Delta'_1 = [p'_1, q'_1]$, $\Delta_2 = [p_2, q_2]$, $\Delta'_2 = [p'_2, q'_2], \ldots$ where $p_1 = 1$, $p'_k = q_k$ and $p_k = q'_{k-1}$. We choose these intervals so that

$$|\Delta_k| = [k^{1/2}], \qquad |\Delta'_k| = [k^\tau]$$

hold, where $|\Delta|$ denotes the length of the interval Δ and $1/4 \leqq \tau < 1/2$ satisfies

$$\min(\alpha,\beta) > 2/\tau, \qquad \delta > 2(\tau+1)/3. \tag{2.6}$$

In view of (1.10), $\tau = 1/4$ will do, but other values of τ are also of interest: for example, choosing τ close to 1/2 shows that for δ close enough to 1, the value 8 in (1.10) can be replaced by 4. Set

$$\xi_k = \sum_{i=2^{p_k}+1}^{2^{q_k}} \frac{1}{i} f\left(\frac{S_i}{a_i}\right), \qquad \xi_k^* = \sum_{i=2^{p_k}+1}^{2^{q_k}} \frac{1}{i} f\left(\frac{S_i - S_{2^{q_{k-1}}}}{a_i}\right)$$
$$\eta_k = \sum_{i=2^{p'_k}+1}^{2^{q'_k}} \frac{1}{i} f\left(\frac{S_i}{a_i}\right), \qquad \eta_k^* = \sum_{i=2^{p'_k}+1}^{2^{q'_k}} \frac{1}{i} f\left(\frac{S_i - S_{2^{q'_{k-1}}}}{a_i}\right).$$

By Lemma 1 we have for any integer $d \geqq 1$

$$\begin{aligned} \mathbf{E}|\xi_k - \xi_k^*|^d &\leqq \text{const} \cdot \frac{1}{(p_k - q_{k-1})^{\gamma_1}} (q_k - p_k)^d \\ &\leqq \text{const} \cdot k^{\frac{d}{2} - \gamma_1 \tau} \leqq \text{const} \cdot k^{\frac{d}{2} - 2 - \varepsilon} \end{aligned} \tag{2.7}$$

and similarly

$$\mathbf{E}|\eta_k - \eta_k^*|^d \leqq \text{const} \cdot k^{d\tau - \frac{\gamma_1}{2}} \leqq \text{const} \cdot k^{\frac{d}{2} - \gamma_1 \tau} \leqq \text{const} \cdot k^{\frac{d}{2} - 2 - \varepsilon} \tag{2.8}$$

for some $\varepsilon > 0$ since $\gamma_1 \tau > 2$ by (2.6).

Lemma 3. *We have*

$$\sum_{i=1}^{k} (\eta_i - \mathbf{E}\eta_i) = O(k^{(\tau+1)/2} \log k) \quad a.s.. \tag{2.9}$$

Proof. Applying (2.8) with $d = 1$ and using the monotone convergence theorem we get

$$\sum_{i=1}^{\infty} |(\eta_i - \mathbf{E}\eta_i) - (\eta_i^* - \mathbf{E}\eta_i^*)| < +\infty \quad \text{a.s..} \tag{2.10}$$

Also, (2.8) with $d = 2$ and Lemma 2 give

$$\|\eta_k^* - \mathbf{E}\eta_k^*\| \leqq \|\eta_k - \mathbf{E}\eta_k\| + O(1) \leqq \text{const} \cdot (q'_k - p'_k)^{1/2} \leqq \text{const} \cdot k^{\tau/2}$$

where $\|\cdot\|$ denotes the L_2 norm. Thus

$$\sum_{k=1}^{\infty} \frac{\mathbf{E}|\eta_k^* - \mathbf{E}\eta_k^*|^2}{k^{\tau+1}\log^2 k} < +\infty. \tag{2.11}$$

Since the r.v.'s η_k^* are independent with zero means, (2.11) implies that the series

$$\sum_{k=1}^{\infty} \frac{\eta_k^* - \mathbf{E}\eta_k^*}{k^{(\tau+1)/2}\log k}$$

is a.s. convergent and thus by the Kronecker lemma

$$\sum_{i=1}^{k}(\eta_i^* - \mathbf{E}\eta_i^*) = O(k^{(\tau+1)/2}\log k) \quad \text{a.s.}.$$

Together with (2.10), the last relation implies (2.9).

LEMMA 4. *Let* $N_k = \sum_{i=1}^{k}([i^{1/2}] + [i^\tau])$. *Then we have*

$$\sum_{i=1}^{k} \mathbf{E}((\xi_i^* - \mathbf{E}\xi_i^*)^2) = \lambda_{2^{N_k}}(1 + O(k^{-\varrho})) \tag{2.12}$$

for some $\varrho > 0$.

PROOF. Lemma 2 implies that

$$\mathbf{Var}\,\eta_i \leqq \text{const} \cdot i^\tau$$

and thus using (2.8) with $d = 2$ we get

$$\mathbf{Var}\,\eta_i^* \leqq \text{const} \cdot i^\tau.$$

Hence by the independence of the η_i^* it follows that

$$\mathbf{Var}\left(\sum_{i=1}^{k}\eta_i^*\right) \leqq \text{const} \cdot k^{\tau+1}. \tag{2.13}$$

Using (2.13), the Minkowski inequality and (2.7), (2.8) with $d = 2$ we see that the first two of the quantities

$$\mathbf{Var}^{1/2}\left(\sum_{i=1}^{k}\xi_i^*\right), \quad \mathbf{Var}^{1/2}\left(\sum_{i=1}^{k}(\xi_i^* + \eta_i^*)\right), \quad \mathbf{Var}^{1/2}\left(\sum_{i=1}^{k}(\xi_i + \eta_i)\right) \tag{2.14}$$

differ at most by $O(k^{(\tau+1)/2})$ while the second and third differ at most by $O(k^{(1-\varepsilon)/2})$. Since the third expression in (2.14) equals $\lambda_{2^{N_k}}^{1/2}$, we proved that

$$\mathbf{Var}^{1/2}\left(\sum_{i=1}^{k}\xi_i^*\right)=\lambda_{2^{N_k}}^{1/2}+O(k^{(\tau+1)/2})=\lambda_{2^{N_k}}^{1/2}(1+O(k^{-\varrho})) \tag{2.15}$$

for some $\varrho>0$, where the second equality follows by observing that by (1.9), (2.6) and $N_k\sim\text{const}\cdot k^{3/2}$ we have

$$\lambda_{2^{N_k}}^{1/2}\geqq\text{const}\cdot(\log 2^{N_k})^{\delta/2}\geqq\text{const}\cdot k^{3\delta/4}\geqq\text{const}\cdot k^{\frac{\tau+1}{2}(1+\varepsilon')} \tag{2.16}$$

for some $\varepsilon'>0$. Since ξ_k^* are independent, (2.15) implies (2.12).

LEMMA 5. *There exists a Wiener process W such that*

$$\sum_{i=1}^{k}(\xi_i-\mathbf{E}\xi_i)=W(\lambda_{2^{N_k}})+O(\lambda_{2^{N_k}}^{\frac{1}{2}-\eta})\quad a.s. \tag{2.17}$$

for some $\eta>0$.

PROOF. Let

$$Z_i=\xi_i^*-\mathbf{E}\xi_i^*,\qquad s_k^2=\sum_{i=1}^{k}\mathbf{E}((\xi_i^*-\mathbf{E}\xi_i^*)^2).$$

By Lemma 2 we have

$$\mathbf{E}(\xi_k-\mathbf{E}\xi_k)^4\leqq\text{const}\cdot(q_k-p_k)^3\leqq\text{const}\cdot k^{3/2} \tag{2.18}$$

and thus applying (2.7) with $d=4$ it follows that (2.18) remains valid if we replace $\xi_k-\mathbf{E}\xi_k$ by $\xi_k^*-\mathbf{E}\xi_k^*$. Thus

$$\mathbf{E}Z_k^4\leqq\text{const}\cdot k^{3/2}.$$

Also, by Lemma 4,

$$s_k^2=\lambda_{2^{N_k}}(1+O(k^{-\varrho})). \tag{2.19}$$

Thus using (1.9) we get for any $0<\vartheta<1$, sufficiently close to 1,

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{s_k^{2\vartheta}}\int_{x^2>s_k^{2\vartheta}}x^2d\mathbf{P}(Z_k<x)&\leqq\sum_{k=1}^{\infty}\frac{1}{s_k^{4\vartheta}}\int_{-\infty}^{+\infty}x^4d\mathbf{P}(Z_k<x)\\&=\sum_{k=1}^{\infty}\frac{1}{s_k^{4\vartheta}}\mathbf{E}(Z_k^4)\leqq\text{const}\cdot\sum_{k=1}^{\infty}\frac{k^{3/2}}{\lambda_{2^{N_k}}^{2\vartheta}}\\&\leqq\text{const}\cdot\sum_{k=1}^{\infty}\frac{k^{3/2}}{N_k^{2\vartheta\delta}}\leqq\text{const}\cdot\sum_{k=1}^{\infty}\frac{k^{3/2}}{k^{3\vartheta\delta}}<+\infty\end{aligned}$$

since the second relation of (2.6) and $\tau \geqq 1/4$ imply $\delta > 5/6$ and thus we have $3\vartheta\delta > 5/2$ if ϑ is sufficiently close to 1. Thus using an a.s. invariance principle of Strassen ([17], Theorem 4.4) we get

$$\sum_{i=1}^{k}(\xi_i^* - \mathbf{E}\xi_i^*) = W(s_k^2) + O(s_k^{(1+\vartheta)/2}\log s_k) \quad \text{a.s.} \tag{2.20}$$

with some Wiener process W. Now by (2.19), the relation $\text{const}\cdot(\log N)^\delta \leqq \lambda_N \leqq \text{const}\cdot(\log N)$ (cf. (1.9) and Lemma 2) and well-known properties of Wiener fluctuations (see, e.g., Csörgő and Révész [6], Theorem 1.2.1) we have

$$|W(s_k^2) - W(\lambda_{2^{N_k}})| = O(\lambda_{2^{N_k}}^{\frac{1}{2}-\eta}) \quad \text{a.s.} \tag{2.21}$$

for some constant $\eta > 0$. Also, (2.7) with $d = 1$ and the monotone convergence theorem imply

$$\sum_{i=1}^{\infty}|(\xi_i - \mathbf{E}\xi_i) - (\xi_i^* - \mathbf{E}\xi_i^*)| < +\infty \qquad \text{a.s.}. \tag{2.22}$$

Now (2.17) follows from (2.20), (2.21), (2.22), (2.19) and $\vartheta < 1$.

We can now easily complete the proof of Theorem 1. By Lemma 3 and Lemma 5 we get

$$\begin{aligned}
&\sum_{i=1}^{N_k}\frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E}f\left(\frac{S_i}{a_i}\right)\right)\\
&= W(\lambda_{2^{N_k}}) + O(\lambda_{2^{N_k}}^{1/2-\eta}) + O(k^{(\tau+1)/2}\log k)\\
&= W(\lambda_{2^{N_k}}) + O(\lambda_{2^{N_k}}^{1/2-\eta'}) \quad \text{a.s.}
\end{aligned} \tag{2.23}$$

for some $\eta' > 0$ by using (2.16). Now if $2^{N_k} \leqq N < 2^{N_{k+1}}$, then the expression

$$\sum_{i=1}^{N}\frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E}f\left(\frac{S_i}{a_i}\right)\right)$$

differs from its value at $N = 2^{N_k}$ by at most

$$\begin{aligned}
&O\left(\sum_{i=2^{N_k}}^{2^{N_{k+1}}}\frac{1}{i}\right) = O(N_{k+1} - N_k) = O(k^{1/2}) = O(N_k^{1/3})\\
&= O((\log N)^{1/3}) = O(\lambda_N^{1/2-\eta''})
\end{aligned} \tag{2.24}$$

for some $\eta'' > 0$, where in the last step we used (1.9) and (1.10). Also, Minkowski's inequality and (2.24) imply for $2^{N_k} \leqq N \leqq 2^{N_{k+1}}$

$$|\lambda_N^{1/2} - \lambda_{2^{N_k}}^{1/2}| = O\left(\sum_{i=2^{N_k}}^{2^{N_k+1}} \frac{1}{i}\right) = O(\lambda_N^{1/2-\eta''}) \tag{2.25}$$

and thus we have $1/2 \leqq \lambda_N/\lambda_{2^{N_k}} \leqq 2$ for $k \geqq k_0$. Hence (2.25) yields

$$|\lambda_N - \lambda_{2^{N_k}}| = O(\lambda_N^{1-\eta''})$$

and thus using again Csörgő and Révész [6], Theorem 1.2.1 we get

$$|W(\lambda_N) - W(\lambda_{2^{N_k}})| = O(\lambda_N^{1/2-\eta'''}) \quad \text{a.s.}$$

for some $\eta''' > 0$. Thus (2.23) implies (1.11), completing the proof of Theorem 1.

PROOF OF THEOREM 2. Since conditions (1.16) and (1.17) of Theorem 2 imply, respectively, conditions (1.6) and (1.7) of Theorem 1, in order to prove Theorem 2 it suffices to prove that the conditions of Theorem 2 imply (1.9). This statement is contained in Lemma 8 below; for the proof we need two auxiliary lemmas.

LEMMA 6. *Assume the conditions of Theorem 2 and let $r < p < q$ be positive integers. Put*

$$X = \sum_{i=2^p+1}^{2^q} \frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E}f\left(\frac{S_i}{a_i}\right)\right)$$

$$X' = \sum_{i=2^p+1}^{2^q} \frac{1}{i}\left(f\left(\frac{S_i - S_{2^r}}{a_i}\right) - \mathbf{E}f\left(\frac{S_i - S_{2^r}}{a_i}\right)\right).$$

Then

$$\mathbf{E}|X - X'|^2 \leqq c_1 2^{-c_2(p-r)} \tag{2.26}$$

where c_1 and c_2 are positive constants.

PROOF. Our argument is similar to that used in the proof of Lemma 1. Put

$$Y = \sum_{i=2^p+1}^{2^q} \frac{1}{i} f\left(\frac{S_i}{a_i}\right), \qquad Y' = \sum_{i=2^p+1}^{2^q} \frac{1}{i} f\left(\frac{S_i - S_{2^r}}{a_i}\right).$$

Since $X - X' = (Y - Y') - \mathbf{E}(Y - Y')$, it suffices to show that

$$\mathbf{E}|Y - Y'|^2 \leqq c_1 \cdot 2^{-c_2(p-r)}. \tag{2.27}$$

By (1.16) we have

$$\mathbf{E}\left|\frac{S_{2^r}}{a_i}\right|^p \leqq \text{const}\cdot\left(\frac{a_{2^r}}{a_i}\right)^p$$

and thus by the Markov inequality

$$\mathbf{P}\left(\left|\frac{S_{2^r}}{a_i}\right| \geqq \sqrt{\frac{a_{2^r}}{a_i}}\right) \leqq \text{const}\cdot\left(\frac{a_{2^r}}{a_i}\right)^{p/2}.$$

Hence letting

$$Q(i) = f\left(\frac{S_i}{a_i}\right) - f\left(\frac{S_i - S_{2^r}}{a_i}\right), \qquad h = (a_{2^r}/a_i)^{1/2}$$

we get, using $|f| \leqq 1$, (1.17) and the fact that $h \geqq \text{const}/\sqrt{a_i}$,

$$\begin{aligned}
&\mathbf{E}|Q(i)|^2 \leqq 2\mathbf{E}|Q(i)| \\
&\leqq 2\mathbf{E}\sup_{|t|\leqq h}\left|f\left(\frac{S_i}{a_i}\right) - f\left(\frac{S_i}{a_i} - t\right)\right| + \text{const}\cdot\left(\frac{a_{2^r}}{a_i}\right)^{p/2} \qquad (2.28)\\
&\leqq \text{const}\cdot\left(\frac{a_{2^r}}{a_i}\right)^{\varrho}
\end{aligned}$$

for some constant $\varrho > 0$. Now by (2.28)

$$\begin{aligned}
\mathbf{E}|Y - Y'|^2 &\leqq \sum_{i=2^p+1}^{2^q}\sum_{j=2^p+1}^{2^q}\frac{1}{ij}\mathbf{E}|Q(i)Q(j)| \\
&\leqq \sum_{i=2^p+1}^{2^q}\sum_{j=2^p+1}^{2^q}\frac{1}{ij}(\mathbf{E}Q(i)^2)^{1/2}(\mathbf{E}Q(j)^2)^{1/2} \qquad (2.29)\\
&\leqq \text{const}\cdot a_{2^r}^{\varrho}\sum_{i=2^p+1}^{2^q}\sum_{j=2^p+1}^{2^q}\frac{1}{ij a_i^{\varrho/2} a_j^{\varrho/2}} = \text{const}\cdot a_{2^r}^{\varrho}\left(\sum_{i=2^p+1}^{2^q}\frac{1}{i a_i^{\varrho/2}}\right)^2.
\end{aligned}$$

Relation (1.18) implies that for any $n \geqq 1$

$$\min_{2^n \leqq i \leqq 2^{n+1}} a_i \geqq C a_{2^n}$$

and thus

$$\sum_{i=2^n+1}^{2^{n+1}}\frac{1}{i a_i^{\varrho/2}} \leqq \text{const}\cdot\frac{1}{a_{2^n}^{\varrho/2}}\sum_{i=2^n+1}^{2^{n+1}}\frac{1}{i} \leqq \text{const}\cdot\frac{1}{a_{2^n}^{\varrho/2}}.$$

Hence using (1.18) once more we get

$$\sum_{i=2^p+1}^{2^q} \frac{1}{i a_i^{\varrho/2}} \leqq \text{const} \cdot \sum_{n=p}^{q} \frac{1}{a_{2^n}^{\varrho/2}}$$
$$\leqq \text{const} \cdot \frac{1}{a_{2^p}^{\varrho/2}} \sum_{n=p}^{q} \frac{1}{(2^{n-p})^{\gamma \cdot \varrho/2}} \leqq \text{const} \cdot \frac{1}{a_{2^p}^{\varrho/2}}.$$

Substituting this into (2.29) and using (1.18) we get

$$\mathbf{E}|Y-Y'|^2 \leqq \text{const} \cdot \left(\frac{a_{2^r}}{a_{2^p}}\right)^{\varrho} \leqq c_1 2^{-c_2(p-r)}$$

proving (2.27).

LEMMA 7. *Assume the conditions of Theorem 2 and let* $m<n<p<q$ *be positive integers. Set*

$$\xi = \sum_{i=2^m+1}^{2^n} \frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E} f\left(\frac{S_i}{a_i}\right)\right)$$
$$\eta = \sum_{i=2^p+1}^{2^q} \frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right) - \mathbf{E} f\left(\frac{S_i}{a_i}\right)\right).$$

Then

$$|\mathbf{E}(\xi\eta)| \leqq c_3 \|\xi\| 2^{-c_4(p-n)}$$

where c_3, c_4 *are positive constants.*

PROOF. This follows immediately from the previous lemma. Put

$$\eta' = \sum_{i=2^p+1}^{2^q} \frac{1}{i}\left(f\left(\frac{S_i - S_{2^n}}{a_i}\right) - \mathbf{E} f\left(\frac{S_i - S_{2^n}}{a_i}\right)\right).$$

Clearly ξ and η' are independent and $\mathbf{E}\xi = \mathbf{E}\eta' = 0$. Thus by Lemma 6

$$|\mathbf{E}(\xi\eta)| = |\mathbf{E}(\xi(\eta-\eta'))| \leqq \|\xi\| \, \|\eta-\eta'\| \leqq c_3 \|\xi\| \cdot 2^{-c_4(p-n)}.$$

LEMMA 8. *Assume the conditions of Theorem 2. Then for any* $0<\varepsilon<1$ *there exists a constant* $c>0$ *such that*

(2.30) $$\mathbf{Var} \sum_{i=2^M+1}^{2^{M+N}} \frac{1}{i} f\left(\frac{S_i}{a_i}\right) \geqq c N^{1-\varepsilon} \quad \textit{for any } M \geqq 0,\ N \geqq 1.$$

PROOF. Clearly if (1.19) holds then it remains valid if we replace $\omega(N)$ by any $0<\omega_1(N)\leqq\omega(N)$. Thus without loss of generality we may assume that ω is nondecreasing and slowly varying. Hence given $0<\varepsilon<1$ we can choose $N_0=N_0(\varepsilon)$ so large that

$$2\omega(N)^{1/4}/\omega([N/3])^{1/2}\leqq\varepsilon^2,\qquad \omega(N)^{1/4}\leqq\varepsilon^2N/6\qquad (N\geqq N_0) \tag{2.31}$$

and

$$\omega([N/3])\geqq 1,\qquad c_3\cdot 2^{-c_4[\omega(N)^{1/4}]}\leqq\varepsilon^2\qquad (N\geqq N_0), \tag{2.32}$$

where c_3 and c_4 are the constants in Lemma 7. With N_0 chosen, choose $c>0$ so small that

$$\omega(k)\geqq ck^{1-\varepsilon}\qquad \text{for } 1\leqq k\leqq N_0. \tag{2.33}$$

We shall prove by induction on N that (2.30) holds for all $M\geqq 0$, $N\geqq 1$.

By (2.33) and (1.19), relation (2.30) is valid for all $M\geqq 0$ and $1\leqq\leqq N\leqq N_0$. Let now $N>N_0$ and assume that (2.30) is valid if N is replaced by any $1\leqq N'<N$ and $M\geqq 1$ is arbitrary. Assume, e.g., that N is even; the argument is similar if N is odd. Put

$$\xi_i=\frac{1}{i}\left(f\left(\frac{S_i}{a_i}\right)-\mathbf{E}f\left(\frac{S_i}{a_i}\right)\right)$$

then

$$S:=\sum_{i=2^M+1}^{2^{M+N}}\xi_i=S_1+S_2+S_3,$$

where

$$S_1=\sum_{i=2^M+1}^{2^{M+N/2-r(N)}}\xi_i,\quad S_2=\sum_{i=2^{M+N/2-r(N)}+1}^{2^{M+N/2+r(N)}}\xi_i,\quad S_3=\sum_{i=2^{M+N/2+r(N)}+1}^{2^{M+N}}\xi_i$$

with

$$r(N)=[\omega(N)^{1/4}]. \tag{2.34}$$

By the induction hypothesis, (2.34) and the second relation of (2.31) we have

$$\mathbf{E}S_1^2\geqq c\left(\frac{N}{2}-r(N)\right)^{1-\varepsilon}\geqq c\left(\frac{1-\varepsilon^2}{2}\right)^{1-\varepsilon}N^{1-\varepsilon}$$

and similarly

$$\mathbf{E}S_3^2\geqq c\left(\frac{1-\varepsilon^2}{2}\right)^{1-\varepsilon}N^{1-\varepsilon}.$$

On the other hand, using $|f| \leqq 1$, the Minkowski inequality, (2.31), (2.34) and the fact that $\sum_{i=2^n+1}^{2^{n+1}} 1/i \leqq 1$ for all $n \geqq 1$, we get

$$\begin{aligned}\|S_2\| &\leqq \sum_{i=2^{M+N/2-r(N)}+1}^{2^{M+N/2+r(N)}} \frac{2}{i} \leqq 2r(N) \leqq 2\omega(N)^{1/4} \\ &\leqq \varepsilon^2\omega\left([N/3]\right)^{1/2} \leqq \varepsilon^2\|S_1\|,\end{aligned} \tag{2.35}$$

where the last inequality follows from (1.19) and $N/2 - r(N) \geqq N/3$ (which is valid by (2.34) and the second relation of (2.31)). Thus

$$|\mathbf{E}S_1S_2| \leqq \|S_1\|\,\|S_2\| \leqq \varepsilon^2\mathbf{E}S_1^2$$

and similarly

$$|\mathbf{E}S_3S_2| \leqq \varepsilon^2\mathbf{E}S_3^2.$$

Finally Lemma 7, (2.32) and (2.34) imply

$$|\mathbf{E}S_1S_3| \leqq c_3\|S_1\|2^{-c_4r(N)} \leqq \varepsilon^2\|S_1\| \leqq \varepsilon^2\|S_1\|^2$$

since $\|S_1\| \geqq 1$ by the last inequality of (2.35) and (2.32). Collecting now all our estimates, we get

$$\begin{aligned}\mathbf{E}S^2 &\geqq \mathbf{E}S_1^2 + \mathbf{E}S_3^2 - |2\mathbf{E}S_1S_2| - |2\mathbf{E}S_2S_3| - |2\mathbf{E}S_1S_3| \\ &\geqq \mathbf{E}S_1^2 + \mathbf{E}S_3^2 - 2\varepsilon^2\mathbf{E}S_1^2 - 2\varepsilon^2\mathbf{E}S_3^2 - 2\varepsilon^2\mathbf{E}S_1^2 \\ &\geqq (1-4\varepsilon^2)(\mathbf{E}S_1^2 + \mathbf{E}S_3^2) \geqq 2c\left(\frac{1-\varepsilon^2}{2}\right)^{1-\varepsilon}(1-4\varepsilon^2)N^{1-\varepsilon} \\ &\geqq cN^{1-\varepsilon}\end{aligned} \tag{2.36}$$

provided

$$2\left(\frac{1-\varepsilon^2}{2}\right)^{1-\varepsilon}(1-4\varepsilon^2) > 1$$

which is true for $0 < \varepsilon \leqq \varepsilon_0$ since the derivative of the left-hand side at $\varepsilon = 0$ is $\ln 2 > 0$. (2.36) completes the induction step and thus we proved the lemma for $0 < \varepsilon \leqq \varepsilon_0$; clearly this implies the lemma for all $0 < \varepsilon < 1$.

PROOF OF COROLLARY 1. The approximation (1.11) trivially implies (1.12) and by the LIL for W it also implies the inequality $\leqq 1$ in (1.13). To prove the inequality $\geqq 1$ in (1.13) note that by the Minkowski inequality and $|f| \leqq 1$ we have $|\lambda_{N+1}^{1/2} - \lambda_N^{1/2}| \leqq 2$ which, together with (1.9), implies that $\lambda_{N+1}/\lambda_N \to 1$. Thus for any $a > 1$ there exists a sequence N_k of integers such that $\lambda_{N_k} \sim a^k$. By the standard proof of the LIL for Wiener process we have

$$\limsup_{N\to\infty} (2\lambda_{N_k} \log\log \lambda_{N_k})^{-1/2} W(\lambda_{N_k}) \geqq 1 - \varepsilon(a) \qquad \text{a.s.},$$

where $\varepsilon(a) \to 0$ if $a \to \infty$. This implies the inequality $\geqq 1$ in (1.13).

PROOF OF COROLLARY 2. We will verify the conditions of Theorem 1 with $a_n = \sqrt{n}$. Since (1.16) and (1.17) imply (1.6) and (1.7), respectively, it suffices to check (1.16), (1.17) and (1.9). Clearly (1.16) holds with $p = 2$ and (1.17) is also obvious if $f \in \mathcal{L}$. To verify (1.17) for $f \in \mathcal{I}$ we apply a concentration function inequality of Esseen (see, e.g., [13], Theorem 2.14) to get

$$\sup_a \mathbf{P}\left(a < \frac{S_n}{\sqrt{n}} < a + h\right) \leqq \frac{Ah}{\left(\int\limits_{|x| \leqq h\sqrt{n}} x^2 dF^s(x)\right)^{1/2}}, \tag{2.37}$$

where A is an absolute constant and F^s is the distribution function obtained from the distribution function F of X_1 by symmetrization. Since $\int\limits_{-\infty}^{+\infty} x^2 dF^s(x) = 2$, there exists a constant $c > 0$ such that the integral in (2.37) is at least 1 for $h \geqq c/\sqrt{n}$, but then (2.37) implies (1.17) for $f \in \mathcal{I}$ with $\beta = 1$. To complete the proof of Corollary 2 it remains now to prove the following

LEMMA 9. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and let $f \in \mathcal{L} \cup \mathcal{I}$. Then*

$$\lambda_N = \mathbf{Var} \sum_{k \leqq N} \frac{1}{k} f\left(\frac{S_k}{\sqrt{k}}\right) \sim \sigma_f \log N \quad \textit{as} \quad N \to \infty,$$

where σ_f is a positive constant depending on f.

PROOF. From an a.s. invariance principle of Major (see [12]) it follows that there exists a Wiener process W such that

$$(S_n - W(n))/\sqrt{n} \to 0 \quad \text{in probability.} \tag{2.38}$$

Also, the results of [10] imply that for any $f \in \mathcal{L} \cup \mathcal{I}$ (and in fact for a larger class of functions f defined by conditions C1–C4 there)

$$\lambda_N^* := \mathbf{Var} \sum_{k \leqq N} \frac{1}{k} f\left(\frac{W(k)}{\sqrt{k}}\right) \sim \sigma_f \log N \quad \text{as} \quad N \to \infty \tag{2.39}$$

for some constant $\sigma_f > 0$ depending on f. (Note that $\lambda_N^* = 0$ if f is constant, but this case is excluded in both $\mathcal{L}$ and $\mathcal{I}$.) Thus it suffices to prove that

$$\lim_{N\to\infty} \lambda_N^*/\lambda_N = 1. \tag{2.40}$$

To prove (2.40) we first note that above we verified (1.17) for any $f \in \mathcal{L} \cup \mathcal{I}$ and $a_n = \sqrt{n}$; observe also that

$$\mathbf{P}(|S_k/\sqrt{l}| \geqq h) \leqq h \qquad \text{for} \quad k < l, \quad h = (k/l)^{1/4}$$

by the Chebyshev inequality. Thus using (1.17) and $|f| \leqq 1$ we get for any $k < l$

$$\begin{aligned} &\mathbf{E}\left| f\left(\frac{S_l}{\sqrt{l}}\right) - f\left(\frac{S_l - S_k}{\sqrt{l}}\right)\right| \leqq \mathbf{E} \sup_{|t|\leqq h} \left| f\left(\frac{S_l}{\sqrt{l}}\right) - f\left(\frac{S_l}{\sqrt{l}} - t\right)\right| + 2h \\ &\leqq \text{const} \cdot (h^\beta + h) \leqq \text{const} \cdot \left(\frac{k}{l}\right)^\alpha \end{aligned} \tag{2.41}$$

for some constants $\alpha > 0$, $\beta > 0$. Since for $k < l$ the random variables S_k and $S_l - S_k$ are independent, using (2.41) and $|f| \leqq 1$ we get

$$\begin{aligned} &\left|\mathbf{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)\right| \\ &= \left|\mathbf{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right) - f\left(\frac{S_l - S_k}{\sqrt{l}}\right)\right)\right| \\ &\leqq 2\mathbf{E}\left| f\left(\frac{S_l}{\sqrt{l}}\right) - f\left(\frac{S_l - S_k}{\sqrt{l}}\right)\right| \leqq \text{const} \cdot \left(\frac{k}{l}\right)^\alpha. \end{aligned} \tag{2.42}$$

Fix $\varepsilon > 0$ and set

$$\begin{aligned} c_{k,l} &= \mathbf{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right), \\ c_{k,l}^* &= \mathbf{Cov}\left(f\left(\frac{W(k)}{\sqrt{k}}\right), f\left(\frac{W(l)}{\sqrt{l}}\right)\right). \end{aligned} \tag{2.43}$$

Relation (2.38) implies that $c_{k,l} - c_{k,l}^* \to 0$ if $\min(k,l) \to \infty$ and thus there exists an integer $A = A(\varepsilon)$ such that

$$|c_{k,l} - c_{k,l}^*| \leqq \varepsilon^2 \quad \text{for} \quad A \leqq k < l. \tag{2.44}$$

Set also $B = [1/\varepsilon]$, then for $N \geqq A$ we get

$$\lambda_{N,A} := \mathbf{Var} \sum_{k=A}^{N} \frac{1}{k} f\left(\frac{S_k}{\sqrt{k}}\right)$$

$$= \sum_{k=A}^{N} \frac{1}{k^2} \mathbf{Var}\, f\left(\frac{S_k}{\sqrt{k}}\right) + 2 \sum_{\substack{A \leqq k < l \leqq N \\ l/k \leqq B}} \frac{1}{kl} \mathbf{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)$$

$$+ 2 \sum_{\substack{A \leqq k < l \leqq N \\ l/k > B}} \frac{1}{kl} \mathbf{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)$$

$$=: \lambda_N^{(1)} + \lambda_N^{(2)} + \lambda_N^{(3)}.$$

Clearly $\lambda_N^{(1)} = O(1)$ and by (2.42)

$$\lambda_N^{(3)} \leqq \text{const} \cdot \sum_{\substack{A \leqq k < l \leqq N \\ l/k > B}} \frac{1}{k^{1-\alpha} l^{1+\alpha}}$$

$$\leqq \text{const} \cdot \sum_{k=A}^{N} \frac{1}{k^{1-\alpha}} \sum_{l=Bk+1}^{\infty} \frac{1}{l^{1+\alpha}} \leqq \text{const} \cdot \sum_{k=A}^{N} \frac{1}{B^\alpha k}$$

$$\leqq \text{const} \cdot B^{-\alpha} \log N.$$

Thus we get for $N \geqq N_0$

$$\left|\lambda_{N,A} - 2 \sum_{\substack{A \leqq k < l \leqq N \\ l/k \leqq B}} \frac{1}{kl} c_{k,l}\right| \leqq \text{const} \cdot B^{-\alpha} \log N. \tag{2.45}$$

In a similar fashion we get for $N \geqq N_0$

$$\left|\lambda^*_{N,A} - 2 \sum_{\substack{A \leqq k < l \leqq N \\ l/k \leqq B}} \frac{1}{kl} c^*_{k,l}\right| \leqq \text{const} \cdot B^{-\alpha} \log N, \tag{2.46}$$

where $\lambda^*_{N,A}$ is defined similarly as $\lambda_{N,A}$, just with S_k replaced by $W(k)$. Now by (2.44)

$$\sum_{\substack{A \leqq k < l \leqq N \\ l/k \leqq B}} \frac{1}{kl} |c_{k,l} - c^*_{k,l}| \leqq \varepsilon^2 \sum_{\substack{A \leqq k < l \leqq N \\ l/k \leqq B}} \frac{1}{kl} \leqq \varepsilon^2 \sum_{k=A}^{N} \frac{1}{k} \sum_{k \leqq l \leqq Bk} \frac{1}{l} \leqq \varepsilon^2 \sum_{k=A}^{N} \frac{1}{k} \frac{Bk}{k}$$

$$\leqq \text{const} \cdot B\varepsilon^2 \log N$$

which, together with (2.45) and (2.46), gives

$$|\lambda_{N,A} - \lambda^*_{N,A}| \leqq \text{const} \cdot (\varepsilon + \varepsilon^\alpha) \log N \qquad \text{for } N \geqq N_0(\varepsilon). \tag{2.47}$$

By Minkowski's inequality

(2.48) $$|\lambda_N^{1/2} - \lambda_{N,A}^{1/2}| = O(1), \quad |(\lambda_N^*)^{1/2} - (\lambda_{N,A}^*)^{1/2}| = O(1).$$

Now (2.40) follows from (2.39), (2.47) and (2.48).

PROOF OF COROLLARY 3. In the proof of Corollary 2 above the identical distribution of the X_n was used only at two places: to verify (1.17) and to guarantee (2.38) which was needed, in turn, to show that

(2.49) $$c_{k,l} - c_{k,l}^* \to 0 \quad \text{as} \quad \min(k,l) \to \infty.$$

Since in Corollary 3 we assume $f \in \mathcal{L}$ under which (1.17) is trivially valid, it suffices to prove that under the conditions of Corollary 3 we have (2.49), where $c_{k,l}$ and $c_{k,l}^*$ are defined in (2.43) and W is an arbitrary Wiener process.

Clearly, (2.49) is valid under condition (c) of Corollary 3 and we also note that condition (a) of Corollary 3 implies condition (b) by a generalization of the Berry–Esseen inequality (see [13], relation (5.26)). Thus it remains to verify (2.49) under condition (b). To see this let us observe that the characteristic function $\varphi_{k,l}$ of the vector $(S_k/\sqrt{k}, S_l/\sqrt{l})$ $(k < l)$ can be written as

$$\varphi_{k,l}(t,u) = \mathbf{E} \exp\left(i\left(t + u\sqrt{\frac{k}{l}}\right)\frac{S_k}{\sqrt{k}}\right) \mathbf{E} \exp\left(iu\sqrt{\frac{l-k}{l}}\frac{S_l - S_k}{\sqrt{l-k}}\right)$$

and a similar formula holds for the characteristic function $\varphi_{k,l}^*$ of the vector $(W(k)/\sqrt{k},\ W(l)/\sqrt{l})$. Using this observation, condition (b) and the fact that the variance of the r.v. $(S_l - S_k)/\sqrt{l-k}$ is 1, we get easily that

(2.50) $$\sup_{|t| \leq C, |u| \leq C} |\varphi_{k,l}(t,u) - \varphi_{k,l}^*(t,u)| \longrightarrow 0 \quad \text{as} \quad k < l,\ k \to \infty$$

for any $C > 0$. Using a two-dimensional version of Esseen's inequality (see [14]), (2.50) implies that

(2.51) $$\sup_{x,y} |F_{k,l}(x,y) - F_{k,l}^*(x,y)| \to 0 \quad \text{as} \quad k < l,\ k \to \infty,$$

where $F_{k,l}$ and $F_{k,l}^*$ are the distribution functions corresponding to $\varphi_{k,l}$ and $\varphi_{k,l}^*$. Also, by Chebyshev's inequality

(2.52) $$\iint_{(x,y) \notin A(T)} dF_{k,l}(x,y) \leq 2T^{-2}, \qquad \iint_{(x,y) \notin A(T)} dF_{k,l}^*(x,y) \leq 2T^{-2}$$

for any $T > 0$, where $A(T)$ is the square $\{|x| \leq T, |y| \leq T\}$. Now (2.49) follows from (2.51), (2.52) and integration by parts.

PROOF OF COROLLARY 4. We first show that (X_n) satisfies (1.16) and (1.17) with $a_n = n^{1/\alpha}$. Let $Y_{\alpha,c}$ be a symmetric stable random variable with characteristic function $\exp(-\varrho|t|^\alpha)$, $\varrho = 2c\int_0^\infty y^{-\alpha}\sin y\,dy$. Relation (1.23) implies (see, e.g., [8], p. 544) that X_n are in the domain of normal attraction of $Y_{\alpha,c}$, i.e., $S_n/n^{1/\alpha} \xrightarrow{\mathcal{D}} Y_{\alpha,c}$. Hence [1], Theorem 6.1 implies

$$\sup_n \mathbf{E}\left|\frac{S_n}{n^{1/\alpha}}\right|^p < +\infty \quad \text{for any} \quad p<\alpha, \tag{2.53}$$

i.e., (1.16) is valid. For $f \in \mathcal{L}$ relation (1.17) is obvious; to verify (1.17) for $f \in \mathcal{I}$ we use the concentration function inequality in [13], Theorem 2.14 to get

$$\sup_a \mathbf{P}\left(a < \frac{S_n}{n^{1/\alpha}} < a+h\right) \leqq \frac{Ahn^{1/\alpha}}{\left(n \int_{|x|\leqq hn^{1/\alpha}} x^2 dF^s(x)\right)^{1/2}}, \tag{2.54}$$

where A is an absolute constant and F^s is the distribution function obtained from F by symmetrization. Relation (1.23) implies (see [8], p. 271)

$$1 - F^s(x) \sim 2cx^{-\alpha} \quad \text{as} \quad x \to \infty$$

whence we get by integration by parts

$$\int_{|x|\leqq t} x^2 dF^s(x) \sim \text{const}\cdot t^{2-\alpha} \qquad \text{as} \quad t\to\infty$$

and thus for $h \geqq a_n^{-1/2} = n^{-1/2\alpha}$ the right-hand side of (2.54) is $\leqq \text{const}\cdot h^{\alpha/2}$. Since for $f \in \mathcal{I}$ the left-hand side of (1.17) is bounded by the left-hand side of (2.54), relation (1.17) is valid for $f \in \mathcal{I}$.

By Theorem 3 of Simons and Stout [16] there exists a symmetric stable process $V_{\alpha,c}$ (i.e., a process with independent increments satisfying $V_{\alpha,c}(0)=0$ and $V_{\alpha,c}(t) - V_{\alpha,c}(s) \overset{\mathcal{D}}{=} (t-s)^{1/\alpha} Y_{\alpha,c}$ for all $0 \leqq s < t < +\infty$) such that

$$(S_n - V_{\alpha,c}(n))/n^{1/\alpha} \to 0 \qquad \text{in probability.} \tag{2.55}$$

Let

$$\lambda_N := \mathbf{Var} \sum_{k\leqq N} \frac{1}{k} f\left(\frac{S_k}{k^{1/\alpha}}\right), \quad \lambda_N^* := \mathbf{Var} \sum_{k\leqq N} \frac{1}{k} f\left(\frac{V_{\alpha,c}(k)}{k^{1/\alpha}}\right).$$

In [3] it is proved that the limit

$$\sigma_f = \lim_{N\to\infty} (\log N)^{-1} \lambda_N^* \tag{2.56}$$

exists. Using (2.55) instead of (2.38), the proof of Lemma 9 can be repeated to give

$$\lambda_N^* - \lambda_N = o(\log N) \quad \text{as} \quad N \to \infty,$$

and thus if $\sigma_f \neq 0$ then (1.9) holds. Hence Corollary 4 follows from Theorem 1.

REFERENCES

[1] De Acosta, A. and Giné, E., Convergence of moments and related functionals in the general central limit theorem in Banach spaces, *Z. Wahrsch. Verw. Gebiete* **48** (1979), 213–231. *MR* **80h**:60011

[2] Berkes, I. and Dehling, H., Some limit theorems in log density, *Ann. Probab.* **21** (1993), 1640–1670. *MR* **94h**:60026

[3] Berkes, I., Horváth, L. and Khoshnevisan, D., Logarithmic averages of stable random variables are asymptotically normal (in preparation).

[4] Brosamler, G., An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.* **104** (1988), 561–574. *MR* **89i**:60045

[5] Csörgő, M. and Horváth, L., Invariance principles for logarithmic averages, *Math. Proc. Cambridge Philos. Soc.* **112** (1992), 195–205. *MR* **93e**:60057

[6] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Academic Press, New York, 1981.*MR* **84d**:60050

[7] Dudley, R. M., *Probabilities and metrics. Convergence of laws on metric spaces, with a view to statistical testing*, Aarhus University Lecture Note Series, No. 45, Matematisk Institut, Aarhus Universitet, Aarhus, 1976. *MR* **58** #7764

[8] Feller, W., *An introduction to probability theory and its applications*, Vol. II, Wiley, New York, 1966. *MR* **35** #1048

[9] Fisher, A., A pathwise central limit theorem for random walks, 1989 (preprint).

[10] Horváth, L. and Khoshnevisan, D., A strong approximation for logarithmic averages, *Studia Sci. Math. Hungar.* **31** (1994), 187–196.

[11] Lacey, M. and Philipp, W., A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9** (1990), 201–205. *MR* **91e**:60100

[12] Major, P., An improvement of Strassen's invariance principle, *Ann. Probab.* **7** (1979), 55–61. *MR* **82f**:60069

[13] Petrov, V. V., *Limit theorems of probability theory,* Oxford Science Publications, 1995.

[14] Sadikova, S. M., On two-dimensional analogs of an inequality of Esseen and their application to the central limit theorem, *Teor. Verojatnost. i Primenen.* **11** (1966), 369–380 (in Russian). *MR* **34** #6832. See also *Theory Probab. Appl.* **11** (1966), 325–335.

[15] Schatte, P., On strong versions of the central limit theorem, *Math. Nachr.* **137** (1988), 249–256. *MR* **89i**:60070

[16] Simons, G. and Stout, W., A weak invariance principle with applications to domains of attraction, *Ann. Probability* **6** (1978), 294–315. *MR* **57** #10778

[17] Strassen, V., Almost sure behavior of sums of independent random variables and martingales, *Proc. 5th Berkeley Sympos. Math. Statist. and Probability*, (Berkeley, Calif., 1965/66), Vol. 2. Contributions to probability theory, Part I, University of California Press, Berkeley, Calif., 1967, 315–343. *MR* **35** #4969

[18] Weigl, A., Zwei Sätze über die Belegungszeit beim random walk, Diplomarbeit, Technical University, Vienna, 1989.

(Received October 30, 1996)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

berkes@math-inst.hu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UT 84112
U.S.A.

horvath@math.utah.edu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 25–41

THE FIRST PASSAGE DENSITY OF THE BROWNIAN MOTION TO A LIPSCHITZ-CONTINUOUS BOUNDARY

Zs. BIRÓ and Gy. MICHALETZKY

Dedicated to Endre Csáki for his sixtieth birthday

Abstract

In this paper we are going to investigate the crossing probabilities of the Wiener process to a Lipschitz continuous boundary. Our method is similar to that one applied in Durbin [2] to differentiable boundaries. Also we are going to derive a differential equation for the crossing probabilities.

1. Introduction

Let $W(u)$ be the standard Wiener process on $[0,+\infty)$ and let a be a continuous function on $[0,+\infty)$. We are interested in the distribution and density of the first hitting time

$$\tau_a = \inf\{t \mid W(t) = a(t)\}.$$

There are many papers dealing with this topic. V. Strassen [11] has shown that, if a is a continuously differentiable function, then τ_a has a continuous density. His proof is based on the following lemma: If a and b are piecewise continuous functions on $[0,+\infty)$ and $t>0$ such that $a \leqq b$ in $(0,t]$ and $a \geqq b$ in $(t,t+\delta)$ where $\delta>0$, a and b are differentiable at t, then

$$\limsup_{h,k\to 0+}\left[\frac{F_a(t+k)-F_a(t-h)}{h+k}-\frac{F_b(t+k)-F_b(t-h)}{h+k}\right]\leqq 0,$$

where

$$F_\psi(t) = P[W(u) \geqq \psi(u) \text{ for some } u \in (0,t]].$$

A. A. Novikov [7] has given an estimation for the tail distribution of the first hitting time applying the Girsanov theorem. If the function a satisfies

1991 *Mathematics Subject Classification.* Primary 60G15, 60G17.
Key words and phrases. Wiener process, crossing probability, first passage density, moving boundary.

This research was supported in part by grants from the Hungarian National Science Foundation, No. T15668 and T16665.

0081–6906/97/\$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

some smoothness conditions, and also conditions describing its increasing rate then

$$\ln P(\tau_a > T) = -\frac{1}{2}\int_s^T (a'(t))^2 dt(1+o(1)) \qquad T \to +\infty.$$

P. Salminen [10] with similar methods has derived expressions for the first passage distribution and density. In J. Durbin [3], B. Ferebee [6], C. Park and S. R. Paranjape [8] one can find integral equations used to compute this density numerically.

We shall use the following notations:

$$I(W,a,v) = I\{\sup_{[0,v]}[W(u)-a(u)] \geqq 0\};$$

$$I^c(W,a,v) = I\{W(u) < a(u) \mid u \in [0,v]\};$$

$$I(W,a,v,t) = I\{\sup_{[v,t]}[W(u)-a(u)] \geqq 0\},$$

$$I^c(W,a,v,t) = I\{W(u) < a(u) \mid u \in [v,t]\},$$

where $I\{\ \}$ denotes the indicator function and $I^c = 1 - I$. Denote $p(t)$ the first passage density of $W(u)$ for boundary a at point t, then

$$p(t) = \lim_{v \uparrow t} \frac{1}{t-v}[P(t) - P(v)], \tag{1}$$

if the limit exists, where $P(z) = E[I(W,a,z)]$. Thus $p(t) = P_l'(t)$, where P_l' denotes the left derivative of the function P. Obviously

$$P(t) - P(v) = E[I(W,a,t) - I(W,a,v)] = E[I^c(W,a,v)I(W,a,t)],$$

thus

$$p(t) = \lim_{v \uparrow t} \frac{1}{t-v} E[I^c(W,a,v)I(W,a,t)],$$

if the limit exists.

Remark 1. If P is an absolutely continuous function, then p exists a.e. (with respect to the Lebesgue-measure) and

$$P(t) = \int_0^t p(u)du,$$

so p can be considered as a "density" function, although in general an unnormalized density, because, $\int_0^{+\infty} p(u)du$ is not necessarily 1. Kolmogorov–Petrovsky–Erdős (cf. [5]) give a condition for the function a assuring that $\int_0^{+\infty} p(u)du = 1$.

An explicit formula for this density was given by J. Durbin [3], namely

$$p(t) = b(t) f_t[a(t)], \tag{2}$$

where

$$b(t) = \lim_{s \uparrow t} \frac{1}{t-s} E[I^c(W, a, s)(a(s) - W(s)) \mid W(t) = a(t)]$$

and

$$f_t(x) = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{x^2}{2t}\right]$$

is the density function of the distribution $N(0, t)$.

This is a generalization of the Bachelier–Levy formula

$$p_L(t) = \frac{1}{t} L(0) f_t[L(t)], \tag{3}$$

where $L(t)$ $t \in [0, +\infty)$ is a line, with $L(0) > 0$.

Durbin uses a very elegant but heuristic argument working with sets of zero probability, and in this way eliminating all the problems concerning differentiability and changing the orders of different operators such as limit and integration. In this paper we are going to prove (2) for a larger class of function.

In the first passage problems the boundary a is usually continuously differentiable, but we shall study a more general class of function.

DEFINITION 1. A function f defined on $[0, +\infty)$ is called Lipschitz-continuous on finite intervals if for all $0 < m < +\infty$ the restriction of the function f to the interval $[0, m]$ is Lipschitz-continuous.

We shall use some further notations. Let Y, Z be real random variables and denote f_Y, $f_{Y,Z}$, $f_{Y|Z}$ their probability density, joint density and conditional density. For the sake of simplicity instead of $f_{W(t)}$ we are going to use the notation f_t. The notations $f_{t,Z}$, $f_{t|Z}$ have similar meanings. The function Φ denotes the normal distribution function.

2. The first passage density

LEMMA 1. *If a is a Lipschitz-continuous function on finite intervals then the function P — defined in (1) — is also Lipschitz-continuous function on finite intervals.*

PROOF. As P is an increasing function, it is enough to show that for all m there exists a constant K_m such that

$$P(t) - P(v) \leqq K_m(t - v) \qquad \text{when} \quad 0 \leqq v \leqq t \leqq m.$$

Using the Lipschitz continuity of a there exists a constant C_m such that

$$|a(x)-a(y)| \leqq C_m|x-y| \qquad \text{when} \qquad 0 \leqq x, y \leqq m.$$

Let $0 \leqq v < t \leqq m$ be fixed points and

$$L_v(x) = a(v) - C_m(x-v) \qquad x \in [0,t].$$

Obviously,

$$\begin{aligned} a(x) &\leqq L_v(x) \qquad \text{if} \qquad x \in [0,v]; \\ a(x) &\geqq L_v(x) \qquad \text{if} \qquad x \in [v,t]. \end{aligned}$$

Consequently,

$$\begin{aligned} I^c(W,a,v) &\leqq I^c(W,L_v,v); \\ I(W,a,v,t) &\leqq I(W,L_v,v,t). \end{aligned}$$

Thus

$$0 \leqq P(t) - P(v) = E[I^c(W,a,v)I(W,a,t)] = E[I^c(W,a,v)I(W,a,v,t)] \leqq$$
$$\leqq E[I^c(W,L_v,v)I(W,L_v,v,t)] = \int_v^t p_v(u)du,$$

where $p_v(u)$ is the first passage density of W for the boundary L_v.

Thus

$$P(t) - P(v) \leqq \int_v^t \frac{1}{u} L_v(0) f_u[L_v(u)]du.$$

The function in the argument of this integral is continuous, therefore

$$\limsup_{v \nearrow t} \frac{1}{t-v}[P(t)-P(v)] \leqq \frac{1}{t} L_t(0) f_t[L_t(t)] = \frac{1}{t}[a(t) + C_m t] f_t[a(t)].$$

But if $0 < t \leqq m$ then

$$\frac{1}{t}[a(t) + C_m t] f_t[a(t)] \leqq 2C_m f_t[a(t)],$$

and as $a(0) > 0$

$$\lim_{t \searrow 0} \frac{1}{t} f_t[a(t)] = 0.$$

The function $f_t[a(t)]$ is continuous so taking

$$K_m = 2C_m \sup_{0<t\leqq m} f_t[a(t)]$$

we get that

$$P(t) - P(v) \leqq K_m(t-v). \qquad \square$$

Since Lipschitz continuity implies absolute continuity we get the following

COROLLARY 1. *If the function a is Lipschitz-continuous on finite intervals then P is an absolutely continuous function.*

COROLLARY 2. *If the derivative of the function P exists at t, then $p(t) = P'(t)$ and*

$$p(t) \leqq \frac{1}{t}[a(t) + C_t t] f_t[a(t)].$$

COROLLARY 3. *Let $0 < \varepsilon \leqq m$ be fixed positive numbers. If $0 < \varepsilon \leqq x < y \leqq m$, then there exists a constant K that*

$$\frac{1}{y-x}[P(y) - P(x)] \leqq K.$$

Let $0 < s \leqq v \leqq t$ and consider the functions

$$P_s(t, v \mid x) = E[I^c(W, a, v) I(W, a, t) \mid I^c(W, a, s) = 1, X_s = x]$$
$$g_s(x) = E[I^c(W, a, s) \mid X_s = x] f_{X_s}(x),$$

where $X_s = a(s) - W(s)$ and $f_{X_s}(x)$ is the probability density of X_s at x. (These functions are defined for almost all x values.)

PROPOSITION 1. *Let a be an arbitrary continuous function on $[0, +\infty)$, then*

$$P(t) - P(v) = \int_0^{+\infty} P_s(t, v \mid x) g_s(x) dx.$$

PROOF.

$$\begin{aligned} P(t) - P(v) &= E\left[I^c(W, a, s) I(W, a, t)\right] = \\ &= \int_{\mathbb{R}} E\left[I^c(W, a, s) I(W, a, t) \mid I^c(W, a, s) = 1, X_s = x\right] \times \\ &\quad \times E\left[I^c(W, a, s) \mid X_s = x\right] f_{X_s}(x) dx + \\ &+ \int_{\mathbb{R}} E\left[I^c(W, a, s) I(W, a, t) \mid I^c(W, a, s) = 0, X_s = x\right] \times \\ &\quad \times P\left[I^c(W, a, s) = 0 \mid X_s = x\right] f_{X_s}(x) dx. \end{aligned}$$

Observe that

$$E[I^c(W, a, s) I(W, a, t) \mid I^c(W, a, s) = 0, X_s = x] = 0 \quad \text{for a.e. } x \text{, and}$$
$$E[I^c(W, a, s) \mid X_s = x] = 0 \quad \text{for a.e. } x \text{ in the interval } (-\infty, 0).$$

So

$$P(t) - P(v) = \int_0^{+\infty} P_s(t, v \mid x) g_s(x) dx$$

which is the assertion of Proposition 1. □

DEFINITION 2. Let

$$p_s(t \mid x) = \limsup_{v \nearrow t} \frac{1}{t-v} P_s(t, v \mid x).$$

Using this definition we prove

PROPOSITION 2. *If a is Lipschitz-continuous on finite intervals and the derivative of the function P exists at t, then there*

$$p(t) = \int_0^{+\infty} p_s(t \mid x) g_s(x) dx.$$

PROOF. If $P'(t)$ exists at t, then at this point obviously

$$P'(t) = \lim_{v \uparrow t} \frac{1}{t-v}[P(t+(t-v)) - P(t)] =$$
$$= \liminf_{v \nearrow t} \int_0^{+\infty} \frac{1}{t-v} P_s\left[t+(t-v), t \mid x\right] g_s(x) dx.$$

Thus applying the Fatou lemma we obtain that

$$P'(t) \geqq \int_0^{+\infty} \liminf_{v \nearrow t} \frac{1}{t-v} P_s[t+(t-v), t \mid x] g_s(x) dx \geqq 0. \tag{4}$$

Let

$$P_s(t \mid x) = E[I(W, a, t) \mid I^c(W, a, s) = 1, X_s = x] \qquad (s < t).$$

This function $P_s(t \mid x)$ is increasing in t for every x, thus its generalized derivatives satisfy the following inequality:

$$\limsup_{v \nearrow t} \frac{1}{t-v}[P_s(t \mid x) - P_s(v \mid x)] \leqq \liminf_{v' \searrow t} \frac{1}{v'-t}[P_s(v' \mid x) - P_s(t \mid x)].$$

Therefore (we remark that $P_s(t, v \mid x) = P_s(t \mid x) - P_s(v \mid x)$, $v < t$)

$$\begin{aligned} &\int_0^{+\infty} \limsup_{v \nearrow t} \frac{1}{t-v} P_s(t, v \mid x) g_s(x) dx \leqq \\ &\leqq \int_0^{+\infty} \liminf_{v \nearrow t} \frac{1}{t-v} P_s[t+(t-v), t \mid x] g_s(x) dx. \end{aligned} \tag{5}$$

From the Fatou lemma it would follow that

$$(6)\qquad \begin{aligned}&\limsup_{v\nearrow t}\int_0^{+\infty}\frac{1}{t-v}P_s(t,v\,|\,x)g_s(x)dx\leqq\\&\leqq\int_0^{+\infty}\limsup_{v\nearrow t}\frac{1}{t-v}P_s(t,v\,|\,x)g_s(x)dx.\end{aligned}$$

Obviously,

$$(7)\qquad \limsup_{v\nearrow t}\int_0^{+\infty}\frac{1}{t-v}P_s(t,v\,|\,x)g_s(x)dx=\lim_{v\uparrow t}\frac{1}{t-v}[P(t)-P(v)]=P'(t).$$

According to the inequalities (4), (5), (6) and the equality (7) we obtain that

$$\begin{aligned}P'(t)&\leqq\int_0^{+\infty}\limsup_{v\nearrow t}\frac{1}{t-v}P_s(t,v\,|\,x)g_s(x)dx\leqq\\&\leqq\int_0^{+\infty}\liminf_{v\nearrow t}\frac{1}{t-v}P_s[t+(t-v),t\,|\,x]g_s(x)dx\leqq P'(t),\end{aligned}$$

so

$$p(t)=P'(t)=\int_0^{+\infty}p_s(t\,|\,x)g_s(x)dx,$$

hence

$$p_s(t\,|\,x)=\limsup_{v\nearrow t}\frac{1}{t-v}P_s(t,v\,|\,x).$$

We have to show that the conditions of the Fatou lemma in the inequality (6) are fulfilled. We are going to prove that there exists a constant M depending on s such that

$$\frac{1}{t-v}P_s(t,v\,|\,x)g_s(x)\leqq Mf_{X_s}(x),\qquad \text{if}\qquad \frac{t+s}{2}<v<t.$$

$$P_s(t,v\,|\,x)g_s(x)=$$

$$
\begin{aligned}
&= E[I^c(W,a,v)I(W,a,t) \mid I^c(W,a,s), X_s = x] \times \\
&\quad \times E[I^c(W,a,s) \mid X_s = x] f_{X_s}(x) = \\
&= \frac{E[I^c(W,a,s)I^c(W,a,v)I(W,a,t) \mid X_s = x]}{E[I^c(W,a,s) \mid X_s = x]} \times \\
&\quad \times E[I^c(W,a,s) \mid X_s = x] f_{X_s}(x) = \\
&= E[I^c(W,a,s)I^c(W,a,s,v)I(W,a,v,t) \mid X_s = x] f_{X_s}(x).
\end{aligned}
$$

(If $x > 0$ then $E[I^c(W,a,s) \mid X_s = x] > 0$ for a.e. x.)

Thus

$$P_s(t,v \mid x) g_s(x) \leqq E[I^c(W,a,s,v)I(W,a,v,t) \mid X_s = x] f_{X_s}(x).$$

Since the Wiener process is a homogeneous Markov process, after a simple computation we obtain that considering the point $(s, a(s) - x)$ as the new origin the conditional expectation can be expressed as follows:

$$E[I^c(W,a,s,v)I(W,a,v,t) \mid X_s = x] = E[I^c(W,a^*,v-s)I(W,a^*,t-s)],$$

where

$$a^*(u) = a(s+u) - (a(s) - x), \quad u \in [0, +\infty], \qquad (x \text{ fixed}).$$

$$E[I^c(W,a^*,v-s)I(W,a^*,t-s)] = P^*(t-s) - P^*(v-s),$$

and

$$P^*(u) = E\left[I(W,a^*,u)\right]$$

the crossing probability defined for the function a^*.

If $x > 0$, then the function a^* satisfies the conditions of Corollary 3. It follows that — if v is sufficiently near to t, for example $\frac{t+s}{2} \leqq v \leqq t$ —, then there exists a constant K, which is independent of x such that

$$E[I^c(W,a,s,v)I(W,a,v,t) \mid X_s = x] = P^*(t-s) - P^*(v-s) \leqq K(t-v), \quad \frac{t+s}{2} \leqq v \leqq t.$$

But then there obviously exists a constant M, such that

$$E[I^c(W,a,s,v)I(W,a,v,t) \mid X_s = x] \leqq M(t-v) \qquad s \leqq v \leqq t.$$

Thus

$$\frac{1}{t-v} P_s(t,v \mid x) g_s(x) \leqq M f_{X_s}(x).$$

The inequality

$$\int_0^{+\infty} M f_{X_s}(x)dx \leq M < +\infty$$

gives that the Fatou lemma could be applied in (6). □

Now we would like to approximate the function a by a suitable chord and show that the first passage probability density of this new function approximates $p(t)$. So let us define a function $\bar{a}$:

$$\bar{a}(u) = \begin{cases} a(u), & \text{if } u \in [0, s] \\ a(s) + \dfrac{1}{t-s}[a(t) - a(s)](u - s), & \text{if } u \in (s, t]. \end{cases}$$

We shall show that

$$\lim_{s \uparrow t} \bar{p}_s(t) = p(t),$$

if $P'(t)$ exists at t, where $\bar{p}_s$ denotes the first passage probability density of W for boundary $\bar{a}$.

Let

$$\bar{P}_s(t, v \mid x) = E[I^c(W, \bar{a}, v)I(W, \bar{a}, t) \mid I^c(W, \bar{a}, s) = 1, X_s = x].$$

Obviously, $\bar{a}(s) = a(s)$, $\bar{a}(t) = a(t)$, and $I^c(W, \bar{a}, s) = I^c(W, a, s)$, respectively. Let us denote the chord by L:

$$L(u) = a(s) + \frac{1}{t-s}[a(t) - a(s)](u - s), \ u \in [s, t], \text{ and let}$$

$$\bar{p}_s(t \mid x) = \begin{cases} \limsup\limits_{v \nearrow t} \dfrac{1}{t-v} \bar{P}_s(t, v \mid x), & \text{if } x > 0 \\ 0, & \text{if } x = 0. \end{cases}$$

Proposition 3. *Let a be a Lipschitz-continuous function on finite intervals, with $a(0) > 0$, then*

$$\bar{p}_s(t \mid x) = \lim_{v \uparrow t} \frac{1}{t-v} \bar{p}_s(t, v \mid x) = \frac{1}{t-s} x f_s(t \mid x),$$

where

$$f_s(t \mid x) = f_{t \mid X_s}[a(t) \mid x].$$

PROOF. Let $s<v<t$ and $x>0$, then

$$E[I^c(W,\bar{a},v)I(W,\bar{a},t)\mid I^c(W,a,s)=1,X_s=x]=$$
$$=\frac{E[I^c(W,a,s)I^c(W,\bar{a},s,v)I(W,\bar{a},v,t)\mid X_s=x]}{E[I^c(W,a,s)\mid X_s=x]}=$$
$$=E[I^c(W,a,s)\mid X_s=x]\frac{E[I^c(W,\bar{a},s,v)I(W,\bar{a},v,t)\mid X_s=x]}{E[I^c(W,a,s)\mid X_s=x]}=$$
$$=E[I^c(W,L,s,v)I(W,L,v,t)\mid X_s=x],$$

where we have used that W is a Markov process and $\bar{a}(u)=L(u)$ if $s\leqq u\leqq t$.

As earlier

$$E[I^c(W,L,s,v)I(W,L,v,t)\mid X_s=x]=E[I^c(W,L^*,v-s)I(W,L^*,t-s)],$$

where

$$L^*(u)=L(s+u)-[a(s)-x]\qquad 0\leqq u\leqq t-s.$$

Thus

$$\bar{p}_s(t\mid x)=\limsup_{v\nearrow t}\frac{1}{t-v}E[I^c(W,L^*,v-s)I(W,L^*,t-s)]=$$
$$=\lim_{v\uparrow t}\frac{1}{t-v}E[I^c(W,L^*,v-s)I(W,L^*,t-s)]=$$
$$=\frac{1}{t-s}L^*(0)f_{t-s}[L^*(t-s)]=$$
$$=\frac{1}{t-s}xf_{t-s}[a(t)-(a(s)-x)].$$

But

$$f_{t-s}[a(t)-(a(s)-x)]=f_{t|s}[a(t)\mid a(s)-x]=f_{t|X_s}[a(t)\mid x],$$

therefore

$$\bar{p}_s(t\mid x)=\frac{1}{t-s}xf_{t|X_s}[a(t)\mid x].$$ □

COROLLARY 4. *$\bar{p}_s(t), s<t$, always exists and*

$$\bar{p}_s(t)=\int_0^{+\infty}\frac{1}{t-s}xf_s(t\mid x)g_s(x)dx.$$

REMARK 2. It can be easily proved that

$$f_s(t\mid x)=f_{W(t)|I^c(W,a,s),X_s}(a(t)\mid 1,x),$$

where

$$f_{W(t)|I^c(W,a,s),X_s}(a(t)|1,x)=\frac{d}{dy}P[W(t)<y|I^c(W,a,s)=1,X_s=x]\bigg|_{y=a(t)}.$$

REMARK 3. We would like to emphasize — as we have seen in the proof of Proposition 3 — that

$$\bar{p}_s(t\,|\,x)=\frac{1}{t-s}xf_{t-s}[a(t)-(a(s)-x)].$$

LEMMA 2. *Let a be a Lipschitz-continuous function on finite intervals, with $a(0)>0$ and $t\in[0,+\infty)$ such a point, where $P'(t)=p(t)$ exists, then*

$$\lim_{s\uparrow t}\bar{p}_s(t)=p(t).$$

PROOF. Suppose that $P'(t)=p(t)$ exists at $t\in[0,+\infty)$. In this case

$$\bar{p}_s(t)=\int_0^{+\infty}\bar{p}_s(t\,|\,x)g_s(x)dx,$$

$$p(t)=\int_0^{+\infty}p(t\,|\,x)g_s(x)dx,\quad \text{respectively}\quad (0<s<t).$$

We show that there exists a constant $K<+\infty$ ($K=K(t)<+\infty$ depending on t) such that

$$|\bar{p}_s(t)-p(t)|\leqq K\sqrt{t-s}.$$

Since $g_s(x)>0$ a.e., if $x>0$, the following inequality is obvious:

$$|\bar{p}_s(t)-p(t)|\leqq\int_0^{+\infty}|\bar{p}_s(t\,|\,x)-p(t\,|\,x)|g_s(x)dx. \tag{8}$$

Now we would like to estimate the difference

$$|\bar{p}_s(t\,|\,x)-p(t\,|\,x)|.$$

Let $s\leqq v\leqq t$ and consider the constant $C=C_t$ for which

$$|\,a(x)-a(y)\,|\leqq C\,|\,x-y,|\qquad x,y\in[0,t].$$

Let

$$L^1(u,v)=a(v)+C(v-u),\qquad u\in[s,t];$$
$$L^2(u,v)=a(v)+C(u-v),\qquad u\in[s,t];$$
$$L^s(u)=a(s)+C(s-u),\qquad u\in[0,s].$$

Let

$$P_s^1(t,v\mid x)=E[I^c(W,L^1,s,v)I(W,L^1,v,t)\mid X_s=x];$$
$$P_s^2(t,v\mid x)=E[I^c(W,L^2,s,v)I(W,L^2,v,t)\mid X_s=x].$$

Obviously,

$$P_s^2(t,v\mid x)\leqq P_s(t,v\mid x)\leqq P_s^1(t,v\mid x)\quad \text{for a.e. } x.$$

Thus we obtain that

$$\begin{aligned}|\bar P_s(t,v\mid x)-P_s(t,v\mid x)|&\leqq\\ \leqq\max[\bar P_s(t,v\mid x)-P_s^2(t,v\mid x),&\,P_s^1(t,v\mid x)-\bar P_s(t,v\mid x)].\end{aligned}\tag{9}$$

Let

$$p_s^1(t\mid x)=\lim_{v\uparrow t}\frac{1}{t-v}P_s^1(t,v\mid x),$$
$$p_s^2(t\mid x)=\lim_{v\uparrow t}\frac{1}{t-v}P_s^2(t,v\mid x).$$

After simple computation from the inequality (9) we obtain that

$$|\bar p_s(t\mid x)-p_s(t\mid x)|\leqq\max[|\bar p_s(t\mid x)-p_s^2(t\mid x)|,|p_s^1(t\mid x)-\bar p_s(t\mid x)|].\tag{10}$$

Remark 3 implies that

$$p_s^1(t\mid x)=\frac{1}{t-s}[a(t)+C(t-s)-(a(s)-x)]f_{t-s}[a(t)-(a(s)-x)];$$

$$p_s^2(t\mid x)=\begin{cases}\frac{1}{t-s}[a(t)-C(t-s)-(a(s)-x)]f_{t-s}[a(t)-(a(s)-x)]\\ \qquad\qquad\qquad\text{if}\quad x\geqq a(s)-[a(t)-C(t-s)]\\ 0,\qquad\text{otherwise.}\end{cases}$$

As $\bar p_s(t\mid x)=\frac{1}{t-s}xf_{t-s}[a(t)-(a(s)-x)]$, so

$$\begin{aligned}&|p_s^1(t\mid x)-\bar p_s(t\mid x)|\leqq\\ &\leqq\frac{1}{t-s}|a(t)-a(s)+C(t-s)|\cdot f_{t-s}[a(t)-(a(s)-x)]\leqq\\ &\leqq 2Cf_{t-s}[a(t)-(a(s)-x)],\end{aligned}$$

since $|a(t)-a(s)|\leqq C(t-s)$; and

$$|\bar p_s(t\mid x)-p_s^2(t\mid x)|\leqq$$
$$\leqq\begin{cases}2Cf_{t-s}[a(t)-(a(s)-x)],\ \text{if}\quad x\geqq a(s)-[a(t)-C(t-s)],\\ \frac{1}{t-s}xf_{t-s}[a(t)-(a(s)-x)],\quad\text{otherwise.}\end{cases}$$

But, if $0 \leqq x < a(s) - [a(t) - C(t-s)]$, then $x \leqq 2C(t-s)$ so

$$(11) \qquad |\bar{p}_s(t \mid x) - p_s(t \mid x)| \leqq 2Cf_{t-s}[a(t) - (a(s) - x)].$$

From the inequalities (9), (10) and (11) it follows:

$$|\bar{p}_s(t \mid x) - p_s(t \mid x)| \leqq 2Cf_{t-s}[a(t) - (a(s) - x)] =$$
$$= 2C\frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{[a(t) - (a(s) - x)]^2}{2(t-s)}\right).$$

So we obtain that

$$(12) \qquad |\bar{p}_s(t \mid x) - p_s(t \mid x)| \leqq D\sqrt{t-s}\frac{1}{t-s} \exp\left(-\frac{[a(t) - (a(s) - x)]^2}{2(t-s)}\right),$$
$$\text{where} \qquad D = 2C\frac{1}{\sqrt{2\pi}}.$$

From the inequalities (9) and (12) we obtain that

$$(13) \qquad |\bar{p}_s(t) - p_s(t)| \leqq \leqq D\sqrt{t-s}\int_0^{+\infty} \frac{1}{t-s} \exp\left(-\frac{[a(t) - (a(s) - x)]^2}{2(t-s)}\right) g_s(x)dx.$$

We are going to show that the integral in (13) is bounded in a small left neighbourhood of t. Consider an $\varepsilon > 0$ such that $\varepsilon < \frac{t}{2}$. We prove that

$$(14) \qquad \sup_{t-\varepsilon \leqq s \leqq t} \int_0^{+\infty} \frac{1}{t-s} \exp\left(-\frac{[a(t) - (a(s) - x)]^2}{2(t-s)}\right) g_s(x)dx < +\infty.$$

As

$$g_s(x) = E[I^c(W, a, s) \mid X_s = x]f_{X_s}(x)$$

and

$$a(u) \leqq L^s(u), \qquad u \in [0, s],$$

we obtain that

$$g_s(x) \leqq E[I^c(W, L^s, s) \mid X_s = x]f_{X_s}(x).$$

Since the line L^s can be transformed into a horizontal line using a linear transformation, an easy application of the reflection principle gives that (cf. Durbin [2])

$$E[I^c(W, L^s, s) \mid X_s = x] = 1 - \exp\left[-\frac{(a(s) + Cs)x}{2s}\right], \text{if } x > 0.$$

Since $0 \leqq f_{X_s}(x) \leqq \frac{1}{\sqrt{2\pi s}}$, we obtain that

$$g_s(x) \leqq \frac{1}{\sqrt{2\pi s}} \left[1 - \exp\left(-\frac{(a(s)+Cs)x}{2s} \right) \right].$$

If $\frac{t}{2} \leqq t - \varepsilon \leqq s < t$, then

$$\frac{a(s)+Cs}{2s} \leqq \frac{a(0)+a(s)-a(0)+Cs}{2s} \leqq \frac{a(0)}{t} + C,$$

so using that $1-\exp(-x) \leqq x$ we obtain that

$$g_s(x) \leqq \frac{1}{\sqrt{\pi t}} \left(\frac{a(0)}{t} + 2C \right) x.$$

Thus

$$\int_0^{+\infty} \frac{1}{t-s} \exp\left(-\frac{[a(t)-(a(s)-x)]^2}{2(t-s)} \right) g_s(x)dx \leqq$$

$$\leqq M \int_0^{+\infty} \frac{x}{t-s} \exp\left(-\frac{[a(t)-a(s)+x]^2}{2(t-s)} \right) dx,$$

where

$$M = \frac{1}{\sqrt{\pi t}} \left(\frac{a(0)}{t} + 2C \right).$$

Applying the identity

$$x = [a(t)-a(s)+x] - [a(t)-a(s)]$$

the integral above can be written as a sum of two integrals. The first one can be integrated giving that it is bounded by one, the second one can be transformed into the integral of a normal density function. Finally, using the inequality

$$\frac{|a(t)-a(s)|}{t-s} \leqq C$$

we obtain that

$$\int_0^{+\infty} \frac{1}{t-s} \exp\left(-\frac{[a(t)-a(s)+x]^2}{2(t-s)} \right) g_s(x)dx \leqq$$

$$\leqq M(1+\sqrt{2\pi(t-s)}C) \leqq M(1+C\sqrt{\pi t}) = N < +\infty.$$

So we have shown that the integral (14) is finite. From this and (13) it follows that

$$|\bar{p}_s(t) - p_s(t)| \leqq DN\sqrt{t-s},$$

if $t - \varepsilon \leqq s < t$, i.e. we have also proved that

$$\lim_{s \uparrow t} \bar{p}_s(t) = p(t). \qquad \square$$

COROLLARY 5. *If a is a Lipschitz-continuous function on finite intervals, with $a(0) > 0$ and $p(t) = P'(t)$ exists at point t, then*

$$p(t) = \lim_{s \uparrow t} \frac{1}{t-s} \int_0^{+\infty} x f_s(t \mid x) g_s(x) dx$$

(see Durbin [3]).

PROPOSITION 4. *Let*

$$g_s(x \mid t) = \frac{d}{dx} E[I(X_s < x) I^c(W, a, s) \mid W(t) = a(t)].$$

Then

$$f_s(t \mid x) g_s(x) = f_t[a(t)] g_s(x \mid t) \qquad a.e.\ x > 0$$

(cf. Durbin [3]).

THEOREM 1. *Let a be a Lipschitz-continuous function on finite intervals, with $a(0) > 0$. Then the first passage density of the Wiener process for the boundary a exists and*

$$p(t) = \lim_{s \uparrow t} \frac{1}{t-s} E[I^c(W, a, s)(a(s) - W(s)) \mid W(t) = a(t)] f_t[a(t)]$$

a.e. $t \in [0, +\infty)$, where

$$f_t[a(t)] = \frac{1}{\sqrt{2\pi t}} \exp\left[-\frac{a(t)^2}{2t}\right].$$

PROOF. From Lemma 1 it follows that the function $p(t), t \in [0, +\infty)$, exists, and $p(t) = P'(t)$ a.e. $t \in [0, +\infty)$. Lemma 2 and Corollary 5 imply that

$$p(t) = \lim_{s \uparrow t} \frac{1}{t-s} \int_0^{+\infty} x f_s(t \mid x) g_s(x) dx \qquad \text{a.e. } t \in [0, +\infty).$$

Using Proposition 4,

$$p(t) = \lim_{s \uparrow t} \frac{1}{t-s} f_t[a(t)] \int_0^{+\infty} x g_s(x \mid t) dx.$$

As

$$g_s(x \mid t) = \frac{d}{dx} E[I(X_s < x) I^c(W, a, s) \mid W(t) = a(t)]$$

and

$$X_s = a(s) - W(s),$$

it is obvious that

$$\int_0^{+\infty} x g_s(x \mid t) dx = E[I^c(W, a, s)(a(s) - W(s)) \mid W(t) = a(t)].$$

Thus

$$p(t) = \lim_{s \uparrow t} \frac{1}{t-s} E[I^c(W, a, s)(a(s) - W(s)) \mid W(t) = a(t)] f_t[a(t)]. \qquad \square$$

REFERENCES

[1] Billingsley, P., *Convergence of probability measures*, Wiley, New York, 1968. *MR* **38** #1718

[2] Durbin, J., Boundary-crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov–Smirnov test, *J. Appl. Probab.* **8** (1971), 431–453. *MR* **45** #1248

[3] Durbin, J., The first passage density of a continuous Gaussian process to a general boundary, *J. Appl. Probab.* **22** (1985), 99–122. *MR* **86h**:60083. Correction: *MR* **89i**:60080

[4] Durbin, J., The first passage density of the Brownian motion process to a curved boundary, Technical Report No. 92, Statistics and Applied Probability Program, University of California, Santa Barbara, 1989. *J. Appl. Probab.* **29** (1992), 291–304. *MR* **93d**:60130

[5] Ferebee, B., The tangent approximation to one-sided Brownian exit densities, *Z. Wahrsch. Verw. Gebiete* **61** (1982), 309–326. *MR* **84c**:60120

[6] Lerche, H. R., *Boundary crossing of Brownian motion*, Lecture Notes in Statistics, **40**, Springer-Verlag, New York, 1986. *MR* **88c**:60150

[7] Novikov, A. A., Estimates for and asymptotic behavior of the probabilities of a Wiener process not crossing a moving boundary, *Mat. Sb. (N.S.)* **110** (1979), 539–550 (in Russian). *MR* **82e**:60132

[8] Park, C. and Paranjape, S. R., Probabilities of Wiener paths crossing differentiable curves, *Pacific J. Math.* **53** (1974), 579–583. *MR* **50** #11487

[9] Riesz, F. and Szőkefalvi-Nagy, B., *Functional analysis*, Frederick Ungar, New York, 1955. *MR* **17**, 175

[10] Salminen, P., On the first hitting time and the last exit time for a Brownian motion to/from a moving boundary, *Adv. in Appl. Probab.* **20** (1988), 411–426. *MR* **89i**:60161

[11] Strassen, V., Almost sure behavior of sums of independent random variables and martingales, *Proc. 5th Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II. Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, 315–343. *MR* **35** #4969

(Received June 28, 1995)

MTA SZÁMÍTÁSTECHNIKAI ÉS
AUTOMATIZÁLÁSI KUTATÓ INTÉZETE
LÁGYMÁNYOSI U. 11
H–1111 BUDAPEST
HUNGARY

biro@luna.aszi.sztaki.hu

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS
STATISZTIKA TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

michgy@ludens.elte.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 43–74

ALMOST SURE SUMMABILITY OF PARTIAL SUMS

M. CSÖRGŐ, L. HORVÁTH and Q.-M. SHAO

Dedicated to Endre Csáki on the occasion of his sixtieth birthday

Abstract

Let $\{X_n, n \geqq 1\}$ be a sequence of independent r.v.'s and let $\{S_n, n \geqq 1\}$ be their partial sums. We study the problem of having $\sum_{1 \leqq n < \infty} |S_n|^p/q(n) < \infty$ a.s. for $0 < p < \infty$ and functions $q(n) > 0$ under tight q-weighted summability conditions of moments of X_n and of quantiles of the distribution function of S_n. We show that, in the i.i.d. case, our conditions are optimal for the indicated summability problem of partial sums, and discuss also some applications of our results.

1. Introduction and results

Studying the asymptotics of L_p-functionals of the uniform empirical process, Csörgő, Horváth and Shao [5] reproved the following dichotomy result.

THEOREM A. *Let $\{W(t), 0 \leqq t < \infty\}$ be a Wiener process, $0 < p < \infty$ and q be a positive function on $[1, \infty)$. Then*

$$\int_1^\infty |W(t)|^p/q(t)dt < \infty \quad a.s. \tag{1.1}$$

holds if and only if we have

$$\int_1^\infty t^{p/2}/q(t)dt < \infty. \tag{1.2}$$

Shepp [10] obtained Theorem A when $p = 2$. His proof is based on Radon–Nikodym derivatives of Gaussian measures. Rajput [9] proved Theorem A for $1 \leqq p < \infty$. The proof of Csörgő et al. [5] for the general case of $0 < p < \infty$, as stated in Theorem A, is direct and elementary. The main aim of this paper is to get necessary and sufficient conditions for (1.1), when $W(t)$ in it is replaced by partial sums of independent r.v.'s.

1991 *Mathematics Subject Classification.* Primary 60F15, 60F17; Secondary 60G50.
Key words and phrases. Partial sums, summability, inequalities, L_p-norms, applications.

Throughout this paper we assume that

$$\{X_n, n \geqq 1\} \quad \text{is a sequence of i.i.d.r.v.'s.} \tag{1.3}$$

To make the presentation and the proofs simpler, we impose some regularity conditions on q as follows:

$$q(n) > 0 \quad \text{for all } n \geqq 1, \tag{1.4}$$

(1.5) there is a constant C_1 such that $\max_{1 \leqq i \leqq n} q(i) \leqq C_1 q(n)$ for all $n \geqq 1$,

(1.6) there is a constant C_2 such that $q(2n) \leqq C_2 q(n)$ for all $n \geqq 1$.

We note that if $q(n)$ is a positive regularly varying sequence with positive exponent, then (1.4)–(1.6) hold true.

We wish to prove our theorems under weak moment conditions via using quantiles of the distribution function of $S_n = \sum_{1 \leqq i \leqq n} X_i$. Let

$$\mu_n(\alpha) = \inf\{x : \mathbf{P}\{|S_n| \geqq x\} \leqq \alpha\}, \qquad 0 < \alpha < 1, \tag{1.7}$$

and

$$m_n(\alpha) = \max_{1 \leqq i \leqq n} \mu_i(\alpha), \qquad 0 < \alpha < 1. \tag{1.8}$$

Our main result is summarized in the following theorem.

THEOREM 1.1. *Let* $0 < p < \infty$. *We assume that* (1.3)–(1.6) *hold true. If*

$$\sum_{1 \leqq n < \infty} m_n^p(\alpha)/q(n) < \infty,$$

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty \tag{1.9}$$

for some $0 < \alpha < 1/(2(1 + 3^p))$, *then we have*

$$\sum_{1 \leqq n < \infty} |S_n|^p / q(n) < \infty \quad a.s. \tag{1.10}$$

and

$$\sum_{1 \leqq n < \infty} \max_{1 \leqq i \leqq n} |S_i|^p / q(n) < \infty \quad a.s.. \tag{1.11}$$

In general, (1.11) or (1.10) do not imply (1.9). This is illustrated by a simple example. Let $0<p<\infty$ and define

$$\mathbf{P}\{X_1=k^{1/p}\}=\frac{C_3}{k(\log(2k))^2}, \qquad k=1,2,\ldots,$$

where $C_3=1 \Big/ \sum_{1\leqq k<\infty} \frac{1}{k(\log(2k))^2}$. Let $X_j=0$ for $j\geqq 2$. It is clear that $\{X_n, n\geqq 1\}$ are independent r.v.'s, satisfying

$$\sum_{1\leqq n<\infty} \frac{1}{n^2} \max_{1\leqq i\leqq n} |S_i|^p = \sum_{1\leqq n<\infty} X_1^p/n^2<\infty \quad \text{a.s..}$$

Thus we have (1.11) and, therefore, (1.10) as well. However, it is easily seen that

$$\sum_{1\leqq n<\infty} \mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{n^2+n\max_{1\leqq i\leqq n}|X_i|^p} = \sum_{1\leqq n<\infty} \mathbf{E}\frac{|X_1|^p}{n^2+n|X_1|^p}$$
$$\geqq \sum_{1\leqq n<\infty} \frac{1}{n}\mathbf{P}\{|X_1|^p>n\}=\infty.$$

Consequently, (1.9) fails to be true. Hence it is interesting and somewhat unexpected that (1.11) and (1.9) turn out to be equivalent if $\{X_n, n\geqq 1\}$ are independent, identically distributed random variables (i.i.d.r.v.'s).

THEOREM 1.2. *Let* $0<p<\infty$. *We assume that* (1.3)–(1.6) *hold true and that* $\{X_n, n\geqq 1\}$ *are i.i.d.r.v.'s. Then* (1.11) *holds if and only if* (1.9) *is satisfied for all* $0<\alpha<1$.

If we assume more regularity conditions on the distribution of X_i, then we can get the equivalence of (1.10) and (1.11) as well.

THEOREM 1.3. *Let* $0<p<\infty$. *We assume that* (1.3)–(1.6) *hold true,* $\{X_n, n\geqq 1\}$ *are i.i.d.r.v.'s and that there are constants* C_4 *and* $\alpha_0\in(0,\frac{1}{2}]$ *such that*

(1.12) $$\max_{1\leqq i\leqq n}|\mathrm{med}(S_i)|\leqq C_4\mu_n(\alpha_0) \quad \textit{for all} \quad n\geqq 1.$$

Then the following statements are equivalent:

(1.13) $$\sum_{1\leqq n<\infty} |S_n|^p/q(n)<\infty \quad \textit{a.s.},$$

(1.14) $$\sum_{1\leqq n<\infty} \max_{1\leqq i\leqq n}|S_i|^p/q(n)<\infty \quad \textit{a.s.},$$

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{|S_n|^p}{q(n) + n|S_n|^p} < \infty, \tag{1.15}$$

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |S_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |S_i|^p} < \infty, \tag{1.16}$$

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty \quad \textit{and} \quad \sum_{1 \leqq n < \infty} \frac{\mu_n^p(\alpha)}{q(n)} < \infty \quad \textit{for all} \ \ 0 < \alpha < 1, \tag{1.17}$$

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty \quad \textit{and} \quad \sum_{1 \leqq n < \infty} \frac{m_n^p(\alpha)}{q(n)} < \infty \quad \textit{for all} \ \ 0 < \alpha < 1. \tag{1.18}$$

We note that (1.12) holds true in case of symmetric random variables. If (1.12) does not hold, then (1.13) still implies (1.17) with $\alpha = 1/2$ as follows.

THEOREM 1.4. *Let* $0 < p < \infty$. *We assume that* (1.3)–(1.6) *hold true and that* $\{X_n, \geqq 1\}$ *are i.i.d.r.v.'s. Then* (1.13) *implies*

$$\sum_{1 \leqq n < \infty} \frac{\mu_n^p(1/2)}{q(n)} < \infty \quad \textit{and} \quad \sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty. \tag{1.19}$$

Section 2 contains the proofs of Theorems 1.1–1.4. We give some applications of the results of this section and discuss the optimality of our conditions in Section 3.

2. Proofs

We start with some preliminary lemmas.

LEMMA 2.1. *Let* $\{\xi_n, n \geqq 1\}$ *be a sequence of r.v.'s. If*

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{|\xi_n|}{1 + |\xi_n|} < \infty, \tag{2.1}$$

then we have

$$\sum_{1 \leqq n < \infty} |\xi_n| < \infty \quad a.s.. \tag{2.2}$$

If $\{\xi_n, n \geqq 1\}$ *are independent r.v.'s, then* (2.2) *implies* (2.1).

PROOF. Let $\xi_n' = \xi_n I\{|\xi_n| \leqq 1\}$. On observing

$$\frac{1}{2}|\xi_n'| \leqq \frac{|\xi_n|}{1+|\xi_n|} \quad \text{and} \quad I\{|\xi_n| > 1\} \leqq 2\frac{|\xi_n|}{1+|\xi_n|},$$

we get

$$\sum_{1 \leqq n < \infty} \mathbf{E}|\xi_n'| < \infty \tag{2.3}$$

and

$$\sum_{1 \leqq n < \infty} \mathbf{P}\{|\xi_n| > 1\} = \sum_{1 \leqq n < \infty} \mathbf{P}\{|\xi_n| \neq |\xi_n'|\} < \infty. \tag{2.4}$$

Now (2.3) implies

$$\sum_{1 \leqq n < \infty} |\xi_n'| < \infty \quad \text{a.s.}, \tag{2.5}$$

and applying (2.4) together with the Borel–Cantelli lemma, we get (2.2).

Conversely, we assume that (2.2) holds and that $\{\xi_n, n \geqq 1\}$ are independent r.v.'s. Using the Kolmogorov three series theorem (cf., e.g., Chow and Teicher [3], p. 114), we get (2.3) and (2.4). It is easy to see

$$\frac{|\xi_n|}{1+|\xi_n|} \leqq |\xi_n'| + I\{|\xi_n| > 1\},$$

and hence we have (2.1).

LEMMA 2.2. *Let* ξ *be an arbitrary r.v. and* a, p *be positive constants. Then we have*

$$\mathbf{E}\frac{|\xi|^p}{a+|\xi|^p} = ap \int_0^\infty \frac{x^{p-1}}{(a+x^2)^2} \mathbf{P}\{|X| \geqq x\} dx.$$

PROOF. Setting $f(x) = \dfrac{x^2}{a+x^p}$, we obtain immediately

$$\mathbf{E}f(|\xi|) = f(0) + \int_0^\infty f'(x)\mathbf{P}\{|X| \geqq x\} dx$$

$$= ap \int_0^\infty \frac{x^{p-1}}{(a+x^p)^2} \mathbf{P}\{|X| \geqq x\} dx.$$

LEMMA 2.3. *We assume that* (1.3) *holds true. Then for each* $0<\alpha<1/2$ *and* $x \geqq 6m_n(\alpha)$ *we have*

$$\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}\leqq\frac{2\alpha}{1-2\alpha}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq\frac{x}{3}\Big\}+\frac{1}{1-2\alpha}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq\frac{x}{3}\Big\}. \tag{2.6}$$

PROOF. It is well-known (cf., e.g., Breiman [2] and Peligrad [8]) that

$$\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}\leqq\frac{\mathbf{P}\Big\{|S_n|>\frac{2}{3}x\Big\}}{1-\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}} \tag{2.7}$$

and

$$\mathbf{P}\Big\{|S_n|\geqq\frac{2}{3}x\Big\}\leqq\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq\frac{x}{3}\Big\}+\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq\frac{x}{3}\Big\} \tag{2.8}$$

for all $x>0$. Putting together (2.7) and (2.8), we obtain

$$\begin{aligned}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}&\leqq\frac{\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}}{1-\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq\frac{x}{3}\Big\}\\&\quad+\frac{\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq\frac{x}{3}\Big\}}{1-\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}}.\end{aligned} \tag{2.9}$$

Using the definition of $m_n(\alpha)$, we get for all $x\geqq 6m_n(\alpha)$ that

$$\begin{aligned}\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_n-S_i|\geqq\frac{x}{3}\Big\}&\leqq 2\max_{1\leqq i\leqq n}\mathbf{P}\Big\{|S_i|\geqq\frac{x}{6}\Big\}\\&\leqq 2\max_{1\leqq i\leqq n}\mathbf{P}\{|S_i|\geqq m_n(\alpha)\}\\&\leqq 2\max_{1\leqq i\leqq n}\mathbf{P}\{|S_i|\geqq \mu_i(\alpha)\}\\&\leqq 2\alpha.\end{aligned} \tag{2.10}$$

Now (2.9) and (2.10) yield Lemma 2.3.

LEMMA 2.4. *Let $0 < p < \infty$ and $\{a(n), n \geqq 1\}$ be a sequence of positive numbers. We assume that (1.3) holds true. Then, for all $0 < \alpha < \dfrac{1}{2(1+3^p)}$, we have*

$$\frac{\alpha}{2^{p+1}}\left\{\frac{m_n^p(\alpha)}{a(n)+m_n^p(\alpha)}+\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|X_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|X_i|^p}\right\}$$

$$\leqq \mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|S_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|S_i|^p}$$

$$\leqq \frac{6^p}{1-2\alpha(1+3^p)}\left\{\frac{m_n^p(\alpha)}{a(n)+m_n^p(\alpha)}+\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|X_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|X_i|^p}\right\}.$$

PROOF. It follows from the definition of $m_n(\alpha)$ that

$$\mathbf{P}\left\{\max_{1\leqq i\leqq n}|S_i|\geqq x\right\}\geqq \alpha \quad \text{for all} \quad 0<x<m_n(\alpha).$$

Hence, by Lemma 2.2, we have

$$\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|S_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|S_i|^p}$$

$$\geqq pa(n)\int_0^{m_n(\alpha)}\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\left\{\max_{1\leqq i\leqq n}|S_i|\geqq x\right\}dx \tag{2.11}$$

$$\geqq \alpha pa(n)\int_0^{m_n(\alpha)}\frac{x^{p-1}}{(a(n)+x^p)^2}dx=\alpha\frac{m_n^p(\alpha)}{a(n)+m_n^p(\alpha)}.$$

An elementary argument gives

$$\mathbf{P}\left\{\max_{1\leqq i\leqq n}|S_i|\geqq x\right\}\geqq \mathbf{P}\left\{\max_{1\leqq i\leqq n}|X_i|\geqq 2x\right\}$$

for all $x > 0$. Hence, using again Lemma 2.2, we obtain

$$\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|S_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|S_i|^p}=pa(n)\int_0^{\infty}\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\left\{\max_{1\leqq i\leqq n}|S_i|\geqq x\right\}dx$$

$$
\begin{aligned}
&\geqq pa(n)\int_0^\infty \frac{x^{p-1}}{(a(n)+x^p)^2}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq 2x\Big\}dx\\
(2.12)\qquad&\\
&\geqq 2^{-p}pa(n)\int_0^\infty \frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq x\Big\}dx\\
&=2^{-p}\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|X_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|X_i|^p}.
\end{aligned}
$$

Observing that $\frac{\alpha}{2^{p+1}}\leqq 1\Big/\Big(\frac{1}{\alpha}+2^p\Big)$, (2.11) and (2.12) imply the left-hand side inequality of Lemma 2.4.

Lemmas 2.2 and 2.3 yield

$$
\begin{aligned}
&\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|S_i|^p}{a(n)+\max\limits_{1\leqq i\leqq n}|S_i|^p}\\
(2.13)\qquad&\\
&=pa(n)\int_0^{6m_n(\alpha)}\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}dx\\
&\quad+pa(n)\int_{6m_n(\alpha)}^\infty\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}dx=A_n^{(1)}+A_n^{(2)},
\end{aligned}
$$

$$
(2.14)\qquad A_n^{(1)}\leqq pa(n)\int_0^{6m_n(\alpha)}\frac{x^{p-1}}{(a(n)+x^p)^2}dx=\frac{6^p m_n^p(\alpha)}{a(n)+m_n^p(\alpha)}
$$

and

$$
\begin{aligned}
&A_n^{(2)}\leqq\frac{2\alpha pa(n)}{1-2\alpha}\int_{6m_n(\alpha)}^\infty\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq\frac{x}{3}\Big\}dx\\
(2.15)\qquad&\\
&\quad+\frac{pa(n)}{1-2\alpha}\int_{6m_n(\alpha)}^\infty\frac{x^{p-1}}{(a(n)+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq\frac{x}{3}\Big\}dx\\
&=A_n^{(3)}+A_n^{(4)}.
\end{aligned}
$$

Using again Lemma 2.2, we get

$$A_n^{(3)} \leqq \frac{2\alpha p a(n) 3^p}{1-2\alpha} \int\limits_{6m_n(\alpha)}^{\infty} \frac{x^{p-1}}{(a(n)+x^p)^2} \mathbf{P}\Big\{ \max_{1\leqq i\leqq n} |S_i| > x \Big\} dx$$

$$(2.16) \qquad \leqq \frac{2\alpha p a(n) 3^p}{1-2\alpha} \int\limits_{0}^{\infty} \frac{x^{p-1}}{(a(n)+x^p)^2} \mathbf{P}\Big\{ \max_{1\leqq i\leqq n} |S_i| > x \Big\} dx$$

$$= \frac{2\alpha 3^p}{1-2\alpha} \mathbf{E} \frac{\max_{1\leqq i\leqq n} |S_i|^p}{a(n) + \max_{1\leqq i\leqq n} |S_i|^p},$$

and a similar argument gives

$$(2.17) \qquad A_n^{(4)} \leqq \frac{3^p}{1-2\alpha} \mathbf{E} \frac{\max_{1\leqq i\leqq n} |X_i|^p}{a(n) + \max_{1\leqq i\leqq n} |X_i|^p}.$$

Putting together (2.13)–(2.17), we obtain immediately

$$\Big(1 - \frac{2\alpha 3^p}{1-2\alpha}\Big) \mathbf{E} \frac{\max_{1\leqq i\leqq n} |S_i|^p}{a(n) + \max_{1\leqq i\leqq n} |S_i|^p} \leqq \frac{6^p m_n^p(\alpha)}{a(n) + m_n^p(\alpha)}$$

$$+ \frac{3^p}{1-2\alpha} \mathbf{E} \frac{\max_{1\leqq i\leqq n} |X_i|^p}{a(n) + \max_{1\leqq i\leqq n} |X_i|^p},$$

and this completes also the proof of Lemma 2.4.

LEMMA 2.5. *Let* $0 < p < \infty$ *and* $\{q(n), n \geqq 1\}$ *be a sequence of positive numbers. If*

$$(2.18) \qquad \sum_{1\leqq n<\infty} |S_n|^p / q(n) < \infty \quad a.s.,$$

then we have

$$(2.19) \qquad \sum_{1\leqq n<\infty} \frac{\mu_n^p(\alpha)}{q(n)} < \infty \quad \textit{for all } \ 0 < \alpha < 1.$$

PROOF. It follows from Lemma 2.1 in Csörgő et al. [5].

PROOF OF THEOREM 1.1. By (1.5) and (1.6) we have

$$\sum_{1\leqq n<\infty} \max_{1\leqq i\leqq n} |S_i|^p / q(n) \leqq \sum_{0\leqq k<\infty} \Big(\sum_{2^k\leqq n<2^{k+1}} \frac{1}{q(n)} \Big) \max_{1\leqq i<2^{k+1}} |S_i|^p$$

$$\leqq C_1 \sum_{0\leqq k<\infty} \frac{2^k}{q(2^k)} \max_{1\leqq i<2^{k+1}} |S_i|^p$$

$$\leqq C_1 C_2^2 \sum_{0\leqq k<\infty} \frac{2^k}{q(2^{k+2})} \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p.$$

Thus, by Lemma 2.1, it is enough to prove

$$\sum_{0\leqq k<\infty} \mathbf{E} \frac{\max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}{2^{-k} q(2^{k+2}) + \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p} < \infty. \tag{2.20}$$

Applying Lemma 2.4 and the conditions (1.5) and (1.6), we get

$$\frac{\max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}{2^{-k} q(2^{k+2}) + \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p} = \frac{1}{2} \sum_{2^{k+1}\leqq n<2^{k+2}} \frac{\max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}{q(2^{k+2}) + 2^k \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}$$

$$\leqq \frac{1}{2} \sum_{2^{k+1}\leqq n<2^{k+2}} \frac{\max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}{q(n)/C_1 + 2^k \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}$$

$$\leqq 2 \sum_{2^{k+1}\leqq n<2^{k+2}} \frac{\max_{1\leqq i\leqq n} |S_i|^p}{q(n)/C_1 + n \max_{1\leqq i\leqq n} |S_i|^p}$$

$$\leqq 2C_1 \sum_{2^{k+1}\leqq n<2^{k+2}} \frac{\max_{1\leqq i\leqq n} |S_i|^p}{q(n) + n \max_{1\leqq i\leqq n} |S_i|^p}.$$

This yields

$$\sum_{0\leqq k<\infty} \mathbf{E} \frac{\max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}{2^{-k} q(2^{k+2}) + \max_{1\leqq i\leqq 2^{k+1}} |S_i|^p}$$

$$\leqq 2C_1 \sum_{1\leqq n<\infty} \mathbf{E} \frac{\max_{1\leqq i\leqq n} |S_i|^p}{q(n) + n \max_{1\leqq i\leqq n} |S_i|^p}$$

$$\leqq \frac{2C_1 6^p}{1-2\alpha(1k+3^p)} \left\{ \sum_{1\leqq n<\infty} \frac{m_n^p(\alpha)}{q(n)+n m_n^p(\alpha)} + \sum_{1\leqq n<\infty} \mathbf{E} \frac{\max_{1\leqq i\leqq n} |X_i|^p}{q(n)+n \max_{1\leqq i\leqq n} |X_i|^p} \right\}.$$

Thus we have (2.20), which completes also the proof of Theorem 1.1.

PROOF OF THEOREM 1.2. By Theorem 1.1 it is enough to prove that (1.11) implies (1.9) for all $0<\alpha<1$. If (1.11) holds, then by (1.6) we have

$$\sum_{0\leq k<\infty} \frac{2^k}{q(2^k)} \max_{1\leq i\leq 2^k} |S_i|^p < \infty \quad \text{a.s.,}$$

which implies

$$\sum_{0\leq k<\infty} \frac{2^k}{q(2^k)} \max_{2^k<i\leq 2^{k+1}} |S_i - S_{2^k}|^p < \infty \quad \text{a.s..} \tag{2.21}$$

Observing that $\left\{ \max_{2^k<i\leq 2^{k+1}} |S_i - S_{2^k}|^p,\ k\geq 1 \right\}$ are independent r.v.'s, Lemma 2.1 and (2.21) yield

$$\sum_{0\leq k<\infty} \mathbf{E} \frac{\max_{2^k<i\leq 2^{k+1}} |S_i - S_{2^k}|^p}{2^{-k}q(2^k) + \max_{2^k<i\leq 2^{k+1}} |S_i - S_{2^k}|^p} < \infty. \tag{2.22}$$

We assumed that $\{X_n, n\geq 1\}$ are i.i.d.r.v.'s. Consequently, (2.22) holds if and only if

$$\sum_{0\leq k<\infty} \mathbf{E} \frac{\max_{1\leq i\leq 2^k} |S_i|^p}{2^{-k}q(2^k) + \max_{1\leq i\leq 2^k} |S_i|^p} < \infty. \tag{2.23}$$

By Lemma 2.4 we have (2.23) if and only if

$$\sum_{0\leq k<\infty} \frac{m_{2^k}^p(\alpha)}{2^{-k}q(2^k) + m_{2^k}^p(\alpha)} < \infty \quad \text{for all } 0<\alpha<\frac{1}{2(1+3^p)} \tag{2.24}$$

and

$$\sum_{0\leq k<\infty} \mathbf{E} \frac{\max_{1\leq i\leq 2^k} |X_i|^p}{2^{-k}q(2^k) + \max_{1\leq i\leq 2^k} |X_i|^p} < \infty. \tag{2.25}$$

Using (2.24), we get

$$m_{2^k}^p(\alpha) = o(2^{-k}q(2^k)) \quad \text{as } k\to\infty,$$

and, therefore, we have

$$\sum_{0\leq k<\infty} 2^k \frac{m_{2^k}^p(\alpha)}{q(2^k)} < \infty. \tag{2.26}$$

By definition, $m_k(\alpha) \leqq m_{k+1}(\alpha)$ for all α, so applying (1.5) and (1.6) we obtain from (2.26) that

$$\sum_{1 \leqq n < \infty} \frac{m_n^p(\alpha)}{q(n)} < \infty \quad \text{for all } 0 < \alpha < \frac{1}{2(1+3^p)}. \tag{2.27}$$

Also, $m_n(\alpha') \leqq m_n(\alpha)$ if $\alpha' \geqq \alpha$, so by (2.27) we have for all $0<\alpha<1$ that

$$\sum_{1 \leqq n < \infty} \frac{m_n^p(\alpha)}{q(n)} < \infty.$$

Using (2.25), similar arguments give

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + N \max_{1 \leqq i \leqq n} |X_i|^p} < \infty,$$

which also completes the proof of Theorem 1.2.

Proof of Theorem 1.3. Let $\{Y_n, n \geqq 1\}$ be i.i.d.r.v.'s satisfying $Y_n \overset{\mathcal{D}}{=} X_1$. We assume also that $\{X_n, n \geqq 1\}$ and $\{Y_n, n \geqq 1\}$ are independent sequences. We define

$$\tilde{X}_n = X_n - Y_n \quad \text{and} \quad \tilde{S}_n = \sum_{1 \leqq i \leqq n} \tilde{X}_i.$$

Then $\{\tilde{X}_n, n \geqq 1\}$ are i.i.d. symmetric r.v.'s. Using the Lévy inequalities (cf., e.g., Chow and Teicher [3], pp. 71 and 325), we get for all $1 \leqq i \leqq n$

$$\begin{aligned}
&\mathbf{P}\{|S_i| \geqq 2\mu_n(\alpha/8) + |\mathrm{med}(S_i)|\} \leqq \mathbf{P}\{|S_i - \mathrm{med}(S_i)| \geqq 2\mu_n(\alpha/8)\} \\
&\leqq 2\mathbf{P}\{|\tilde{S}_i| \geqq 2\mu_n(\alpha/8)\} \leqq 4\mathbf{P}\{|\tilde{S}_n| \geqq 2\mu_n(\alpha/8)\} \\
&\leqq 8\mathbf{P}\{|S_n| \geqq \mu_n(\alpha/8)\} \leqq \alpha.
\end{aligned}$$

Thus we have

$$m_n(\alpha) \leqq 2\mu_n(\alpha/8) + \max_{1 \leqq i \leqq n} |\mathrm{med}(S_i)|. \tag{2.28}$$

Now (1.12) and the monotonicity of $\mu_n(\alpha)$ in α give

$$\max_{1 \leqq i \leqq n} |\mathrm{med}(S_i)| \leqq C_4 \mu_n(\alpha) \quad \text{for all } 0 < \alpha \leqq \alpha_0. \tag{2.29}$$

Combining (2.28) and (2.29), we get

$$\mu_n(\alpha) \leqq m_n(\alpha) \leqq (C_4 + 2)\mu_n(\alpha/8) \quad \text{for all } 0 < \alpha \leqq \alpha_0. \tag{2.30}$$

Thus the equivalence of (1.17) and (1.18) is now proven.

If $0<\alpha<\min(\alpha_0, \frac{1}{2})$, then by Lemma 2.2 we have

$$\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|S_i|^p}{q(n)+n\max\limits_{1\leqq i\leqq n}|S_i|^p}$$

(2.31)

$$\leqq \frac{pq(n)}{n^2}\int\limits_0^{6m_n(\alpha)}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}dx$$

$$+\frac{pq(n)}{n^2}\int\limits_{6m_n(\alpha)}^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|S_i|\geqq x\Big\}dx=A_n^{(5)}+A_n^{(6)}.$$

It is easy to see that we have

$$\begin{aligned}A_n^{(5)}&\leqq\frac{1}{n}\frac{(6m_n(\alpha))^p}{\frac{q(n)}{n}+m_n^p(\alpha)}\leqq\frac{6^p}{n}\frac{(2+C_4)^p\mu_n^p(\alpha/8)}{\frac{q(n)}{n}+m_n^p(\alpha)}\\&\leqq 6^p(2+C_4)^p\frac{\mu_n^p(\alpha/8)}{q(n)+n\mu_n^p(\alpha/8)}.\end{aligned}\tag{2.32}$$

Using (2.7), we get

$$A_n^{(6)}\leqq\frac{pq(n)}{n^2}\int\limits_{6m_n(\alpha)}^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\frac{\mathbf{P}\{|S_n|\geqq\frac{2}{3}x\}}{1-\max\limits_{1\leqq i\leqq n}\mathbf{P}\{|S_n-S_i|\geqq\frac{x}{3}\}}dx$$

$$\leqq\frac{pq(n)}{n^2}\int\limits_{6m_n(\alpha)}^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\frac{\mathbf{P}\{|S_n|\geqq\frac{2}{3}x\}}{1-2\max\limits_{1\leqq i\leqq n}\mathbf{P}\{|S_i|\geqq\frac{x}{6}\}}dx$$

$$\leqq\frac{pq(n)}{n^2(1-2\alpha)}\int\limits_{6m_n(\alpha)}^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\mathbf{P}\Big\{|S_n|\geqq\frac{2}{3}x\Big\}dx\tag{2.33}$$

$$\leqq\frac{(2/3)^p}{1-2\alpha}\frac{pq(n)}{n^2}\int\limits_0^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\mathbf{P}\{|S_n|\geqq x\}dx$$

$$=\frac{(2/3)^p}{1-2\alpha}\mathbf{E}\frac{|S_n|^p}{q(n)+n|S_n|^p}.$$

Applying again Lemma 2.2, we obtain

$$\mathbf{E}\frac{|S_n|^p}{\frac{q(n)}{n}+|S_n|^p}=\frac{pq(n)}{n}\int\limits_0^{\infty}\frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2}\mathbf{P}\{|S_n|\geqq x\}dx$$

$$\geq \frac{pq(n)}{n} \int_0^{\mu_n(\alpha/8)} \frac{x^{p-1}}{(\frac{q(n)}{n}+x^p)^2} \mathbf{P}\{|S_n| \geq x\} dx \tag{2.34}$$

$$\geq \frac{\alpha}{8} \frac{\mu_n^p(\alpha/8)}{\frac{q(n)}{n}+\mu_n^p(\alpha/8)}.$$

Thus, by (2.31)–(2.34), we can find $C_5 = C_5(p)$ such that

$$\mathbf{E} \frac{\max\limits_{1 \leq i \leq n} |S_i|^p}{q(n)+n \max\limits_{1 \leq i \leq n} |S_i|^p} \leq C_5 \mathbf{E} \frac{|S_n|^p}{q(n)+n|S_n|^p},$$

and, therefore, (1.15) implies (1.16). It is trivial that (1.16) implies (1.15).

According to Lemma 2.4, the statements in (1.16) and (1.18) are equivalent. By Theorem 1.2 we have (1.14) if and only if (1.18) holds true. Thus, it suffices to show that (1.13) implies (1.17). According to Lemma 2.5, (1.13) yields

$$\sum_{1 \leq n < \infty} \frac{\mu_n^p(\alpha)}{q(nb)} < \infty \quad \text{for all } 0 < \alpha < 1. \tag{2.35}$$

Recalling the definition of $\widetilde{S}_n$, we have

$$\sum_{1 \leq n < \infty} |\widetilde{S}_n|^p / q(n) < \infty \quad \text{a.s..} \tag{2.36}$$

Hence Lemma 2.5 gives

$$\sum_{1 \leq n < \infty} \frac{\widetilde{\mu}_n^p(\alpha)}{q(n)} < \infty, \tag{2.37}$$

where

$$\widetilde{\mu}_n(\alpha) = \inf\{x : \mathbf{P}\{|\widetilde{S}_n| \geq x\} \leq \alpha\}.$$

By (2.36) we obtain

$$\sum_{1 \leq k < \infty} \sum_{2^k+1 \leq n \leq 2^{k+1}} |\widetilde{S}_{n+2^{k+1}}|^p / q(n+2^{k+1}) < \infty \quad \text{a.s.,} \tag{2.38}$$

and hence, by (1.5) and (1.6), we have

$$\sum_{1 \leq k < \infty} \sum_{2^k+1 \leq n \leq 2^{k+1}} |\widetilde{S}_{n+2^{k+1}}|^p / q(n) < \infty \quad \text{a.s..} \tag{2.39}$$

Putting together (2.36) and (2.39), we get

$$\sum_{1\le k<\infty}\ \sum_{2^k+1\le n\le 2^{k+1}} |\widetilde S_{n+2^{k+1}}-\widetilde S_n|^p/q(n)<\infty \quad \text{a.s.,} \tag{2.40}$$

which implies

$$\sum_{1\le k<\infty}\ \sum_{2^{2k}+1\le n\le 2^{2k+1}} |\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p/q(2^{2k})<\infty \quad \text{a.s..} \tag{2.41}$$

Observing that $\Big\{\sum_{2^{2k}+1\le n<2^{2k+1}} |\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p,\ k\ge 1\Big\}$ are independent r.v.'s, it follows from Lemma 2.1 that

(2.42)

$$\sum_{1\le k<\infty}\mathbf{E}\Big\{\sum_{2^{2k}+1\le n\le 2^{2k+1}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p\Big/\Big(q(2^{2k})+\sum_{2^{2k}+1\le n\le 2^{2k+1}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p\Big)\Big\}<\infty.$$

We have for all $x>0$

$$\begin{aligned}
&\mathbf{P}\Big\{\sum_{2^{2k}+1\le n\le 2^{2k+1}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p\ge x\Big\}\\
&=\mathbf{P}\Big\{\sum_{1\le n\le 2^{2k}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_n|^p\ge x\Big\}\\
&\ge\mathbf{P}\Big\{\sum_{1\le n\le 2^{2k}}(2^{-p}|\widetilde S_{2^{2k+1}}-\widetilde S_{2^{2k}}|^p-2^p|\widetilde S_{n+2^{2k+1}}-\widetilde S_{2^{2k+1}}|^p-2^p|\widetilde S_{2^{2k}}-\widetilde S_n|^p)\ge x\Big\}\\
&\ge\mathbf{P}\Big\{|\widetilde S_{2^{2k+1}}-\widetilde S_{2^{2k}}|^p\ge\frac{x2^p}{2^{2k}}+2^{2p}\max_{1\le n\le 2^{2k}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_{2^{2k+1}}|^p\\
&\qquad+2^{2p}\max_{1\le n\le 2^{2k}}|\widetilde S_{2^{2k}}-\widetilde S_n|^p\Big\}\\
&\ge\mathbf{P}\{\max_{1\le n\le 2^{2k}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_{2^{2k+1}}|\le\widetilde\mu_{2^{2k}}(\alpha)\}\\
&\qquad\times\mathbf{P}\{\max_{1\le n\le 2^{2k}}|\widetilde S_{2^{2k}}-\widetilde S_n|\le\widetilde\mu_{2^{2k}}(\alpha)\}\\
&\qquad\times\mathbf{P}\Big\{|\widetilde S_{2^{2k+1}}-\widetilde S_{2^{2k}}|^p\ge\frac{x2^p}{2^{2k}}+2^{2p+1}\widetilde\mu^p_{2^{2k}}(\alpha)\Big\}.
\end{aligned}$$

Using again the Lévy inequalities (cf., e.g., Chow and Teicher [3], p. 71), for $0<\alpha\le 1/4$ we get

$$\mathbf{P}\{\max_{1\le n\le 2^{2k}}|\widetilde S_{n+2^{2k+1}}-\widetilde S_{2^{2k+1}}|\ge\widetilde\mu_{2^{2k}}(\alpha)\}$$

$$= \mathbf{P}\{\max_{1\leqq n\leqq 2^{2k}} |\widetilde{S}_{2^{2k}} - \widetilde{S}_n| \geqq \widetilde{\mu}_{2^{2k}}(\alpha)\}$$
$$\leqq 2\mathbf{P}\{|\widetilde{S}_{2^{2k}}| \geqq \widetilde{\mu}_{2^{2k}}(\alpha)\} \leqq 2\alpha \leqq 1/2,$$

and by Chow and Teicher ([3], p. 73) we have

$$\mathbf{P}\left\{|\widetilde{S}_{2^{2k+1}} - \widetilde{S}_{2^{2k}}|^p \geqq \frac{2^p x}{2^{2k}} + 2^{2p+1}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}$$
$$= \mathbf{P}\left\{|\widetilde{S}_{2^{2k}}|^p \geqq \frac{2^p x}{2^{2k}} + 2^{2p+1}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}$$
$$\geqq \frac{1}{2}\mathbf{P}\left\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{S}_i|^p \geqq \frac{2^p x}{2^{2k}} + 2^{2p+1}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}$$
$$\geqq \frac{1}{2}\mathbf{P}\left\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{X}_i|^p \geqq \frac{2^{p+1} x}{2^{2k}} + 2^{2p+2}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}.$$

Thus, for each $\alpha_0 \in (0, 1/4)$, there is a positive constant $C_6 = C_6(\alpha_0)$ such that

(2.43)
$$\mathbf{P}\left\{\sum_{2^{2k}+1\leqq n\leqq 2^{2k+1}} |\widetilde{S}_{n+2^{2k+1}} - \widetilde{S}_n|^p \geqq x\right\}$$
$$\geqq C_6\mathbf{P}\left\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{X}_i|^p \geqq \frac{2^{p+1} x}{2^{2k}} + 2^{2p+2}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}$$

for all $x > 0$ and $0 < \alpha \leqq \alpha_0$.

It follows from Lemma 2.2, (2.42) and (2.43) that

$$\frac{1}{C_6}\mathbf{E}\left\{\sum_{2^{2k}+1\leqq n\leqq 2^{2k+1}} |\widetilde{S}_{n+2^{2k+1}} - \widetilde{S}_n|^p \Big/ \left(q(2^{2k}) + \sum_{2^{2k}+1\leqq n\leqq 2^{2k+1}} |\widetilde{S}_{n+2^{2k+1}} - \widetilde{S}_n|^p\right)\right\}$$
$$\geqq q(2^{2k}) \int_0^\infty \frac{1}{(q(2^{2k}) + x)^2}\mathbf{P}\left\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{X}_i|^p \geqq \frac{2^{p+1} x}{2^{2k}} + 2^{2p+2}\widetilde{\mu}^p_{2^{2k}}(\alpha)\right\}dx$$

(2.44)
$$\geqq 2^{-(p+1)} q(2^k) 2^{2k} \int_{2^{2p+2}\widetilde{\mu}^p_{2^{2k}}(\alpha)}^\infty \frac{1}{(q(2^{2k}) + 2^{2k}x)^2}\mathbf{P}\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{X}_i|^p \geqq x\}dx$$
$$= 2^{-(p+1)} q(2^k) 2^{2k} \int_0^\infty \frac{\mathbf{P}\{\max_{1\leqq i\leqq 2^{2k}} |\widetilde{X}_i|^p \geqq x\}}{(q(2^{2k}) + 2^{2k}x)^2}dx$$

$$-2^{-(p+1)}q(2^k)2^{2k}\int_0^{2^{p+2}\widetilde{\mu}_{2^{2k}}(\alpha)}\frac{\mathbf{P}\{\max_{1\leqq i\leqq 2^{2k}}|\widetilde{X}_i|^p\geqq x\}}{(q(2^{2k})+2^{2k}x)^2}dx$$

$$\geqq 2^{-(p+1)}2^{-2k}\mathbf{E}\frac{\max_{1\leqq i\leqq 2^{2k}}|\widetilde{X}_i|^p}{2^{-2k}q(2^{2k})+\max_{1\leqq i\leqq 2^{2k}}|\widetilde{X}_i|^p}-\frac{2^p 2^{2k}\widetilde{\mu}^p_{2^{2k}}(\alpha)}{q(2^{2k})}.$$

Using (2.37) we get

$$\sum_{1\leqq k<\infty}\sum_{2^{2k}+1\leqq n\leqq 2^{2k+1}}\frac{\widetilde{\mu}_{2^{2k}}(\alpha)}{q(2^{2k})}<\infty. \tag{2.45}$$

Consequently, (2.42) and (2.44) yield

$$\sum_{1\leqq k<\infty}\sum_{2^{2k}+1\leqq n\leqq 2^{2k+2}}\mathbf{E}\frac{\max_{1\leqq i\leqq 2^{2k}}|\widetilde{X}_i|^p}{2^{-2k}q(2^{2k})+\max_{1\leqq i\leqq 2^{2k}}|\widetilde{X}_i|^p}<\infty. \tag{2.46}$$

We apply (1.5) and (1.6) to deduce from (2.46) that

$$\sum_{1\leqq n<\infty}\mathbf{E}\frac{\max_{1\leqq i\leqq n}|\widetilde{X}_i|^p}{q(n)+n\max_{1\leqq i\leqq n}|\widetilde{X}_i|^p}<\infty. \tag{2.47}$$

We note that Lemma 1 in Chow and Teicher ([3], p. 325) gives

$$\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|\widetilde{X}_i|\geqq x\Big\}\geqq\frac{1}{2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i-\operatorname{med}(X_1)|\geqq x\Big\}$$
$$\geqq\frac{1}{2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq x+|\operatorname{med}(X_1)|\Big\}.$$

Hence, by Lemma 2.2, we have

$$\mathbf{E}\frac{\max_{1\leqq i\leqq n}|\widetilde{X}_i|^p}{q(n)+n\max_{1\leqq i\leqq n}|\widetilde{X}_i|^p}=\frac{q(n)}{n^2}\int_0^\infty\frac{\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|\widetilde{X}_i|^p\geqq x\Big\}}{(\frac{q(n)}{n}+x)^2}dx$$

$$\geqq\frac{q(n)}{2n^2}\int_0^\infty\frac{\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p\geqq(x^{1/p}+|\operatorname{med}(X_1)|)^p\Big\}}{(\frac{q(n)}{n}+x)^2}dx$$

$$\geqq \frac{q(n)}{2n^2} \int_0^\infty \frac{\mathbf{P}\left\{ \max_{1 \leqq i \leqq n} |X_i|^p \geqq 2^p x + 2\mathrm{med}(X_1)|^p \right\}}{\left(\frac{q(n)}{n} + x\right)^2} dx$$

(2.48)

$$\geqq 2^{-p-1} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} - \frac{|2\mathrm{med}(X_1)|^p}{q(n)}.$$

According to Lemma 2.5, (1.13) implies (2.35). Using now (2.30), we get

$$\sum_{1 \leqq n < \infty} \frac{m_n^p(\alpha)}{q(n)} < \infty. \tag{2.49}$$

By definition, $m_n(\alpha) \leqq m_{n+1}(\alpha)$ for all α, and, therefore, (2.59) implies

$$\sum_{1 \leqq n < \infty} \frac{1}{q(n)} < \infty. \tag{2.50}$$

Putting together (2.47), (2.49) and (2.50), we obtain

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty,$$

which completes proof of the equivalence of (1.13) and (1.17).

PROOF OF THEOREM 1.4. It follows immediately from the proof of Theorem 1.3.

3. Applications and discussion

We show first that Theorem A remains true if the Wiener process is replaced by partial sums of i.i.d.r.v.'s with finite variance.

THEOREM 3.1. *Let* $0 < p < \infty$. *We assume that* (1.4)–(1.6) *hold true and that* $\{X_n, n \geqq 1\}$ *are i.i.d.r.v.'s with* $\mathbf{E}X_1 = 0$ *and* $0 < \operatorname{var} X_1 < \infty$. *Then* (1.13), (1.14) *and*

$$\sum_{1 \leqq n < \infty} \frac{n^{p/2}}{q(n)} < \infty \tag{3.1}$$

are equivalent.

PROOF. The central limit theorem implies

$$\lim_{n\to\infty} n^{-1/2}\mu_n(\alpha)=\gamma(\alpha)>0 \quad \text{for all } 0<\alpha<1. \tag{3.2}$$

Thus, (3.1) holds if and only if

$$\sum_{1\leqq n<\infty} \frac{\mu_n^p(\alpha)}{q(n)}<\infty. \tag{3.3}$$

Using Lemma 2.2, we obtain

$$\mathbf{E}\frac{\max\limits_{1\leqq i\leqq n}|X_i|^p}{q(n)+n\max\limits_{1\leqq i\leqq n}|X_i|^p}=\frac{q(n)}{n^2}\int_0^\infty \frac{\mathbf{P}\left\{\max\limits_{1\leqq i\leqq n}|X_i|^p\geqq x\right\}}{\left(\frac{q(n)}{n}+x\right)^2}dx$$

$$\leqq \frac{q(n)}{n^2}\left\{\int_0^{n^{p/2}}\left(\frac{q(n)}{n}+x\right)^{-2}dx+\mathbf{P}\left\{\max_{1\leqq i\leqq n}|X_i|^p\geqq n^{p/2}\right\}\int_{n^{p/2}}^\infty\left(\frac{q(n)}{n}+x\right)^{-2}dx\right\}$$

$$\leqq \frac{n^{p/2}}{q(n)}+\mathbf{P}(|X_1|\geqq n^{1/2}),$$

and it is well known that $\mathbf{E}X_1^2<\infty$ if and only if

$$\sum_{1\leqq n<\infty}\mathbf{P}(|X_1|\geqq n^{1/2})<\infty.$$

Theorem 3.1 remains true for not necessarily identically distributed r.v.'s.

THEOREM 3.2. *Let* $0<p<\infty$. *We assume that* (1.3)–(1.6) *hold true and that* $\mathbf{E}X_n=0$, $\mathbf{E}X_n^2<\infty$. *Set* $\sigma_n^2=\sum_{1\leqq i\leqq n}\mathbf{E}X_i^2$. *Then*

$$\sum_{1\leqq n<\infty}\sigma_n^p/q(n)<\infty$$

and

$$\sum_{1\leqq n<\infty}\frac{1}{n}\sum_{1\leqq i\leqq n}\mathbf{P}\{|X_i|\geqq\sigma_n\}<\infty$$

imply (1.14).

PROOF. It goes along the lines of the proof of Theorem 3.1 and hence omitted.

If $\mathbf{E}X_1^2=\infty$, then, in general, condition (3.1) is not enough to have (1.13).

PROPOSITION 3.1. *Let $g(x)$ be a non-decreasing continuous function satisfying $\lim_{x\to\infty} g(x)=\infty$. For any $0<p<\infty$ we can find a sequence $\{X_n, n\geqq 1\}$ of i.i.d.r.v.'s with $\mathbf{E}X_1=0$, $\mathbf{E}X_1^2/g(|X_1|)<\infty$ and a sequence $\{q(n), n\geqq 1\}$ satisfying* (1.4)–(1.6) *such that* (3.1) *holds true but* (1.13) *does not.*

PROOF. We can find a sequence $\{n_k, k\geqq 1\}$ such that

$$n_{k+1}/n_k \geqq 2k \quad \text{and} \quad g(n_k)\geqq e^k, \quad k=1,2,\dots. \tag{3.4}$$

The distribution of X_1 is given by

$$\mathbf{P}\{X_1=\pm n_k\}=\frac{C_6}{n_k k^{1/2}}, \quad k=1,2,\dots,$$

and $\{X_n, n\geqq 1\}$ are independent copies of X_1. Let

$$h(i)=k^{p/4}, \quad \text{if } n_k^2\leqq i<n_{k+1}^2.$$

There is a sequence $\{q(n), n\geqq 1\}$ such that (1.4)–(1.6) hold, and

$$\sum_{1\leqq i<\infty}\frac{i^{p/2}}{q(i)}<\infty, \tag{3.5}$$

$$\sum_{1\leqq i<\infty}\frac{i^{p/2}h(i)}{q(i)}=\infty. \tag{3.6}$$

One can easily verify that $\mathbf{E}X_1=0$, $\mathbf{E}X_1^2=\infty$ and $\mathbf{E}X_1^2/g(|X_1|)<\infty$.

Let $L(x)=\mathbf{E}X_1^2 I\{|X_1|\leqq x\}$. Since $\mathbf{E}X_1^2=\infty$, we have $\lim_{x\to\infty} L(x)=\infty$. By (3.4) we have also

$$|L(ax)-L(x)|=\sum_{k\in U}\frac{2C_6 n_k^2}{n_k^2 k^{1/2}}\leqq\frac{|\log a|}{\log 2}(\inf\{k\colon n_k\geqq x\})^{-1/2}\to 0$$

for all $a>0$, where $U=\{k\colon n_k$ is between x and $ax\}$. Thus, we conclude that L is a slowly varying function at zero. Using Theorem 8.3.1 in Bingham et al. [1], we get

$$\lim_{n\to\infty}\mathbf{P}\Big\{(nL(n^{1/2}))^{-1/2}\sum_{1\leqq i\leqq n}X_i\leqq x\Big\}=\Phi(x)$$

for all x, where Φ is the standard normal distribution function. Thus we get

$$\lim_{n\to\infty}\frac{\operatorname{med}|S_n|}{(nL(n^{1/2}))^{1/2}}>0. \tag{3.7}$$

Next we show that

$$\sum_{1\leqq n<\infty}\frac{(\text{med}\,|S_n|)^p}{q(n)}=\infty. \tag{3.8}$$

Using (3.4) and (3.6) we get

$$\sum_{n_1^2\leqq i<\infty}\frac{(i^{1/2}L(i^{1/2}))^p}{q(i)}=\sum_{1\leqq k<\infty}\sum_{n_k^2\leqq i\leqq n_{k+1}^2-1}\frac{(i^{1/2}L(i^{1/2}))^p}{q(i)}$$
$$\geqq\sum_{1\leqq k<\infty}\sum_{n_k^2\leqq i\leqq n_{k+1}^2-1}\frac{(i^{1/2}L(n_k))^p}{q(i)}$$
$$\geqq\sum_{1\leqq k<\infty}\sum_{n_k^2\leqq i\leqq n_{k+1}^2-1}\frac{i^{p/2}}{q(i)}\Big(\sum_{1\leqq j\leqq k}\frac{C_7}{j^{1/2}}\Big)^p$$
$$\geqq C_8\sum_{1\leqq k<\infty}\sum_{n_k^2\leqq i\leqq n_{k+1}^2-1}\frac{i^{p/2}}{q(i)}k^{p/2}$$
$$\geqq C_8\sum_{1\leqq k<\infty}\frac{i^{p/2}h(i)}{q(i)}=\infty,$$

and hence (3.7) implies (3.8). The distribution of X_1 is symmetric, therefore (1.12) holds true. Applying Theorem 1.3, (3.8) yields

$$\mathbf{P}\Big\{\sum_{1\leqq n<\infty}\frac{|S_n|^p}{q(n)}=\infty\Big\}>0. \tag{3.9}$$

Proposition 3.1 now follows from (3.5) and (3.9).

It is interesting to note that the condition (3.1) is not sufficient for (1.13). We saw in the proof of Theorem 3.1 that $\sum_{1\leqq n<\infty}\frac{\mu_n^p(\alpha)}{q(n)}<\infty$ implies $\sum_{1\leqq n<\infty}\mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{q(n)+n\max_{1\leqq i\leqq n}|X_i|^p}<\infty$, if $\mathbf{E}X_1^2<\infty$, in case of i.i.d.r.v.'s. However, if $\mathbf{E}X_1^2=\infty$, then $\sum_{1\leqq n<\infty}\mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{q(n)+n\max_{1\leqq i\leqq n}|X_i|^p}<\infty$ is the necessary and sufficient condition.

THEOREM 3.3. *Let* $0<p<\infty$. *We assume that* (1.4)–(1.6) *hold true and that* $\{X_n,n\geqq 1\}$ *are i.i.d.r.v.'s. We assume that there is a sequence*

$\{a(n), n \geqq 1\}$ *such that*

$$\frac{S(n)}{a(n)} \xrightarrow{\mathcal{D}} Y,$$

where Y *is a non-degenerate stable r.v. with exponent* $0<\nu<2$. *Then* (1.13), (1.14) *and*

$$\sum_{1 \leqq n < \infty} \mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} < \infty \tag{3.11}$$

are equivalent.

Proof. Let $G(t) = \mathbf{P}\{|X_1| \geqq t\}$. According to Theorem 8.3.1 in Bingham et al. [1], (3.10) implies that $G(t)$ is a regularly varying function with index $-\nu$ with some $0<\nu<2$. Also, if G^{-1} denotes the inverse of G, then we have as well

$$0 < \lim_{n \to \infty} \frac{G^{-1}(1/n)}{a(n)} < \infty. \tag{3.12}$$

Using (3.10), we get that (1.12) holds true, and

$$\lim_{n \to \infty} \frac{\mu_n(\alpha)}{a(n)} = \gamma(\alpha). \tag{3.13}$$

Thus we have

$$\sum_{1 \leqq n < \infty} \frac{\mu_n^p(\alpha)}{q(n)} < \infty \tag{3.14}$$

if and only if

$$\sum_{1 \leqq n < \infty} \frac{(G^{-1}(1/n))^p}{q(n)} < \infty. \tag{3.15}$$

Using Lemma 2.2 we obtain

$$\mathbf{E} \frac{\max_{1 \leqq i \leqq n} |X_i|^p}{q(n) + n \max_{1 \leqq i \leqq n} |X_i|^p} = \frac{q(n)}{n^2} \int_0^\infty \frac{\mathbf{P}\left\{\max_{1 \leqq i \leqq n} |X_i|^p \geqq x\right\}}{\left(\frac{q(n)}{n} + x\right)^2} dx$$

$$\geqq \frac{q(n)}{n^2} \int_0^{(G^{-1}(1/n))^p} \mathbf{P}\left\{\max_{1 \leqq i \leqq n} |X_i|^p \geqq x\right\} \left(\frac{q(n)}{n} + x\right)^{-2} dx$$

$$\geqq \mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq (G^{-1}(1/n))^p\Big\}\frac{(G^{-1}(1/n))^p}{q(n)+n(G^{-1}(1/n))^p}.$$

Since G has a regularly varying tail, we get

$$\lim_{n\to\infty}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq (G^{-1}(1/n))^p\Big\}=1-e^{-1}.$$

Thus, (3.11) yields

$$\sum_{1\leqq n<\infty}\frac{(G^{-1}(1/n))^p}{q(n)+n(G^{-1}(1/n))^p}<\infty,$$

which immediately implies (3.15) and (3.14).

So, if $\{X_n, n\geqq 1\}$ are i.i.d.r.v.'s in the domain of attraction of a stable law, then the necessary and sufficient condition for having (1.13) and (1.14) is given in terms of $\max_{1\leqq i\leqq n}|X_i|$. This observation rhymes very well with the main results in Csörgő, Csörgő, Horváth and Mason [4], and Csörgő, Horváth and Mason [6]. They showed that $S(n)/G^{-1}(1/n)$ is asymptotically stable on account of $\max_{1\leqq i\leqq n}|X_i|$ being very large. Namely, $S(n)$ and $\max_{1\leqq i\leqq n}|X_i|$ are of the same order asymptotically. However, it is still of interest to see whether (3.11) can be replaced by a straightforward moment condition.

THEOREM 3.4. *Let $2\leqq p<\infty$. We assume that $\{X_n, n\geqq 1\}$ are i.i.d.r.v.'s and $\theta>1+p/2$, $\theta\neq 1+p$. Then the following statements are equivalent:*

(3.16)
$$\sum_{1\leqq n<\infty}|S_n|^p/n^\theta<\infty \quad a.s.,$$

(3.17)
$$\sum_{1\leqq n<\infty}\max_{1\leqq i\leqq n}|S_i|^p/n^\theta<\infty \quad a.s.$$

and

(3.18)
$$\mathbf{E}|X_1|^{p/(\theta-1)}<\infty \quad and \quad \mathbf{E}X_1=0 \quad if \ \ \theta<1+p.$$

PROOF. First we show that (3.16) implies (3.18). By Theorem 1.4 we have

$$\sum_{1\leqq n<\infty}\mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{n^\theta+n\max_{1\leqq i\leqq n}|X_i|^p}<\infty.$$

It is easy to see

$$I\Big\{\max_{1\leqq i\leqq n}|X_i|^p\geqq n^{\theta-1}\Big\}\frac{n^{\theta-1}}{2n^\theta}\leqq\frac{\max_{1\leqq i\leqq n}|X_i|^p}{n^\theta+n\max_{1\leqq i\leqq n}|X_i|^p},$$

and, therefore, we get

$$\sum_{1\leqq n<\infty} \frac{1}{n}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\}<\infty.$$

Elementary calculations show

$$\begin{aligned}
&\sum_{1\leqq n<\infty} \frac{1}{n}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\}\\
&=\sum_{0\leqq k<\infty}\sum_{2^k\leqq n<2^{k+1}} \frac{1}{n}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\}\\
&\geqq \sum_{0\leqq k<\infty} 2^{-k}\sum_{2^k\leqq n<2^{k+2}} \mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\}\\
&\geqq \sum_{0\leqq k<\infty} \mathbf{P}\Big\{\max_{1\leqq i\leqq 2^{k+1}}|X_i|^p \geqq 2^{(k+1)}(\theta-1)\Big\}.
\end{aligned}$$

Therefore, we have

$$\sum_{1\leqq k<\infty} \mathbf{P}\Big\{\max_{1\leqq i\leqq 2^k}|X_i|^p \geqq 2^{k(\theta-1)}\Big\}<\infty. \tag{3.19}$$

Hence, $\lim_{k\to\infty}\mathbf{P}\Big\{\max_{1\leqq i\leqq 2^k}|X_i|^p \geqq 2^{2(\theta-1)}\Big\}=0$, and thus we get

$$\lim_{k\to\infty}\mathbf{P}\Big\{\max_{1\leqq i\leqq 2^k}|X_i|^p \geqq 2^{k(\theta-1)}\Big\}/(2^k\mathbf{P}\{|X_1|^p\geqq 2^{k(\theta-1)}\})=1.$$

Using (3.19), we obtain

$$\sum_{1\leqq k<\infty} 2^k\mathbf{P}\{|X_1|^p\geqq 2^{k(\theta-1)}\}<\infty. \tag{3.20}$$

From (3.20) we can easily derive

$$\sum_{1\leqq n<\infty}\mathbf{P}\{|X_1|>n^{\frac{\theta-1}{p}}\}<\infty,$$

which is equivalent to $\mathbf{E}|X_1|^{\frac{p}{\theta-1}}<\infty$. If $\theta<1+p$, then $\mathbf{E}|X_1|<\infty$ and, therefore, the strong law of large numbers and 3.16 imply that $\mathbf{E}X_1=0$.

Now we show that (3.18) implies (3.17). By Lemma 2.2 we have

$$\mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{n^\theta+n\max_{1\leqq i\leqq n}|X_i|^p}=n^{\theta-2}\int_0^\infty\frac{\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p\geqq x\Big\}}{(n^{\theta-1}+x)^2}dx$$

$$= n^{\theta-2}\int_0^{n^{\theta-1}} (n^{\theta-1}+x)^{-2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq x\Big\}dx \tag{3.21}$$
$$+ n^{\theta-2}\int_{n^{\theta-1}}^{\infty} (n^{\theta-1}+x)^{-2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq x\Big\}dx$$
$$= C_n^{(1)} + C_n^{(2)}.$$

It is elementary to check

$$C_n^{(2)} \leqq n^{-1}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\} \leqq \mathbf{P}\Big\{|X_1| \geqq n^{\frac{\theta-1}{p}}\Big\}. \tag{3.22}$$

Next we note

$$C_n^{(1)} \leqq n^{\theta-2}\int_0^{n^{\theta-1}} (n_x^{\theta-1})^{-2}\mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|^p \geqq n^{\theta-1}\Big\}dx$$
$$+ n^{\theta-2}\int_0^{n^{\theta-1}} (n^{\theta-1}+x)^{-2}\mathbf{P}\Big\{x \leqq \max_{1\leqq i\leqq n}|X_i|^p \leqq n^{\theta-1}\Big\}dx \tag{3.23}$$
$$\leqq \mathbf{P}\{|X_1|^p \geqq n^{\theta-1}\} + n^{1-\theta}\int_0^{n^{\theta-1}} \mathbf{P}\{x \leqq |X_1|^p \leqq n^{\theta-1}\}dx$$

and

$$n^{1-\theta}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}} I\{|X_1|^p \leqq n^{\theta-1}\}) \int_0^{n^{1-\theta}} x^{-\frac{1}{p}\left(1+\frac{p}{2(\theta-1)}\right)}dx \tag{3.24}$$
$$\leqq \frac{1}{1-(\frac{1}{p}+\frac{1}{2(\theta-1)})} n^{-(\theta-1)\left(\frac{1}{p}+\frac{1}{2(\theta-1)}\right)}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}} I\{|X_1|^p \leqq n^{\theta-1}\}),$$

on account of $\frac{1}{p}+\frac{1}{2(\theta+1)} < 1$. The assumptions of Theorem 3.4 give that $(\theta-1)(\frac{1}{p}+\frac{1}{2(\theta-1)}) > 1$, and hence we arrive at

$$\sum_{1\leqq n<\infty} n^{-(\theta-1)\left(\frac{1}{p}+\frac{1}{2(\theta-1)}\right)}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}} I\{|X_1|^p \leqq n^{\theta-1}\}$$
$$= \sum_{1\leqq n<\infty} n^{-(\theta-1)\left(\frac{1}{p}+\frac{1}{2(\theta-1)}\right)} \sum_{1\leqq j\leqq n}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}} I\{(j-1)^{\theta-1} < |x_1|^P \leqq j^{\theta}\})$$

$$(3.25) \leqq \sum_{1\leqq j<\infty}\sum_{j\leqq n<\infty} n^{-(\theta-1)\left(\frac{1}{p}+\frac{1}{2(\theta-1)}\right)}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}}I\{(j-1)^{\theta-1}<|X_1|^p\leqq j^\theta\})$$

$$\leqq \frac{1}{\frac{\theta-1}{p}-\frac{1}{2}}\sum_{1\leqq j<\infty} j^{-\frac{\theta-1}{p}+\frac{1}{2}}\mathbf{E}(|X_1|^{1+\frac{p}{2(\theta-1)}}I\{(j-1)^{\theta-1}<|X_1|^p\leqq j^{\theta-1}\})$$

$$\leqq \frac{2p}{2(\theta-1)-p}\sum_{1\leqq j<\infty}\mathbf{E}(|X_1|^{\frac{p}{\theta-1}}I\{(j-1)^{\theta-1}<|X_1|^p\leqq j^{\theta-1}\})$$

$$=\frac{2p}{2(\theta-1)-p}\mathbf{E}|X_1|^{\frac{p}{\theta-1}}.$$

Combining (3.21)–(3.25), we obtain

$$(3.26)\qquad \sum_{1\leqq n<\infty}\frac{\max\limits_{1\leqq i\leqq n}|X_i|^p}{q(n)+n\max\limits_{1\leqq i\leqq n}|X_i|^p}<\infty.$$

Next we show

$$(3.27)\qquad \sum_{1\leqq n<\infty}\frac{m_n^p(\alpha)}{n^\theta}<\infty.$$

First we note

$$(3.28)\qquad \lim_{n\to\infty} n\mathbf{P}\{|X_1|^p\geqq n^{\theta-1}\}=0,$$

and, by the Markov inequality, we have

$$\mathbf{P}\{|S_n|\geqq x\}\leqq \mathbf{P}\Big\{\max_{1\leqq i\leqq n}|X_i|\geqq n^{\frac{\theta-1}{p}}\Big\}$$

$$(3.29)\qquad +\mathbf{P}\Big\{\Big|\sum_{1\leqq i\leqq n}X_iI\{|X_i|\leqq n^{(\theta-1)/p}\}\Big|\geqq x\Big\}$$

$$\leqq n\mathbf{P}\{|X_1|^p\geqq n^{\theta-1}\}+x^{-2}\{n\mathbf{E}(X_1^2I\{|X_1|^p\leqq n^{\theta-1}\})$$

$$+n\mathbf{E}(X_1I(|X_1|^p\leqq n^{\theta-1}))\}.$$

It follows from (3.28) and (3.29) that

$$\limsup_{n\to\infty} m_n(\alpha)\{n\mathbf{E}(X_1^2I\{|X_1|^p\leqq n^{\theta-1}\})+n\mathbf{E}(X_1I\{|X_1|^p\leqq n^{\theta-1}\})\}^{-1/2}<\infty.$$

Thus, it is enough to verify

$$(3.30)\qquad \sum_{1\leqq n<\infty} n^{-\theta}(n\mathbf{E}(X_1^2I\{|X_1|\leqq n^{(\theta-1)/p}\}))^{p/2}<\infty$$

and

$$\sum_{1 \leqq n < \infty} n^{-\theta} |n\mathbf{E}(X_1 I\{|X_1| \leqq n^{(\theta-1)/p}\})|^p < \infty. \tag{3.31}$$

Using the Hölder inequality, we obtain

$$\begin{aligned}
&\mathbf{E}(X_1^2 I\{|X_1| \leqq n^{(\theta-1)/p}\}) \\
&\leqq (\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{p}} (\mathbf{E}(|X_1|^{(2-\frac{p-2}{\theta-1})\frac{p}{2}} I\{|X_1| \leqq n^{(\theta-1)/p}\}))^{\frac{2}{p}} \\
&\leqq (\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{p}} (\mathbf{E}(|X_1|^2 I\{|X_1| \leqq n^{\frac{\theta-1}{p}}\}))^{\frac{2}{p}} n^{-\frac{(p-2)(\theta-1)}{p^2}(2-\frac{p}{\theta-1})},
\end{aligned}$$

and therefore, we get

$$\begin{aligned}
&\sum_{1 \leqq n < \infty} n^{-\theta} (n\mathbf{E}(X_1^2 I\{X_1| \leqq n^{(\theta-1)/p}\}))^{p/2} \\
&\leqq (\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{2}} \sum_{1 \leqq n < \infty} n^{\frac{p}{2}-\theta+(2-\frac{p}{\theta-1})\frac{(\theta-1)(p-2)}{2p}} \mathbf{E}(X_1^2 I\{|X_1| \leqq n^{\frac{\theta-1}{p}}\}) \\
&\leqq (\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{2}} \sum_{1 \leqq n < \infty} n^{-\frac{2}{p}(\theta-1)} \sum_{1 \leqq j \leqq n} \mathbf{E}(X_1^2 I\{(j-1)^{\frac{\theta-1}{p}} < |X_1| \leqq j^{\frac{\theta-1}{p}}\}) \\
&\leqq (\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{2}} \sum_{1 \leqq j < \infty} \sum_{j \leqq n < \infty} n^{-\frac{2}{p}(\theta-1)} \mathbf{E}(X_1^2 I\{j-1 < |X_1|^{\frac{p}{\theta-1}} \leqq j\}) \\
&\leqq \frac{p(\mathbf{E}|X_1|^{\frac{p}{\theta-j}})^{\frac{p-2}{2}}}{2(\theta-1)-p} \sum_{1 \leqq j < \infty} j^{1-\frac{2}{p}(\theta-1)} \mathbf{E}(X_1^2 I\{j-1 < |X_1|^{\frac{p}{\theta-1}} \leqq j\}) \\
&\leqq \frac{p(\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p-2}{2}}}{2(\theta-1)-p} \sum_{1 \leqq j < \infty} \mathbf{E}(|X_1|^{\frac{p}{\theta-1}} I\{j-1 < |X_1|^{\frac{p}{\theta-1}} \leqq j\}) \\
&\leqq \frac{p(\mathbf{E}|X_1|^{\frac{p}{\theta-1}})^{\frac{p}{2}}}{2(\theta-1)-p}.
\end{aligned}$$

The proof of (3.31) is similar to that of (3.30) and hence omitted.

It may seem to be curious that the case of $\theta = 1 + p$ is excluded in Theorem 3.4. However, this is not due to the lack of power of our method. Indeed, Theorem 3.4 may fail if $\theta = 1 + p$.

EXAMPLE 3.1. Let $\{X_n, n \geqq 1\}$ be i.i.d.r.v.'s with distribution

$$\mathbf{P}\{X_1 = -1\} = c_0,$$

$$\mathbf{P}\{X_1 \geqq x\} = c_0 \int_x^\infty \frac{1}{y^2(\log y)(\log\log y)^2} dy, \quad \text{if } x \geqq e^e,$$

where

$$c_0=\left(1+\int_{e^e}^{\infty}\frac{1}{y^2(\log y)(\log\log y)^2}dy\right)^{-1}.$$

We show that this sequence of r.v.'s has the following properties:

$$\mathbf{E}X_1=0, \tag{3.32}$$

$$\sum_{1\le n<\infty} n^{-(1+p)}(\mathrm{med}|S_n|)^p=\infty \quad \text{for all} \quad p>0, \tag{3.33}$$

and

$$\sum_{1\le n<\infty} n^{-(1+p)}\max_{1\le i\le n}|S_i|^p=\infty \quad \text{a.s.}. \tag{3.34}$$

The definition of X_1 implies (3.32) without any calculations. By Theorem 1.2, (3.34) follows from (3.33). Hence, it suffices to verify only (3.33). Let $F(x)=\mathbf{P}\{X_1\le x\}$ and $L(x)=c_0/((\log x)(\log\log x)^2)$. Clearly,

$$\lim_{x\to\infty}\frac{x(1-F(x)+F(-x))}{L(x)}=1,$$

and

$$\lim_{x\to\infty}\frac{F(-x)}{1-F(x)+F(-x)}=0.$$

Therefore, F belongs to the domain of attraction of a stable law with exponent one (cf. Theorem 8.3.1 in Bingham et al. [1]). This means that

$$\frac{\sum_{1\le i\le n} X_i-b(n)}{a(n)}\xrightarrow{\mathcal{D}}Y, \tag{3.35}$$

where Y is a random variable with characteristic function

$$f(t)=\exp\left(-|t|-i\frac{2}{\pi}t\log|t|\right),$$

and

$$a(n)=\frac{\pi n}{2c_0(\log\log n)^2\log n},$$

$$b(n)=na(n)\int_{-\infty}^{\infty}\sin(x/a(n))dF(x)$$

(cf. Theorem 8.3.2 in Bingham et al. [1] or Mijnheer [7], p. 16). We note that

$$b(n)=na(n)\int_{-\infty}^{\infty}\Big(\sin\frac{x}{a(n)}-\frac{x}{a(n)}\Big)dF(x)$$

$$=c_0na(n)\Big(-\sin\Big(\frac{1}{a(n)}\Big)+\frac{1}{a(n)}+\int_{e^e}^{\infty}\frac{\sin\frac{y}{a(n)}-\frac{y}{a(n)}}{y^2(\log\log y)^2\log y}dy\Big),$$

and, therefore, we have

$$b(n)\geqq c_0na(n)\Big\{\int_{e^e}^{a(n)}\frac{-\frac{1}{6}(y/a(n))^3}{y^2(\log\log y)^2\log y}dy$$

$$+\int_{a(n)}^{\infty}\frac{-2y}{a(n)y^2(\log\log y)^2\log y}dy\Big\}=D_n^{(1)},$$

$$b(n)\leqq c_0na(n)\Big\{\frac{1}{6a^3(n)}+\int_{2a(n)}^{\infty}\frac{-y}{2a(n)y^2(\log\log y)^2\log y}dy\Big\}=D_n^{(2)}.$$

Elementary properties of slowly varying functions give

$$\lim_{n\to\infty}D_n^{(1)}\Big/\Big(\frac{-2c_0n}{\log\log a(n)}\Big)=1,$$

$$\lim_{n\to\infty}D_n^{(2)}\Big/\Big(\frac{-c_0n}{2\log\log a(n)}\Big)=1,$$

and thus we conclude

$$-2c_0\leqq\liminf_{n\to\infty}\frac{b(n)\log\log n}{n}\leqq\limsup_{n\to\infty}\frac{b(n)\log\log n}{n}\leqq-\frac{c_0}{2}. \tag{3.36}$$

Using (3.35) and (3.36), we obtain

$$\liminf_{n\to\infty}\mathbf{P}\Big\{\Big|\sum_{1\leqq i\leqq n}X_i\Big|\geqq\frac{c_0n}{3\log\log n}\Big\}$$

$$\geqq\liminf_{n\to\infty}\mathbf{P}\left\{\frac{\sum_{1\leqq i\leqq n}X_i-b(n)}{a(n)}\leqq\frac{-b(n)-c_0n/(3\log\log n)}{a(n)}\right\}$$

$$\begin{aligned}
&\geqq \liminf_{n\to\infty} \mathbf{P}\left\{\frac{\sum_{1\leqq i\leqq n} X_i - b(n)}{a(n)} \leqq \frac{c_0 n}{7a(n)\log\log n}\right\} \\
(3.37)\qquad &= \liminf_{n\to\infty} \mathbf{P}\left\{\frac{\sum_{1\leqq i\leqq n} X_i - b(n)}{a(n)} \leqq \frac{2c_0(\log\log n)\log n}{7\pi}\right\} \\
&= 1.
\end{aligned}$$

Now (3.37) yields

$$\liminf_{n\to\infty} \frac{\mathrm{med}|S_n|}{(n/\log\log n)} \geqq \frac{c_0}{3},$$

which implies (3.33) immediately.

Only Theorem 3.4 has a restriction on the value of p. The following example shows that Theorem 3.4 may not be true if $0<p<2$.

EXAMPLE 3.2. Let $0<p<2$ and $1+p/2<\theta<2$. Let $\{X_n, n\geqq 1\}$ be i.i.d.r.v.'s with density function

$$g(x) = \begin{cases} a_0|x|^{1+\frac{p}{\theta-1}}(\log|x|)(\log\log|x|)^2, & \text{if } |x|\geqq e^e \\ 0, & \text{if } |x|<e^e. \end{cases}$$

We establish the following properties of $\{X_n, n\geqq 1\}$:

$$X_1 \text{ is symmetric,} \tag{3.38}$$

$$\mathbf{E}|X_1|^{\frac{p}{\theta-1}} < \infty \quad \text{and} \quad \mathbf{E}X_1 = 0 \quad \text{if } \theta\leqq 1+p, \tag{3.39}$$

$$\sum_{1\leqq n<\infty} \mathbf{E}\frac{\max_{1\leqq i\leqq n}|X_i|^p}{n^\theta + n\max_{1\leqq i\leqq n}|X_i|^p} = \infty, \tag{3.40}$$

and

$$\sum_{1\leqq n<\infty} n^{-\theta}|S_n|^p = \infty \quad \text{a.s.}. \tag{3.41}$$

Elementary calculations give (3.38) and (3.39). Applying Theorem 1.3, we can see that (3.40) implies (3.41). So we need to verify (3.40) only. Let

$$a(n) = \left(\frac{6a_0(\theta-1)n}{p(\log\log n)^2\log n}\right)^{\frac{\theta-1}{p}}.$$

It is easy to see

$$\lim_{x\to\infty} \mathbf{P}\{|X_1| \geqq x\} \Big/ \Big(\frac{2a_0(\theta-1)}{px^{p/(\theta-1)}(\log\log x)^2 \log x}\Big) = 1,$$

and hence we have

$$\lim_{n\to\infty} n\mathbf{P}\{|X_1| \geqq a(n)\} = \frac{1}{3}.$$

If $0 < x \leqq a(n)$, then we obtain

$$\mathbf{P}\Big\{\max_{1\leqq i\leqq n} |X_i| \geqq x\Big\} \geqq \mathbf{P}\Big\{\max_{1\leqq i\leqq n} |X_i| \geqq a(n)\Big\} \geqq \frac{1}{6},$$

assuming that n is large enough. Using Lemma 2.2, we get

$$\mathbf{E}\frac{\max_{1\leqq i\leqq n} |X_i|^p}{n^\theta + n \max_{1\leqq i\leqq n} |X_i|^p} \geqq pn^{\theta-2} \int_0^{a(n)} \frac{x^{p-1}\mathbf{P}\Big\{\max_{1\leqq i\leqq n} |X_i| \geqq x\Big\}}{(n^{\theta-1}+x^p)^2} dx$$

$$\geqq \frac{pn^{\theta-2}}{6} \int_0^{a(n)} \frac{x^{p-1}}{(n^{\theta-1}+x^p)^2} dx = \frac{a^p(n)}{6n(n^{\theta-1}+a^p(n))}$$

$$\geqq \frac{1}{12}\Big(\frac{6a_0(\theta-1)}{p}\Big)^{\theta-1} \frac{1}{n((\log n)(\log\log n)^2)^{\theta-1}},$$

if n is large. Since $\theta < 2$, the proof of (3.40) is complete.

ACKNOWLEDGEMENT. Miklós Csörgő was supported by an NSERC Canada grant at Carleton University. Qi-Man Shao was supported by the Fok Ying-tung Education Foundation and by an NSERC Canada Scientific Exchange Award at Carleton University, Ottawa, Canada.

REFERENCES

[1] BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L., *Regular variation*, Encyclopedia of mathematics and its applications, Vol. 27, Cambridge University Press, Cambridge – New York, 1989. *MR* **90i**:26003

[2] BREIMAN, L., *Probability*, Addison-Wesley, Reading, Mass., 1968. *MR* **37** #4841. Corrected reprint in: Classics in applied mathematics, Vol. 7, SIAM, Philadelphia, PA, 1992. *MR* **93d**:60001

[3] CHOW, Y. S. and TEICHER, H., *Probability theory*, Independence, interchangeability, martingales, Springer-Verlag, New York – Heidelberg, 1978. *MR* **80a**:60004. Second edition in: Springer texts in statistics, 1988. *MR* **89e**:60001

[4] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D. M., Normal and stable convergence of integral functions of the empirical distribution function, *Ann. Probab.* **14** (1986), 86–118. *MR* **87d**:60036

[5] CSÖRGŐ, M., HORVÁTH, L. and SHAO, Q.-M., Convergence of integrals of uniform empirical and quantile processes, *Stochastic Process. Appl.* **45** (1993), 283–294. *MR* **94f**:60052

[6] Csörgő, M., Horváth, L. and Mason, D. M., What portion of the sample makes a partial sum asymptotically stable or normal?, *Probab. Theory Relat. Fields* **72** (1986), 1–16. *MR* **87k**:60063

[7] Mijnheer, J. L., *Sample path properties of stable processes*, Math. Centre Tracts No. 59, Mathematisch Centrum, Amsterdam, 1975. *MR* **51** #7008

[8] Peligrad, M., An invariance principle for ϕ-mixing sequences, *Ann. Probab.* **13** (1985), 1304–1313. *MR* **87b**:60056

[9] Rajput, B., Gaussian measures on L_p spaces, $1 \leqq p < \infty$, *J. Multivariate Anal.* **2** (1972), 382–403. *MR* **49** #9896

[10] Shepp, L. A., Radon–Nikodym derivatives of Gaussian measures, *Ann. Math. Statist.* **37** (1966), 321–354. *MR* **32** # 8408

(Received February 20, 1995)

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
1125 COLONEL BY DRIVE
OTTAWA, ONTARIO
K1S 5B6
CANADA

mcsorgo@math.carleton.ca

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UT 84112
U.S.A.

horvath@math.utah.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OREGON
EUGENE, OR 97403
U.S.A.

shao@bright.uoregon.edu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 75–110

ON THE HAUSDORFF DIMENSION OF THE SET GENERATED BY EXCEPTIONAL OSCILLATIONS OF A WIENER PROCESS

P. DEHEUVELS and M. A. LIFSHITS

Dedicated to Endre Csáki on his sixtieth birthday

Abstract

The rescaled h-increments $Y_{t,h}(u) = (2h\log(1/h))^{-1/2}\{W(t+hu) - W(t)\}$, for $u \in [0,1]$, of a Wiener process $\{W(t) : t \geqq 0\}$, are considered as elements of the space $C_0[0,1]$ of all continuous functions g on $[0,1]$ with $g(0) = 0$. We endow $C_0[0,1]$ with the topology defined by a norm $||\cdot||_\nu$ chosen within a general class $\mathcal{C}$ for which the limit law $\lim_{h\downarrow 0}\{\sup_{0\leqq t\leqq 1}||Y_{t,h}||_\nu\} < \infty$ holds with probability 1. We show that, for each $f \in C_0[0,1]$ with $\int_0^1\{\frac{d}{du}f(u)\}^2du \leqq 1$, the set $\mathcal{L}_\nu(f) = \{t \in [0,1] : \liminf_{h\downarrow 0}||Y_{t,h} - f||_\nu = 0\}$ contains, with probability 1 for each $\nu \in \mathcal{C}$, a subset $\mathcal{L}(f)$, independent of $||\cdot||_\nu \in \mathcal{C}$ and with Hausdorff dimension equal to $\dim(\mathcal{L}(f)) = 1 - \int_0^1\{\frac{d}{du}f(u)\}^2du$.

1. Introduction and statement of main results

Let $\{W_1(t) : t \geqq 0\}$ denote a standard Wiener process. For each $h > 0$ and $t \geqq 0$, set $X_{t,h}(u) = h^{-1/2}(W(t+hu) - W(t))$, $L_h = L(h) = (2\log(1/h))^{1/2}$ and $Y_{t,h}(u) = L_h^{-1}X_{t,h}(u)$ for $u \geqq 0$. Fix any $t_0 \in [0,1]$. Lévy [14] established that with probability 1,

$$(1.1)\qquad \begin{aligned} &\text{(i)} \quad \lim_{h\downarrow 0}\Big\{\sup_{t\in[0,1]}||Y_{t,h}||_{\mathcal{U}}\Big\} = 1,\\ &\text{(ii)} \quad \limsup_{h\downarrow 0}\Big\{\frac{\log(1/h)}{\operatorname{loglog}(1/h)}\Big\}^{1/2}||Y_{t_0,h}||_{\mathcal{U}} = 1, \end{aligned}$$

where $||f||_{\mathcal{U}} = \sup_{u\in[0,1]}|f(u)|$ stands for the *sup-norm* of f. Orey and Taylor [17] precised (1.1)(i) by proving the existence with probability 1 of a

1991 *Mathematics Subject Classification.* Primary 60F17, 60F15, 60F10, 60G15.
Key words and phrases. Wiener process, law of the iterated logarithm, modulus of continuity, strong laws, Hausdorff dimension, fractals.

*Supported by International Science Foundation and Russian Foundation of Basic Research.

(random) subset $\mathcal{L} \subseteq [0,1]$ such that, for each $t \in \mathcal{L}$,

$$\limsup_{h\downarrow 0} ||Y_{t,h}||_{\mathcal{U}} = 1. \tag{1.2}$$

The following notation is needed for the exposition of refinements of (1.1)–(1.2) obtained in the recent literature. Denote by $C[0,1]$ the set of all continuous functions defined on $[0,1]$, and set $C_0[0,1]=\{g\in C[0,1]:g(0)=0\}$. Aside of the *uniform topology* $\mathcal{U}$ generated by $||f||_{\mathcal{U}}$, there is a large choice of normed topologies on $C_0[0,1]$ (or $C[0,1]$) which are appropriate with respect to the derivation of limit laws for the Wiener process. To characterize a general class including most of the possible norms of interest, we follow the notation and vocabulary of Deheuvels and Lifshits [5], [6]. By a *norm* on a vector space $\mathcal{X}$ (with emphasis on the cases where $\mathcal{X} = C_0[0,1]$ or $\mathcal{X} = C[0,1]$), is meant a mapping $\nu : f \to \nu(f) = ||f||_\nu$ of $\mathcal{X}$ onto $[0,\infty]$ fulfilling (A) below, with the conventions $0\times\infty = 0$ and $0+\infty=\infty$. When it is defined on either $\mathcal{X} = C_0[0,1]$ or $\mathcal{X} = C[0,1]$, ν is said to be a *consistent norm* if it satisfies (A), (B) and (C) below. By a *consistent semi-norm* on $\mathcal{X} = C_0[0,1]$ (resp. $\mathcal{X} = C[0,1]$) is meant a mapping $\nu : f \to \nu(f) = ||f||_\nu$ of $\mathcal{X}$ onto $[0,\infty]$ satisfying (A)(i,ii)–(B)–(C) but not necessarily (A)(iii).

(A) For all $f, g \in \mathcal{X}$ and $c \in \mathbf{R}$

(i) $\quad \nu(f+g) \leqq \nu(f) + \nu(g),$

(ii) $\quad \nu(cf) = |c|\nu(f),$

(iii) $\quad \nu(f) = 0 \Rightarrow f = 0.$

(B) ν is lower semicontinuous with respect to the uniform topology $\mathcal{U}$.

(C) There exists an $\varepsilon > 0$ such that with probability one

$$\sup_{\theta_1,\theta_2\in[0,\varepsilon]} \nu\Big(W(\theta_1 + (1-\theta_2)\cdot) - W(\theta_1)\Big) < \infty. \tag{1.3}$$

The space $\mathcal{X}$ endowed with the topology defined by $\nu = ||\cdot||_\nu$ is denoted by $(\mathcal{X},\nu)$. We note that, with the above definitions a norm (resp. a semi-norm) is *possibly infinite*. Moreover, any norm (resp. semi-norm) ν on $C_0[0,1]$ may be extended to a norm (resp. semi-norm) on $C[0,1]$ by setting $\nu(f) = \infty$ for all $f \in C[0,1] - C_0[0,1]$. Therefore, in the sequel we will not distinguish the case where ν is defined on $C[0,1]$ from that where ν is initially defined on $C_0[0,1]$ only. Throughout, we will denote by $\mathcal{C}$ the set of all consistent norms on $C_0[0,1]$. We will repeatedly make use of the observation that (A)–(C), when combined with the scaling property of the Wiener process, jointly imply that

$$\mathbf{P}(\nu(W) < \infty) = 1. \tag{1.4}$$

We refer to Deheuvels and Lifshits [5], [6] for further discussions and examples of norms which satisfy (A)–(B)–(C). In particular, the *weighted* *sup-norm*, the *Hölder norm* and the *L^p-norm* defined, respectively, for $f \in C_0[0,1]$, $\alpha < 1/2$, $\beta < 1/2$ and $p \geqq 1$, by

$$\sup_{0<t<1} t^{-\alpha}(1-t)^{-\beta}|f(t)|, \qquad \sup_{0\leqq s\neq t\leqq 1} (t-s)^{-\alpha}|f(t)-f(s)|,$$
$$\Big(\int_0^1 |f(u)|^p du\Big)^{1/p},$$

are consistent norms in the above sense. Following the discussion in [5], [6], we mention that the technical assumption (B) may be relaxed in part in the statement of our forthcoming results at the expense of huge technical difficulties. Since, to our best knowledge, there is no interesting norm with respect to the Wiener process which does not satisfy the latter condition, we will limit ourselves to the present framework.

The following notation and facts, taken from probability theory in Banach spaces will be useful. Let $\mathcal{X} = (\mathcal{X}, \tau)$ denote a vector space $\mathcal{X}$, endowed with a Hausdorff locally convex topology τ, and algebra of Borel sets $\mathcal{B}_{\mathcal{X}}$. We denote by $\mathcal{X}^*$ the space of all τ-continuous linear forms on $\mathcal{X}$. An $\mathcal{X}$-valued random variable Z whose distribution $\mathbf{P}_Z(B) = \mathbf{P}(Z \in B)$ for $B \in \mathcal{B}_{\mathcal{X}}$ is a Radon measure on $\mathcal{X}$ is said to be *centered Gaussian* whenever the distribution of $\pi(Z)$ is centered Gaussian for all $\pi \in \mathcal{X}^*$. The Gaussian measure $\mathbf{P}_Z$ allows to imbed $\mathcal{X}^*$ into $L^2 = L^2(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mathbf{P}_Z)$ via the mapping $\mathrm{I}_Z : \pi \in \mathcal{X}^* \to \pi(Z) \in L^2$. The closure of $\mathcal{X}^* = \mathrm{I}_Z(\mathcal{X}^*)$ into L^2 is called the space of *measurable linear forms* (with respect to $\mathbf{P}_Z$) on $\mathcal{X}$ and denoted by $\mathcal{X}_Z^*$. For each centered Gaussian distribution $\mathbf{P}_Z$ on $\mathcal{X}$, there exists a *kernel* $\mathbf{H}_Z$ which is a linear subspace of $\mathcal{X}$, endowed with a Hilbert norm $|\cdot|_{\mathbf{H}_Z}$, such that the following property holds (see e.g. Section 9 in Lifshits [15]). For each $h \in \mathbf{H}_Z$, there exists an $\tilde{h} \in \mathcal{X}_Z^*$ satisfying the equalities

$$\text{(1.5)}\qquad \begin{aligned} &\mathbf{P}_Z(B+h) = \int_B \Big(\tilde{h}(z) - \frac{1}{2}|h|^2_{\mathbf{H}_Z}\Big)\mathbf{P}_Z(dz) \quad \text{for each} \quad B \in \mathcal{B}_Z, \quad \text{and} \\ &\int_{\mathcal{X}} \tilde{h}(z)^2 \mathbf{P}_Z(dz) = |h|^2_{\mathbf{H}_Z}. \end{aligned}$$

The space $\mathbf{H}_Z$ is called the *reproducing kernel Hilbert space* [RKHS] of $\mathbf{P}_Z$. The unit ball of $\mathbf{H}_Z$ will be denoted by $\mathbf{K}_Z = \{h \in \mathbf{H}_Z : |h|_{\mathbf{H}_Z} \leqq 1\}$.

We will specialize in the case where $\mathcal{X} = C_0[0,1]$ (or $C[0,1]$), $\tau = \mathcal{U}$, $Z = W$ is (the restriction on $[0,1]$ of) a Wiener process, and $\mathbf{P}_W$ is the *Wiener measure*. In this case, (1.5) is the Cameron–Martin formula (Cameron and Martin [4]) and $|\cdot|_{\mathbf{H}_W} = |\cdot|_{\mathbf{H}}$ is the *Wiener process Hilbert norm* $|\cdot|_{\mathbf{H}}$, conveniently defined by setting, for each $f \in C[0,1]$

$$|f|_{\mathbf{H}} = \Big(\int_0^1 \dot{f}(s)^2 ds \Big)^{1/2},$$

when $f \in C_0[0,1]$ is absolutely continuous on $[0,1]$ with Lebesgue derivative $\dot{f}(u) = \frac{d}{du} f(u)$, and $|f|_{\mathbf{H}} = \infty$ else. The RKHS of $\mathbf{P}_W$ is then $\mathbf{H}_W = \mathbf{H} = \{f \in C_0[0,1] : |f|_{\mathbf{H}} < \infty\}$, with unit ball $\mathbf{K}_W = \mathbf{K} = \{f \in \mathbf{H} : |f|_{\mathbf{H}} \leqq 1\}$ equal to the *Strassen set* ([21]).

For each consistent semi-norm ν on $C_0[0,1]$, introduce the limit sets

$$\begin{aligned} C_t(\nu) &= \{f \in C_0[0,1] : \liminf_{h \downarrow 0} \nu(Y_{t,h} - f) = 0\}, \\ C(\nu) &= \{f \in C_0[0,1] : \liminf_{h \downarrow 0} \big(\inf_{t \in [0,1]} \nu(Y_{t,h} - f)\big) = 0\}. \end{aligned} \tag{1.6}$$

For each consistent semi-norm ν, $f \in C_0[0,1]$ and $\alpha \geqq 0$ introduce the sets of *exceptional points* (in $[0,1]$, and with respect to ν, f, α and W), defined by

$$\begin{aligned} T(\nu, f) &= \{t \in [0,1] : f \in C_t(\nu)\}, \\ T(\nu, \alpha) &= \{t \in [0,1] : \limsup_{h \downarrow 0} \big(\inf_{f \in \alpha \mathbf{K}} \nu(Y_{t,h} - f)\big) > 0\}. \end{aligned} \tag{1.7}$$

Here, and elsewhere, we set $\lambda K = \{\lambda z : z \in K\}$ for $\lambda \in \mathbf{R}$ and $K \subseteqq \mathcal{X}$. By extending the functional law of the iterated logarithm of Strassen [21] to increments, Révész [18] and Mueller [16] improved (1.1) by showing that, with probability 1,

$$C(\mathcal{U}) = \mathbf{K}.$$

Deheuvels and Lifshits [5] obtained a general version of this statement by showing that, for any consistent norm $\nu \in \mathcal{C}$, we have, with probability 1,

$$C(\nu) = \bigcup_{t \in [0,1]} C_t(\nu) = \mathbf{K}. \tag{1.8}$$

It follows from the versions of (1.1)(ii) which hold with $\mathcal{U}$ replaced by $\nu \in \mathcal{C}$ that, with probability 1 for each specified $t_0 \in [0,1]$, the set $C_{t_0}(\nu)$ contains only of the null function (see [5], [6]). This is not in contradiction with (1.8) because of the continuum cardinality of $[0,1]$. In addition, (1.8) implies that, with probability 1, $T(\nu, f) = \emptyset$ for each $f \in C_0[0,1]$ with $|f|_H > 1$, and $T(\nu, \alpha) = \emptyset$ for each $\alpha \geqq 1$.

On the other hand, it follows from (1.8) that, with probability 1, the sets $T(\nu, f)$ and $T(\nu, \alpha)$ are not empty for $|f|_H \leqq 1$ and $\alpha \in [0,1)$. We will now show that each of these sets, being of Lebesgue measure zero and dense in [0,1], constitutes a *random fractal*, whose Hausdorff dimension is independent of $\nu \in \mathcal{C}$ with probability 1.

We refer to Falconer [9], [10] and Stoyan and Stoyan [20], for expositions of the theory of fractals. Below, we limit ourselves to simple definitions. By a *fractal* subset of $[0,1]$, is meant here a set $A \subseteqq [0,1]$ with an arbitrary *Hausdorff dimension* $\dim(A) \in [0,1]$. The latter is defined by

$$(1.9) \qquad \dim(A) = \inf\Big\{\rho > 0 : \lim_{\delta\downarrow 0}\Big(\inf \sum_{i\in\mathcal{I}} |b_i - a_i|^\rho\Big) = 0\Big\},$$

where the infimum is taken among all $\{(a_i, b_i) : i \in \mathcal{I}\} \in \mathbf{R}^2$ such that $A \subseteqq \bigcup_{i\in\mathcal{I}}[a_i, b_i]$ and $0 < b_i - a_i \leqq \delta$ for each $i \in \mathcal{I}$. The first steps in the evaluation of Hausdorff dimension for sets generated by exceptional oscillations of Wiener processes were made by Orey and Taylor [17]. They established that, for each $\alpha \in [0,1]$, with probability 1,

$$(1.10) \qquad \dim\Big\{t \in [0,1] : \limsup_{h\downarrow 0} ||Y_{t,h}||_{\mathcal{U}} \geqq \alpha\Big\} = 1 - \alpha^2.$$

Deheuvels and Mason [7], [8] obtained the functional version of (1.10) stated in Theorem 1.1 below.

THEOREM 1.1. (a) *For any* $\alpha \in [0,1]$, *with probability* 1,

$$(1.11) \qquad \dim(T(\mathcal{U}, \alpha)) = 1 - \alpha^2.$$

(b) *For any* $f \in \mathbf{K}$, *with probability* 1,

$$(1.12) \qquad \dim(T(\mathcal{U}, f)) = 1 - |f|^2_{\mathbf{H}}.$$

The aim of this paper is three-fold. First, we will extend the validity of Theorem 1.1 to the case where the uniform norm $\mathcal{U}$ is replaced by any norm ν within the class $\mathcal{C}$ of consistent norms. The corresponding result is stated in Theorem 1.2 below in the somewhat more general setting of a countable family $\mathcal{N} = \{\nu_n : n \geqq 1\} \subseteqq \mathcal{C}$ of consistent norms.

THEOREM 1.2. *Let* $\mathcal{N} = \{\nu_n : n \geqq 1\} \subseteqq \mathcal{C}$ *be a countable family of consistent norms. Then, with probability* 1,

$$(1.13) \qquad \dim\Big(\bigcup_{\nu\in\mathcal{N}} T(\nu, \alpha)\Big) = 1 - \alpha^2, \quad \forall \alpha \in [0,1],$$

and

$$(1.14) \qquad \dim\Big(\bigcap_{\nu\in\mathcal{N}} T(\nu, f)\Big) = 1 - |f|^2_{\mathbf{H}}, \quad \forall f \in \mathbf{K}.$$

Second, we will make use of Theorem 1.2 to establish in Theorem 1.3 the existence with probability 1 of *norm-independent* sets of exceptional points (in the sense of (1.6)–(1.7)) with the same fractal dimension than that which is obtained for a single norm $\nu \in \mathcal{C}$.

THEOREM 1.3. *There exist families of random subsets* $\{T'_\alpha : \alpha \in [0,1)\}$, $\{T''_\alpha : \alpha \in [0,1)\}$ *and* $\{T_f, f \in \mathbf{K}\}$ *of* $[0,1]$, *indexed by* $\alpha \in [0,1)$ *and* $f \in \mathbf{K}$, *respectively, such that, with probability* 1,

$$\dim(T'_\alpha) = \dim(T''_\alpha) = 1 - \alpha^2, \qquad \forall \alpha \in [0,1), \tag{1.15}$$

$$\dim(T_f) = 1 - |f|^2_{\mathbf{H}}, \qquad \forall f \in \mathbf{K}, \tag{1.16}$$

and, for each consistent norm $\nu \in \mathcal{C}$, *with probability* 1

$$T'_\alpha \subseteqq T(\nu, \alpha) \subseteqq T''_\alpha, \qquad \forall \alpha \in [0,1), \tag{1.17}$$

$$T_f \subseteqq T(\nu, f), \qquad \forall f \in \mathbf{K}. \tag{1.18}$$

REMARK 1.1. The meaning of the second half of Theorem 1.3 is that, for each specified consistent norm $\nu \in \mathcal{C}$, there exists an event Ω_ν of probability 1 on which (1.17)–(1.18) hold. In particular, on the event Ω_ν, for each $f \in \mathbf{K}$ and $t \in T_f$, the function f belongs to the limit set $C_t(\nu)$ of the increment functions $\{Y_{t,h} : h > 0\}$ as $h \downarrow 0$. In other words,

$$\mathbf{P}\Big(\liminf_{h \downarrow 0} \nu(Y_{t,h} - f) = 0 \ \forall t \in T_f, \ \forall f \in \mathbf{K} \Big) = 1 \quad \forall \nu \in \mathcal{C}. \tag{1.19}$$

The arguments given in the sequel show that, even though T'_α, T''_α and T_f may be defined independently of $\nu \in \mathcal{C}$, the event Ω_ν of probability 1 on which (1.17)–(1.18) hold *depends upon* $\nu \in \mathcal{C}$. The existence of an event of probability 1 implying (1.17)–(1.18) independently of $\nu \in \mathcal{C}$ is unlikely even though we have not been able to disprove its existence.

The proofs of Theorems 1.2 and 1.3 are given in Sections 2 and 4.

To motivate the remaining third part of our paper, it is useful to outline some of the ideas which will be used to prove Theorem 1.3. We will make an instrumental use of the observation that the set $\mathcal{C}$ of consistent norms may be equipped with a *separable metric topology*. This will allow us to infer Theorem 1.3 from an application of Theorem 1.2 to a properly chosen countable dense subset $\mathcal{N}$ of $\mathcal{C}$. The large deviation bounds for distances of norms within $\mathcal{C}$ which are needed to complete this part of our proof have interest in and of themselves and will be established in the forthcoming Section 3.

2. Proof of Theorem 1.2

2.1. Introduction. To prove Theorem 1.2, it is enough to check the validity of the upper bound in (1.11), and the lower bound in (1.12). In other words, we need only show that, under the assumptions of the theorem, with probability 1,

$$\dim\Big(\bigcup_{\nu \in \mathcal{N}} T(\nu, \alpha) \Big) \leqq 1 - \alpha^2 \quad \forall \alpha \in [0,1], \tag{2.1}$$

and

$$\dim\Big(\bigcap_{\nu\in\mathcal{N}} T(\nu,f)\Big) \geqq 1-|f|_{\mathbf{H}}^2 \quad \forall f\in\mathbf{K}. \tag{2.2}$$

By combining (2.1)–(2.2) with the implications, holding for $f\in\mathbf{H}$ and $\alpha\geqq 0$,

$$|f|_{\mathbf{H}}>\alpha \Rightarrow T(\nu,f)\subseteqq T(\nu,\alpha), \tag{2.3}$$

it is readily seen that (2.1)–(2.2) hold as equalities, which establishes (1.11)–(1.12).

We postpone the verification of (2.1) and (2.2) when $\mathcal{N}=\{\nu\}$ consists of a single norm $\nu\in\mathcal{C}$ to Sections 2.2 and 2.3. Below, we limit ourselves to show how the result for $\mathcal{N}=\{\nu\}$ can be extended to the case where $\mathcal{N}\subseteqq\mathcal{C}$ is an arbitrary countable set of consistent norms.

Assume therefore from now on that, for each specified $\nu\in\mathcal{C}$, (2.1)–(2.2) hold with probability 1 for $\mathcal{N}=\{\nu\}$. Consider any countable set $\mathcal{N}=$ $=\{\nu_i : i\geqq 1\}\subseteqq\mathcal{C}$ of consistent norms. It follows from (A) in combination with the scaling property of Wiener processes that a norm ν satisfying (1.3) for *some* $\varepsilon>0$, also satisfies (1.3) for *all* $\varepsilon\in(0,1)$. In addition, the integrability properties of Gaussian seminorms (see (2.24) and Remark 2.1 in the sequel) imply that for each $\varepsilon\in(0,1)$ and each consistent norm $\nu\in\mathcal{C}$

$$0<\mathbf{E}\Big(\sup_{\theta_1,\theta_2\in[0,\varepsilon]} \nu(W(\theta_1+(1-\theta_2)\cdot)-W(\theta_1))\Big)<\infty. \tag{2.4}$$

By (2.4), we may let, for $i=1,2,\ldots,$

$$b_i=2^i\,\mathbf{E}\Big(\sup_{\theta_1,\theta_2\in[0,1/2]} \nu_i(W(\theta_1+(1-\theta_2)\cdot)-W(\theta_1))\Big)$$

and define properly a new norm ν_0 on $C_0[0,1]$ by setting

$$\nu_0=\sum_{i\geqq 1}\nu_i/b_i.$$

In view of the easily verified fact that ν_0 is a consistent norm on $C_0[0,1]$, we conclude the proof of Theorem 1.2 by an application of (2.1)–(2.2)–(2.3) to $\mathcal{N}=\{\nu_0\}\subseteqq\mathcal{C}$, in combination with the straightforward inclusions of sets, for all $f\in\mathbf{H}$ and $\alpha\geqq 0$,

$$\bigcup_{i\geqq 1} T(\nu_i,\alpha)\subseteqq T(\nu_0,\alpha) \quad\text{and}\quad \bigcap_{i\geqq 1} T(\nu_i,f)\supseteqq T(\nu_0,f).$$

2.2. Upper bounds. In this subsection, we prove (2.1) when $\mathcal{N}=\{\nu\}\subseteq\mathcal{C}$ reduces to a single element. We first observe that (2.1) is trivial for $\alpha=0$ since the Hausdorff dimension of a subset of $[0,1]$ is always bounded above by 1. We may therefore limit ourselves to the case where $\alpha\in(0,1]$, and assume from now on that this condition holds. Select an arbitrary $\varepsilon>0$ together with a $\rho\in(1-\alpha^2,1)$. Below, we will show that the Hausdorff dimension of the set

$$T(\nu,\alpha,\varepsilon):=\Big\{t\in[0,1]:\limsup_{h\downarrow 0}\Big(\inf_{f\in\alpha\mathbf{K}}\nu(Y_{t,h}-f)\Big)>\varepsilon\Big\}, \tag{2.5}$$

is bounded above by ρ. This will be achieved by the construction of an "economic" covering of $T(\nu,\alpha,\varepsilon)$, as follows. Throughout the sequel, we will set $\gamma_n=e^{-\sqrt{n}}$ and $t_{j,n}=jn^{-1}\gamma_n$, for $j\in\mathbf{N}=\{0,1,\ldots\}$ and $n\geqq 1$. Moreover we set, for $n\geqq 1$,

$$J_n=\{j\in\mathbf{N}:0\leqq t_{j,n}\leqq 1,\ \inf_{f\in\alpha\mathbf{K}}\nu(Y_{t_{j,n},\gamma_n}-f)\geqq\varepsilon/2\}.$$

We will show that, with probability 1 for all large n_0, the inclusion

$$T(\nu,\alpha,\varepsilon)\subseteq\bigcup_{n\geqq n_0}\Big\{\bigcup_{j\in J_n}[t_{j,n},t_{j+1,n}]\Big\} \tag{2.6}$$

holds. In addition, setting $|A|$ for the Lebesgue measure of $A\subseteq\mathbf{R}$ and $\#J$ for the cardinality of J, we will show that, with probability 1,

$$\sum_{n=1}^{\infty}\Big|\bigcup_{j\in J_n}[t_j,t_{j+1}]\Big|=\sum_{n=1}^{\infty}(\#J_n)(n^{-1}\gamma_n)^{\rho}<\infty. \tag{2.7}$$

Let us first assume that the claims (2.6)–(2.7) hold with probability 1 for each choice of $\varepsilon>0$ and $\rho\in(1-\alpha^2,1)$. The definition (1.9) of Hausdorff dimension, implies in this case that, with probability 1, for each choice of $\varepsilon>0$ and $\rho\in(1-\alpha^2,1)$,

$$\dim(T(\nu,\alpha,\varepsilon))\leqq\rho. \tag{2.8}$$

We note from the definitions (1.7) and (2.5) that

$$T(\nu,\alpha)=\bigcup_{n\geqq 1}T(\nu,\alpha,1/n).$$

By applying (2.8) to $\varepsilon=1/n$ for $n=1,2,\ldots$, we obtain readily that $\dim(T(\nu,\alpha,1/n))\leqq\rho$ with probability 1, for each $\rho\in(1-\alpha^2,1)$ and all $n\geqq 1$. The σ-stability of Hausdorff dimension (see e.g. (2.20), p. 17 in [20]) implies in turn that, with probability 1, for each $\rho\in(1-\alpha^2,1)$,

$$\dim(T(\nu,\alpha))=\sup_{n\geqq 1}\Big\{\dim(T(\nu,\alpha,1/n))\Big\}\leqq\rho.$$

By choosing in this last inequality $\rho = \rho_m$, where $\rho_m \in (1-\alpha^2, 1)$, $m = 1, 2, \ldots$ is any sequence such that $\rho_m \downarrow 1-\alpha^2$ as $m \to \infty$, we obtain readily that, with probability 1, for each $\alpha \in [0,1]$

$$\dim(T(\nu, \alpha)) \leqq 1 - \alpha^2. \tag{2.9}$$

This in turn implies that (2.9) holds with probability 1 for all $\alpha \in A$, where A is a countable dense subset of $[0,1]$. Since the function $\alpha \in [0,1] \to \dim(T(\nu, \alpha))$ is obviously nonincreasing, an easy argument shows that (2.9) holds with probability 1 for all $\alpha \in [0,1]$.

By all this, the assertion (2.1) when $\mathcal{N} = \{\nu\}$ consists of a single element is a consequence of (2.6)–(2.7). The following arguments are oriented towards proving the latter two assertions.

We will make use of the *isoperimetric inequality* of Borell, Sudakov and Tsyrelson, which we cite for convenience below in the general framework of Section 1 (see e.g. Section 11 in [15]). We will apply this inequality in the special case of $Z = W$, $\mathcal{X} = (C_0[0,1], \mathcal{U})$, $\mathbf{H}_Z = \mathbf{H}$ and $\mathbf{K}_Z = \mathbf{K}$. Denote by $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} dt$ the standard normal distribution function, and define $\Phi^{-1}(s)$ for $s \in (0,1)$ by the equality $\Phi(\Phi^{-1}(s)) = s$ for all $s \in (0,1)$.

FACT 1. *Let $\mathcal{X}$ be a Hausdorff locally convex space with Borel σ-algebra $\mathcal{B}_{\mathcal{X}}$. Let Z denote an $\mathcal{X}$-valued random vector with distribution given by a centered Gaussian Radon measure $\mathbf{P}_Z$ on $\mathcal{X}$. Denote by $\mathbf{H}_Z$ the RKHS of Z and by $\mathbf{K}_Z$ its unit ball. Then for any $r \geqq 0$, $A \in \mathcal{B}_{\mathcal{X}}$ and $B \in \mathcal{B}_{\mathcal{X}}$ with $B \cap (A + r\mathbf{K}_Z) = \emptyset$, we have*

$$\mathbf{P}(Z \in B) \leqq 1 - \Phi\Big(\Phi^{-1}(\mathbf{P}(Z \in A)) + r\Big). \tag{2.10}$$

PROOF. See Borell [1], Sudakov and Tsyrelson [22], and e.g. Section 11 in [15]. □

Since $\mathbf{E}(|X|) < \infty \Rightarrow \mathbf{P}(|X| < \infty) = 1$, to establish (2.7), we need only show that

$$\sum_{n=1}^{\infty} \mathbf{E}(\#J_n)(n^{-1}\gamma_n)^\rho < \infty. \tag{2.11}$$

Recall the definitions of $X_{t,h}$, $Y_{t,h}$ and $L(h) = (2\log(1/h))^{1/2}$. By the scaling property of the Wiener process in combination with (A), we obtain readily that, for all large n,

$$\begin{aligned} \mathbf{E}(\#J_n) &\leqq \lfloor n\gamma_n^{-1} + 1 \rfloor \ \mathbf{P}\Big(\inf_{f \in \alpha\mathbf{K}} \nu(Y_{0,\gamma_n} - f) \geqq \varepsilon/2\Big) \\ &\leqq 2n\gamma_n^{-1} \ \mathbf{P}\Big(\inf_{h \in \alpha L(\gamma_n)\mathbf{K}} \nu(X_{0,\gamma_n} - h) \geqq \varepsilon L(\gamma_n)/2\Big) \\ &= 2n\gamma_n^{-1} \ \mathbf{P}\Big(\inf_{h \in \alpha L(\gamma_n)\mathbf{K}} \nu(W - h) \geqq \varepsilon L(\gamma_n)/2\Big), \end{aligned} \tag{2.12}$$

where, here and elsewhere, $\lfloor u \rfloor \leqq u < \lfloor u \rfloor + 1$ denotes integer part of u. We apply Fact 1 with $Z = W$, $\mathcal{X} = (C_0[0,1], \mathcal{U})$, $\mathbf{H}_Z = \mathbf{H}$, $\mathbf{K}_Z = \mathbf{K}$. Letting $D_M = \{f \in C_0[0,1] : \nu(h) < M\}$, we observe that (1.4) implies the existence of a large $M > 0$ such that

$$\mathbf{P}(W \in D_M) = \mathbf{P}(\nu(W) < M) > \Phi(1). \tag{2.13}$$

We apply (2.10) with $A = D_M$, $B = C_0[0,1] - \{D_M + \alpha L(\gamma_n)\mathbf{K}\}$ and $r = \alpha L(\gamma_n)$. Since $L(\gamma_n) = (\log n)^{1/2} \to \infty$, there exists an $n_0 < \infty$ such that $\varepsilon L(\gamma_n)/2 > M$ for all $n \geqq n_0$. We obtain therefore that, for all $n \geqq n_0$,

$$\begin{aligned} \mathbf{P}\Big(\inf_{h \in \alpha L(\gamma_n)\mathbf{K}} \nu(W-h) \geqq \varepsilon L(\gamma_n)/2\Big) &= \mathbf{P}\Big(W \notin D_{\varepsilon L(\gamma_n)/2} + \alpha L(\gamma_n)\mathbf{K}\Big) \\ &\leqq \mathbf{P}\Big(W \notin D_M + \alpha L(\gamma_n)\mathbf{K}\Big) = \mathbf{P}\Big(W \in B\Big) \\ &\leqq 1 - \Phi\Big(\Phi^{-1}(\mathbf{P}(W \in A)) + r\Big) \\ &= 1 - \Phi\Big(\Phi^{-1}(\mathbf{P}(W \in D_M)) + \alpha L(\gamma_n)\Big) \\ &\leqq 1 - \Phi\Big(1 + \alpha L(\gamma_n)\Big). \end{aligned} \tag{2.14}$$

Since $1 - \Phi(x) \leqq x^{-1}(2\pi)^{-1/2} e^{-x^2/2} \leqq e^{-x^2/2}$ for $x \geqq 1$, it follows readily from (2.14) that for all large n,

$$\mathbf{P}\Big(\inf_{h \in \alpha L(\gamma_n)\mathbf{K}} \nu(W-h) \geqq \varepsilon L(\gamma_n)/2\Big) \leqq \exp\Big(-\alpha^2 L(\gamma_n)^2/2\Big) = \gamma_n^{\alpha^2}. \tag{2.15}$$

We infer from (2.12) and (2.15) that, for all n sufficiently large,

$$\begin{aligned} \mathbf{E}(\#J_n)(n^{-1}\gamma_n)^\rho &\leqq (2n\gamma_n^{-1}) \times \gamma_n^{\alpha^2} \times (n^{-1}\gamma_n)^\rho \\ &= 2n^{1-\rho} \exp\Big(-(\alpha^2 + \rho - 1)\sqrt{n}\Big), \end{aligned}$$

which in turn, given that $\alpha^2 + \rho - 1 > 0$, readily implies (2.11), and hence (2.7).

The following Lemmas 2.1–2.2 are oriented towards proving (2.6).

LEMMA 2.1. *For each consistent norm $\nu \in \mathcal{C}$ and each $\varepsilon > 0$ we have*

$$\sum_{n=1}^{\infty} \Big\{ \sum_{j:\, 0 \leqq t_{j,n} \leqq 1} \mathbf{P}\Big(\sup_{t \in [t_{j,n}, t_{j+1,n}]} \sup_{h \in [\gamma_{n+1}, \gamma_n]} \nu(Y_{t,h} - Y_{t_{j,n},\gamma_n}) \geqq \varepsilon/2 \Big) \Big\} < \infty. \tag{2.16}$$

Before giving the details of the proof of Lemma 2.1, we will show that its conclusion (2.16) implies (2.6). Assume therefore that (2.16) holds. By combining this claim with the Borel–Cantelli lemma, we obtain readily that the property $(\mathcal{P})$ below is satisfied with probability 1 for all large n. Recall that $t_{j,n} = jn^{-1}\gamma_n$ for $j \geqq 0$.

$(\mathcal{P})$ For any $t \in [0,1]$, the choice of $j = j(t) = \lfloor tn\gamma_n^{-1} \rfloor$ ensures that $\nu(Y_{t,h} - Y_{t_{j,n},\gamma_n}) \leqq \varepsilon/2$ for all $h \in [\gamma_{n+1}, \gamma_n]$, so that

$$\begin{aligned} t \in T(\nu,\alpha,\varepsilon) &\Leftrightarrow \Big\{ \inf_{f\in\alpha\mathbf{K}} \nu(Y_{t,h} - f) \geqq \varepsilon \ \text{ for some } h \in [\gamma_{n+1}, \gamma_n] \Big\} \\ &\Rightarrow \Big\{ \inf_{f\in\alpha\mathbf{K}} \nu(Y_{t_{j,n},\gamma_n} - f) \geqq \varepsilon/2 \Big\} \Rightarrow j \in J_n \\ &\Leftrightarrow t \in \bigcup_{n\geqq n_0} \Big\{ \bigcup_{j\in J_n} [t_{j,n}, t_{j+1,n}] \Big\}. \end{aligned}$$

Since $(\mathcal{P})$ holds with probability 1 *uniformly over* $t \in [0,1]$, we obtain (2.6) as sought.

The following Lemma 2.2 gives a version of the isoperimetric inequality which will be instrumental in the forthcoming proof of Lemma 2.1.

LEMMA 2.2. *Let Z denote a random vector with distribution given by a centered Gaussian Radon measure on a Hausdorff locally convex space $\mathcal{X}$. Let $\|\cdot\|$ be a seminorm on $\mathcal{X}$, measurable with respect to $\mathbf{P}_Z$ and such that $\mathbf{P}(\|Z\| < \infty) = 1$. Let m be a median of the distribution of $\|Z\|$. For any $c > m$ set $\beta = \beta(\|\cdot\|, c, Z) = c/\Phi^{-1}(\mathbf{P}(\|Z\| < c))$. Then for any $R \geqq m$ the following inequality holds*

$$\mathbf{P}(\|Z\| \geqq R) \leqq 1 - \Phi((R-m)/\beta). \tag{2.17}$$

PROOF. See e.g. Lemma 3.2 in [6], or Lemma 2.1 in [5]. □

REMARK 2.1. It is easy to check from (2.17) that, for any measurable norm $\|\cdot\|$ on $\mathcal{X}$, the condition $\mathbf{P}(\|Z\| < \infty) = 1$ is equivalent to $\mathbf{E}(\|Z\|^m) < \infty$ for an arbitrary $m > 0$.

PROOF OF LEMMA 2.1. Recall that $t_{j,n} = jn^{-1}\gamma_n$ and $\gamma_n = e^{-\sqrt{n}}$ for $j \geqq 0$ and $n \geqq 1$. The scaling property of the Wiener process in combination with (A) show that, for all n sufficiently large, the left-hand-side of (2.13) is bounded above by

$$\begin{aligned} &\lfloor n\gamma_n^{-1} + 1 \rfloor \, \mathbf{P}\Big(\sup_{t\in[0,n^{-1}\gamma_n]} \ \sup_{h\in[\gamma_{n+1},\gamma_n]} \nu(Y_{t,h} - Y_{0,\gamma_n}) \geqq \varepsilon/2 \Big) \\ &\leqq 2n\gamma_n^{-1} \, \mathbf{P}\Big(\sup_{t\in[0,n^{-1}\gamma_n]} \ \sup_{h\in[\gamma_{n+1},\gamma_n]} \nu(Y_{t,h} - Y_{0,\gamma_n}) \geqq \varepsilon/2 \Big). \end{aligned} \tag{2.18}$$

Set $X = X_{0,g}$, $g = \gamma_n$ and $\Gamma = \gamma_{n+1}/\gamma_n$. Recalling that $\{X(u) = g^{-1/2}W(gu) : u \geqq 0\}$ is a Wiener process, we have the following key identity for $h \in [\Gamma g, g]$, $t \in [0, bg]$ and $u \geqq 0$

(2.19)
$$\begin{aligned} X_{t,h}(u) =& h^{-1/2}(W(t+hu) - W(t)) = (g/h)^{1/2} g^{-1/2}(W(t+hu) - W(t)) \\ =& (g/h)^{1/2} g^{-1/2} W(gu) + (g/h)^{1/2} g^{-1/2}\{W(t+hu) + W(t) - W(gu)\} \\ =& (g/h)^{1/2} X(u) + (g/h)^{1/2}\{X(\theta_1 + (1-\theta_2)u) - X(\theta_1) - X(u)\}, \end{aligned}$$

where $\theta_1 := t/g \leqq b$ and $\theta_2 := 1 - h/g \in [0, 1-\Gamma]$. Therefore, by (2.19),

$$\begin{aligned} Y_{t,h} - Y_{0,g} =& L_h^{-1} X_{t,h} - L_g^{-1} X_{0,g} = \{(g/h)^{1/2} L_h^{-1} - L_g^{-1}\} X(\cdot) \\ &+ (g/h)^{1/2} L_h^{-1}\{X(\theta_1 + (1-\theta_2)\cdot) - X(\theta_1) - X(\cdot)\}. \end{aligned} \tag{2.20}$$

Next, uniformly over $\Gamma \leqq h/g \leqq 1$, we have ultimately in $\Gamma \uparrow 1$ and $g \downarrow 0$,

$$0 \leqq (g/h)^{1/2} L_h^{-1} - L_g^{-1} \leqq 3(1-\Gamma) L_g^{-1} \quad \text{and}$$
$$(g/h)^{1/2} L_h^{-1} \leqq \Gamma^{-1/2} L_g^{-1} \leqq 2 L_g^{-1}.$$

The combination of these estimates with (2.20) yields

$$\begin{aligned} \mathbf{P}\Big(\sup_{t\in[0,bg]} \sup_{h\in[\Gamma g,g]} \nu(Y_{t,h} - Y_{0,g}) \geqq \varepsilon/2\Big) &\leqq \mathbf{P}_1(g,\Gamma) + \mathbf{P}_2(b,g,\Gamma) \\ &:= \mathbf{P}\Big(\nu(X) \geqq L_g\varepsilon/\{12(1-\Gamma)\}\Big) \\ + \mathbf{P}\Big(\sup_{\theta_1\in[0,b]} \sup_{\theta_2\in[0,1-\Gamma]} \nu(X(\theta_1 + &(1-\theta_2)\cdot) - X(\theta_1) - X(\cdot)) \geqq L_g\varepsilon/8\Big). \end{aligned} \tag{2.21}$$

Since for the choices of $g = \gamma_n$ and $\Gamma = \gamma_{n+1}/\gamma_n$ we consider, we have $\Gamma \uparrow 1$ and $g \downarrow 0$, we note that (2.21) holds uniformly over $\Gamma \leqq h/g \leqq 1$ for large n. Below, we derive upper bounds for $\mathbf{P}_1(g,\Gamma)$ and $\mathbf{P}_2(b,g,\Gamma)$, making use of Lemma 2.2.

Recalling that X and W are identically distributed, we first apply Lemma 2.2 to $Z = W$, $\|\cdot\| = \nu(\cdot)$, $R = L_g\varepsilon/\{12(1-\Gamma)\}$ and $c = c_1$, chosen, via (2.13), in such a way that $\mathbf{P}(\nu(W) < c_1) > \Phi(1)$. Since then, $\beta = \beta(\nu, c_1, W) = c_1/\Phi^{-1}(\mathbf{P}(\nu(W) < c_1)) \leqq c_1$, the observation that, for all large n, $(R-m)/\beta > \frac{12}{13}R/\beta \geqq \frac{12}{13}R/c_1$ allows us to write, via (2.17), the following inequalities. For all large n, with $g = \gamma_n$ and $\Gamma = \gamma_{n+1}/\gamma_n$,

$$\begin{aligned} \mathbf{P}_1(g,\Gamma) = \mathbf{P}\Big(\nu(W) \geqq L_g\varepsilon/\{12(1-\Gamma)\}\Big) &\leqq 1 - \Phi\Big(L_g\varepsilon/\{13c_1(1-\Gamma)\}\Big) \\ \leqq \exp\Big(-\frac{L_g^2\varepsilon^2}{2\{13c_1(1-\Gamma)\}^2}\Big) &= \gamma_n^{q_n} \leqq \gamma_n^{(9/8)^2}, \end{aligned} \tag{2.22}$$

where $q_n := \varepsilon^2/\{13c_1(1-\Gamma)\}^2$. Here, the fact that $q_n \to \infty$ (whence $q_n \geqq (9/8)^2$ ultimately), follows from the observation that $\Gamma = \gamma_{n+1}/\gamma_n = e^{\sqrt{n}-\sqrt{n+1}} \to 1$ as $n \to \infty$.

Second, we apply Lemma 2.2 to $Z = W$ and $\|\cdot\| = \|\cdot\|_\theta$ given for $f \in C_0[0,1]$ by

$$\|f\|_\theta = \sup_{\theta_1,\theta_2\in[0,\theta]} \nu(f(\theta_1+(1-\theta_2)\cdot) - f(\theta_1) - f(\cdot)). \tag{2.23}$$

Our assumption that ν satisfies (A) and (B) readily implies that $\|\cdot\|_\theta$ defines on $C_0[0,1]$ a lower semi-continuous norm with respect to $\mathcal{U}$. The fact that $\mathbf{P}(\|W\|_\theta < \infty) = 1$ (which is needed in Lemma 2.2) follows from the last consistency condition (C) via (1.3). This, together with an application of the zero-one law (see Cameron and Graves [3], Kallianpur [12], Jain [11]) following the lines of proof of Lemma 2.4 in [6], implies the existence of a constant $c_2 < \infty$ such that, with probability 1,

$$\lim_{\theta\downarrow 0} \|W\|_\theta = c_2. \tag{2.24}$$

Since (2.24) implies that $\mathbf{P}(\|W\|_\theta \leqq c_2+1) \to 1$ as $\theta \downarrow 0$ and $\Phi(10\varepsilon^{-1}(c_2+1)) \to 1$ as $\varepsilon \downarrow 0$, there exists an $\varepsilon_0 > 0$ such that, for each $0 < \varepsilon \leqq \varepsilon_0$, a choice of $\theta_0 = \theta_0(\varepsilon) > 0$ sufficiently small guarantees that, for all $0 < \theta \leqq \theta_0 = \theta_0(\varepsilon)$,

$$\mathbf{P}\Big(\|W\|_\theta \leqq c_2+1\Big) \geqq \Phi(10\varepsilon^{-1}(c_2+1)) > 1/2. \tag{2.25}$$

By (2.25), any median $m(\theta)$ of the distribution of $\|W\|_\theta$ is such that $c_2+1 > m(\theta)$. Moreover, with the notation of Lemma 2.2, uniformly over $0 < \theta \leqq \theta_0$,

$$\beta(\|\cdot\|_\theta, c_2+1, W) = \{c_2+1\}/\Phi^{-1}(\mathbf{P}(\|W\|_\theta < c_2+1) \leqq \varepsilon/10. \tag{2.26}$$

By combining (2.24)–(2.25) with (2.17), taken with $\beta = \beta(\|\cdot\|_{\theta_0}, c_2+1, W)$, $\|\cdot\| = \|\cdot\|_{\theta_0}$ and $m = m(\theta_0)$, we see that there exists an $R_0 \geqq c_2+1$ such that, for all $R \geqq R_0$,

$$\begin{aligned}\mathbf{P}\Big(\|W\|_{\theta_0} \geqq R\Big) &\leqq 1-\Phi((R-c_2-1)/\beta) \leqq 1-\Phi(10(R-c_2-1)/\varepsilon)\\ &\leqq 1-\Phi(9R/\varepsilon) \leqq \exp\big(-9^2R^2/2\varepsilon^2\big).\end{aligned} \tag{2.27}$$

We now choose $R = L_g\varepsilon/8$ in (2.27) with $g = \gamma_n$, and assume n to be so large that $R \geqq R_0$, $n^{-1} \leqq \theta_0$ and $1-\Gamma = 1-\gamma_{n+1}/\gamma_n \leqq \theta_0$. By combining (2.21) with the definition (2.23) of $\|\cdot\|_\theta$, we infer from (2.27) that, for all large n, with $b = n^{-1}$, $g = \gamma_n$ and $\Gamma = \gamma_{n+1}/\gamma_n$,

$$\begin{aligned}\mathbf{P}_2(b,g,\Gamma) &= \mathbf{P}\Big(\sup_{\theta_1\in[0,b]}\ \sup_{\theta_2\in[0,1-\Gamma]} \nu(X(\theta_1+(1-\theta_2)\cdot) - X(\theta_1) - X(\cdot)) \geqq L_g\varepsilon/8\Big)\\ &\leqq \mathbf{P}\Big(\|X\|_{\theta_0} \geqq R\Big) = \mathbf{P}\Big(\|W\|_{\theta_0} \geqq R\Big)\\ &\leqq \exp\Big(-9^2L_g^2/\{2\times 8^2\}\Big) = \gamma_n^{(9/8)^2}.\end{aligned} \tag{2.28}$$

By combining (2.21) with (2.22) and (2.28) we see that the left-hand side of (2.18) is ultimately bounded above by

$$(2n\gamma_n^{-1}) \times (2\gamma_n^{(9/8)^2}) = 4n \times \exp\Big(-\frac{17}{81}\sqrt{n}\Big),$$

which is summable in n. The convergence of the series (2.16) is therefore established. □

2.3. Lower bounds. In this section, we establish the validity of (2.2) when $\mathcal{N} = \{\nu\} \subseteq \mathcal{C}$ consists of a single element. We start by proving that a version of (2.2) holds for a fixed $f \in \mathbf{K}$. Namely, we claim that, for each $\nu \in \mathcal{C}$ and $f \in \mathbf{K}$, we have with probability 1

$$\dim(T(\nu, f)) \geqq 1 - |f|_{\mathbf{H}}^2. \tag{2.29}$$

We postpone until the end of the section the proof that (2.29) holds with probability 1 uniformly over $f \in \mathbf{K}$.

Since (2.29) is trivial when $|f|_{\mathbf{H}} = 1$, we may limit ourselves without loss of generality to prove this claim in the case where $|f|_{\mathbf{H}} < 1$. We will give below the arguments needed when $0 < |f|_{\mathbf{H}} < 1$ and assume from now on that this assumption holds. The case where $f = 0$ is obtained by routine modifications which we omit for the sake of conciseness. We will obtain a lower bound for the Hausdorff dimension of $T(\nu, f)$ by following the ideas of Orey and Taylor [17] (see also [8]). The next two facts will be instrumental for our needs.

FACT 2. *Let $T \subseteq [0,1]$ be such that $T = \bigcap_{n=n_0}^{\infty} T_n$, where $T_{n_0} \supseteq \dots \supseteq T_n \supseteq \dots$ for $n \geqq n_0$, and $T_n = \bigcup_{k=1}^{M_n} I_{n,k}$ with $\{I_{n,k} : 1 \leqq k \leqq M_n\}$ being for each $n \geqq n_0$ a collection of disjoint closed intervals. Let $\rho > 0$ be a constant, and assume that there exist two constants $c > 0$ and $\delta > 0$ such that the following property holds. For each interval $J \subseteq [0,1]$ with $|J| \leqq \delta$ there exists a finite integer $n(J)$ such that for all $n \geqq n(J)$*

$$M_n(J) := \#\{I_{n,k} \subseteq J : 1 \leqq k \leqq M_n\} \leqq c|J|^{\rho} M_n\,. \tag{2.30}$$

Then $\dim(T) \geqq \rho$.

PROOF. See e.g. Lemma 2.2 in [17] and Lemma 3.5 in [8]. □

FACT 3. *Let A be a symmetric Borel subset of $(C_0[0,1], \mathcal{U})$. Then*

$$\mathbf{P}(W - h \in A) \geqq \mathbf{P}(W \in A) \exp\Big(-\frac{1}{2}|h|_{\mathbf{H}}^2\Big). \tag{2.31}$$

PROOF. See e.g. Lemma 2.2 in [5]. □

The following arguments aim to construct appropriate Cantor-type sets $T = \hat{T}(f) \subseteq T(\nu, f)$ as in Fact 2, satisfying (2.30) for suitable choices of ρ.

We start by choosing an $R>0$ so large that $\mathbf{P}(\nu(W)\leqq R)>1/2$. This is rendered possible by (1.4) and the assumption that $\nu\in\mathcal{C}$ is consistent. For each $b>0$ and $q>0$ with $q^{-1}\in\mathbf{N}$ and $b^{-1}\in\mathbf{N}$, we define, for $g\in\mathbf{K}$, the families of intervals

$$\begin{aligned}\mathcal{J}_q(g)&=\Big\{[iq,(i+1)q]\subseteqq[0,1]:\nu(Y_{iq,q}-g)\leqq R/L_q\,\text{and}\,i\in\mathbf{N}\Big\},\\ \mathcal{J}_q^b(g)&=\Big\{[iq,(i+b)q]:[iq,(i+1)q]\in\mathcal{J}_q(g)\Big\}.\end{aligned}\tag{2.32}$$

In (2.32), we will set $b=b_n=1/n$ and $q=q_n$ for $n\geqq1$, where $\{q_n:n\geqq1\}$ is a rapidly decreasing sequence of positive constants which will be precised later on. We assume that this sequence is such that, for each $n\geqq1$, $q_n^{-1}\in\mathbf{N}$ and $b_nq_n/q_{n+1}=q_n/\{nq_{n+1}\}\in\mathbb{N}$. We construct $\hat{T}(f)$ by induction as follows.

First, we introduce a sequence $\{f_n:n\geqq1\}\subseteqq\mathbf{K}$ such that

(F.1) $\lim_{n\to\infty}\nu(f_n-f)=0$.

The following parameters are needed for the statement of the assumptions (F.2)–(F.4) and (C.2)–(C.5) below. We select a sequence $\{\varepsilon_n:n\geqq1\}$ of constants such that $0<\varepsilon_n\leqq1$ for $n\geqq1$ and $\sum_{n=1}^{\infty}\varepsilon_n<\log2$. Letting $\mathcal{E}=\exp\big(\sum_{n=1}^{\infty}\varepsilon_n\big)$, we observe that $1<\mathcal{E}<2$. Moreover, we select a $\triangle\in(0,\frac18\min\{|f|_{\mathbf{H}}^2,1-|f|_{\mathbf{H}}^2\})$, and set $\rho=1-|f|_{\mathbf{H}}^2-2\triangle$. We note for further use that

$$0<6\triangle<\rho<1-|f|_{\mathbf{H}}^2-\triangle<1-|f|_{\mathbf{H}}^2<1-8\triangle.$$

Second, we consider a sequence $\{\mathcal{F}_n:n\geqq1\}\subseteqq\mathbf{K}$ of finite subsets of $\mathbf{K}$ fulfilling with $\{f_n:n\geqq1\}$ the following properties.

(F.2) For each $n\geqq1$, $\Theta_n:=\#\mathcal{F}_n<\infty$;

(F.3) For each $n\geqq1$, and each $g\in\mathcal{F}_n$, $2\triangle<|g|_{\mathbf{H}}<1-2\triangle$.

(F.4) For each $n\geqq1$, $f_n\in\mathcal{F}_n$ and $|f_n|_{\mathbf{H}}\leqq|f|_{\mathbf{H}}$.

We will give later on some additional conditions which will be imposed upon $\{f_n:n\geqq1\}$ and $\{\mathcal{F}_n:n\geqq1\}$. Third, we select a large $n_0\in\mathbf{N}$ and set

$$\hat{T}_{n_0}(f)=\bigcup\{I:I\in\mathcal{T}_{n_0}(f)\},\quad\mathcal{T}_{n_0}(f)=\mathcal{J}_{q_{n_0}}^{b_{n_0}}(f_{n_0}).\tag{2.33}$$

Fourth, for each $n\geqq n_0$ we define $\hat{T}_{n+1}(f)$ and $\mathcal{T}_{n+1}(f)$ out of $\hat{T}_n(f)$ by setting

$$\begin{aligned}\hat{T}_{n+1}(f)&=\bigcup\{I:I\in\mathcal{T}_{n+1}(f)\},\\ \mathcal{T}_{n+1}(f)&=\{I\in\mathcal{J}_{q_{n+1}}^{b_{n+1}}(f_{n+1}),\ I\subseteqq\hat{T}_n(f)\}.\end{aligned}\tag{2.34}$$

Finally, we put

$$\hat{T}(f)=\bigcap_{n=n_0}^{\infty}\hat{T}_n(f).\tag{2.35}$$

The induction (2.33)–(2.35) may end if, for some $n \geqq n_0$, we have $\mathcal{T}_n(f) = \emptyset$. If such is the case, we will set $\hat{T}_m(f) = \emptyset$ for $m \geqq n$ and $\hat{T}(f) = \emptyset$. Below, we will list a series of conditions which will exclude such a degenerate situation. In particular, we will need the following minimal assumption (C.1) which implies that the first step (2.33) of the induction may be achieved for all sufficiently large n_0.

(C.1) There exists an $n_1 < \infty$ such that

$$\#\mathcal{J}_{q_n}^{b_n}(g) \geqq 1 \quad \text{for all} \quad g \in \mathcal{F}_n \quad \text{with} \quad n \geqq n_1.$$

We assume from now on, unless otherwise specified, that (C.1) holds, and investigate in more detail the conditions we need impose upon $\{q_n : n \geqq n_0\}$ and $\{f_n : n \geqq 1\}$ to ensure that $\hat{T}(f)$ is properly defined by (2.35) and fulfills (2.30) for suitable ρ, c and δ.

We note for further use that the definitions (2.33)–(2.35) imply that, for each $n \geqq n_0$, $\hat{T}_n(f)$ is the union of a finite (eventually void) collection of closed intervals of length $b_n q_n$. If $\hat{T}_n(f) \neq \emptyset$ for all $n \geqq n_0$, then $\hat{T}(f)$ is the intersection of an imbedded sequence of non-void closed subsets of $[0,1]$, which entails that $\hat{T}(f) \neq \emptyset$. In order to check that $\hat{T}_n(f)$ fulfills a version of (2.30), we need evaluate, for an arbitrary interval $J \subseteqq [0,1]$, the number of component intervals $I \subseteqq \hat{T}_n(f)$ included in J, of lengths equal to $b_n q_n$. The following notation will be needed. Recalling (2.32), we set, for each $g \in \mathbf{K}$, each interval $J \subseteqq [0,1]$, each $q > 0$ with $1/q \in \mathbf{N}$ and each $n \geqq n_0$,

$$N_q(J,g) = \#\Big\{I \subseteqq J : I = [iq,(i+1)q] \in \mathcal{J}_q(g) \text{ for some } i \in \mathbf{N}\Big\},$$

$$M_n(J,f) = \#\Big\{I \subseteqq J : I = [iq_n,(i+b_n)q_n] \in \mathcal{T}_n(f) \text{ for some } i \in \mathbf{N}\Big\}, \tag{2.36}$$

$$N_q(f) = N_q([0,1],f), \quad M_n(f) = M_n([0,1],f).$$

By combining (2.34) with (2.36), we get the following recurrent formula. For each interval $J^* \subseteqq [0,1]$ of the form $J^* = [rq_n, sq_n]$ with $r \in \mathbf{N}$, $s \in \mathbf{N}$ and $n \geqq n_0$,

$$M_{n+1}(J^*,f) = \sum_{\substack{I \in \mathcal{T}_n(f) \\ I \subseteqq J^*}} N_{q_{n+1}}(I, f_{n+1}). \tag{2.37}$$

Recalling that $R > 0$ is chosen in such a way that $\mathbf{P}(\nu(W) \leqq R) \geqq 1/2$, an application of (2.31) to $A = \{\psi \in C_0[0,1] : \nu(\psi) \leqq R\}$ and $h = L_{q_n} g$ shows that, uniformly over $g \in \mathbf{K}$,

$$\begin{aligned} p_n(g) :&= \mathbf{P}\Big([0,q_n] \in \mathcal{J}_{q_n}(g)\Big) = \mathbf{P}\Big(\nu(Y_{q_n,0} - g) \leqq R/L_{q_n}\Big) \\ &= \mathbf{P}\Big(\nu(X_{q_n,0} - L_{q_n} g) \leqq R\Big) = \mathbf{P}\Big(\nu(W - L_{q_n} g) \leqq R\Big) \\ &\geqq \mathbf{P}(\nu(W) \leqq R) \exp\Big\{-\frac{1}{2}|g|_{\mathbf{H}}^2 L_{q_n}^2\Big\} \geqq \frac{1}{2} q_n^{|g|_{\mathbf{H}}^2}. \end{aligned} \tag{2.38}$$

We infer from (2.32), (2.36) and (2.38) that, for each $n \geqq n_0$ and each interval $I \subseteqq [0,1]$ of the form $I = [rq_{n+1}, sq_{n+1}]$ with $r, s \in \mathbf{N}$ and $r < s$, we have, uniformly over $g \in \mathbf{K}$,

$$E_{n+1}(|I|, g) := \mathbf{E} N_{q_{n+1}}(I, g) = \frac{|I|}{q_{n+1}} \times p_{n+1}(g) = (s-r) p_{n+1}(g). \tag{2.39}$$

We note for further use that (2.39) holds when $I \in \mathcal{T}_n(f)$ with $|I| = b_n q_n$.

It is convenient to add the following claims (C.2)–(C.5) to (C.1). In a first step, we will assume their validity to complete the proof of the lower bounds we seek. In a second step, we will show that they hold with probability 1. In a last step, we will give the proof that (2.29) holds with probability 1 uniformly over $f \in \mathbf{K}$.

(C.2) Let $\|\cdot\|_\theta$ and c_2 be as in (2.23) and (2.24). For $n \geqq 1$, let

$$u_n = \max\{b_n, \beta(\|\cdot\|_{b_n}, c_2 + 1, W)\}^{1/2}, \tag{2.40}$$

with $\beta(\|\cdot\|, c, W) = c/\Phi^{-1}(\mathbf{P}(\|W\| \leqq c))$ as in Lemma 2.2. There exists an $n_2 < \infty$, such that, uniformly over all $\{f_n : n \geqq 1\} \subseteqq \mathbf{K}$, for each $n \geqq n_2$, $I = [iq_n, (i+1)q_n] \in \mathcal{J}_{q_n}(f_n)$, and $t \in [iq_n, (i+b_n)q_n]$,

$$\nu(Y_{t,(i+1)q_n - t} - f_n) \leqq R/L_{q_n} + u_n. \tag{2.41}$$

(C.3) There exists an $n_3 < \infty$ such that, for each $n \geqq n_3$ and each interval $I \subseteqq [0,1]$ of the form $I = [rq_n, sq_n]$, with $r, s \in \mathbf{N}$ and $|I| = (s-r)q_n \geqq q_{n+1}^{\triangle}$, we have, uniformly over $g \in \mathcal{F}_{n+1}$,

$$\exp\{-\varepsilon_n\} \leqq \frac{N_{q_{n+1}}(I, g)}{\mathbf{E} N_{q_{n+1}}(I, g)} \leqq \exp\{\varepsilon_n\}. \tag{2.42}$$

(C.4) There exists an $n_4 < \infty$, together with a constant c_3 such that, for each $n \geqq n_4$ and each interval $I \subseteqq [0,1]$ of the form $I = [rq_{n+1}, sq_{n+1}]$, with $r, s \in \mathbf{N}$ and $|I| = (s-r)q_{n+1} \geqq q_{n+1}$, we have, uniformly over $g \in \mathcal{F}_{n+1}$,

$$N_{q_{n+1}}(I, g) \leqq c_3 |J|^{\{1 - |g|^2_{\mathbf{H}} - \triangle\}} \, \mathbf{E} N_{q_{n+1}}(g). \tag{2.43}$$

(C.5) Let u_n be as in (2.40). There exists an $n_5 < \infty$ such that, for all $n \geqq n_5$, we have

$$\begin{aligned}
&\text{(i)} \quad 0 < q_{n+1}^{\triangle^2} < b_n q_n \varepsilon_n^2\,; \\
&\text{(ii)} \quad q_{n+1}^{\triangle} \leqq \Big\{\prod_{m=1}^{n+1} b_m\Big\} \prod_{m=1}^{n} (q_m/2) = \frac{1}{(n+1)!} \prod_{m=1}^{n} (q_m/2); \\
&\text{(iii)} \quad q_{n+1} \leqq \exp(-1/u_{n+1}^4)\,; \\
&\text{(iv)} \quad \frac{\Theta_{n+1}}{q_{n+1}^2} \exp\Big\{-\frac{1}{8} q_{n+1}^{-\triangle}\Big\} \leqq \frac{1}{2^n}.
\end{aligned} \tag{2.44}$$

Recalling the relations $1-\rho-|f|_{\mathbf{H}}^2=2\triangle<\frac{1}{4}$ and $|f|_{\mathbf{H}}^2<1$, we observe that, whenever $0<q\leqq 1$ and $|g|_{\mathbf{H}}\leqq|f|_{\mathbf{H}}$,

$$q^{1-\rho-|f|_{\mathbf{H}}^2}=q^{2\triangle}\leqq q^{\triangle}\leqq q^{\triangle^2}\quad\text{and}\quad q\leqq q^{|f|_{\mathbf{H}}^2}\leqq q^{|g|_{\mathbf{H}}^2}.$$

In particular, (2.44)(ii) entails that, for all $m\geqq n_0+1>n_5$,

$$\frac{1}{q_m^{1-\rho-|f|_{\mathbf{H}}^2}}\Big\{\prod_{n=n_0}^{m}b_n\Big\}\Big\{\prod_{n=n_0}^{m-1}(q_n^{|f|_{\mathbf{H}}^2}/2)\Big\}\geqq 1.$$

STEP 1. We start by proving a version of (2.29) under (F.1)–(F.4) and (C.1)–(C.5). We set $n_0=\max\{n_j:1\leqq j\leqq 5\}$. Making use of (C.2) and (C.5), we infer from (2.32)–(2.35) that, for each $t\in\hat{T}(f)$ and $n\geqq n_0$, we have

$$t\in\hat{T}(f)\subseteqq\hat{T}_n(f)\Rightarrow\exists i_n:\, t\in[i_nq_n,(i_n+b_n)q_n]\subseteqq[i_nq_n,(i_n+1)q_n]\in\mathcal{J}_{q_n}(f_n).$$

Since $\varepsilon>0$ in (2.26) may be chosen arbitrarily small, we infer readily from this inequality and our choice of $b_n=1/n\to 0$ that, as $n\to\infty$,

$$\beta(\|\cdot\|_{b_n},c_2+1,W)\to 0. \tag{2.45}$$

By (2.45) and the definition (2.40) of u_n, we see that $u_n\to 0$. By (F.1), it follows that

$$\begin{aligned}\liminf_{h\downarrow 0}\ \nu(Y_{t,h}-f)&\leqq\lim_{n\to\infty}\ \nu(Y_{t,(i_n+1)q_n-t}-f)\\&\leqq\lim_{n\to\infty}\ \nu(Y_{t,(i_n+1)q_n-t}-f_n)+\lim_{n\to\infty}\ \nu(f_n-f)\\&=\lim_{n\to\infty}\ \nu(Y_{t,(i_n+1)q_n-t}-f_n)\leqq\lim_{n\to\infty}\ (R/L_{q_n}+u_n)=0.\end{aligned}$$

This shows that $t\in T(\nu,f)$. Since the just proven implication $t\in\hat{T}(f)\Rightarrow t\in T(\nu,f)$ holds uniformly over $t\in\hat{T}(f)$, we obtain therefore that

$$\hat{T}(f)\subseteqq T(\nu,f). \tag{2.46}$$

The following arguments are oriented to prove that (2.30) holds with $T=\hat{T}(f)$ and $T_n=\hat{T}_n(f)$. By Fact 2 and in view of the notation (2.36), we need only prove the existence of a $\delta>0$ and a $c>0$ such that, for all intervals $J\subseteqq[0,1]$ with $|J|\leqq\delta$,

$$M_n(J,f)\leqq c|J|^{\rho}M_n(f).$$

Consider an arbitrary interval J with $0<|J|\leqq q_{n_0+1}$. By the assumption, implied by (C.5), that $\{q_k:k\geqq n_0\}$ is decreasing, there exists a unique $k\geqq n_0$ such that $|J|\in[q_{k+1},q_k)$. Obviously, we have

$$M_m(J,f)=0\quad\text{for}\quad n_0\leqq m\leqq k, \tag{2.47}$$

so that we need only evaluate an upper bound for $M_m(J,f)$ when $m \geqq k+1$. We start with a simple observation. Let $J=[a,b]$. Recalling that $q_{k+1} \leqq |J| < q_k$, we set $J^* = [r_{k+1}q_{k+1}, s_{k+1}q_{k+1}]$, where $r_{k+1} = \max\{i \in \mathbf{N}: iq_{k+1} \leqq a\}$ and $s_{k+1} = \min\{i \in \mathbf{N} : iq_{k+1} \geqq b\}$. It is straightforward that $J \subseteqq J^*$, and $|J^*|/|J| \leqq 1 + 2q_{k+1}/|J| \leqq 3$. Moreover, our assumptions on $\{q_n : n \geqq 1\}$ entail that $J^* \subseteqq [0,1]$ is of the form $[rq_n, sq_n]$ for all $n \geqq k+1$. By (2.44)(i) in (C.5), for any $n \geqq k+1$ and $I \in \mathcal{T}_n(f)$, we have $|I| = b_n q_n > q_{n+1}^{\triangle^2} > q_{n+1}^{\triangle}$. This, when combined with (F.4), (2.37), (2.39) and (2.42), shows that, for all $n \geqq k+1$,

$$\begin{aligned} M_{n+1}(J,f) \leqq M_{n+1}(J^*,f) &= \sum_{\substack{I \in \mathcal{T}_n(f) \\ I \subseteqq J^*}} N_{q_{n+1}}(I, f_{n+1}) \\ &\leqq \#\{I \subseteqq J^* : I \in \mathcal{T}_n(f)\} \exp\{\varepsilon_n\} E_{n+1}(b_n q_n, f_{n+1}) \\ &= M_n(J^*, f) \exp\{\varepsilon_n\} \frac{b_n q_n}{q_{n+1}} p_{n+1}(f_{n+1}). \end{aligned} \tag{2.48}$$

Likewise, we obtain that

$$\begin{aligned} M_{n+1}(f) &= \sum_{I \in \mathcal{T}_n(f)} N_{q_{n+1}}(I, f_{n+1}) \\ &\geqq \#\{I \in \mathcal{T}_n(f)\} \exp\{-\varepsilon_n\} E_{n+1}(b_n q_n, f_{n+1}) \\ &= M_n(f) \exp\{-\varepsilon_n\} \frac{b_n q_n}{q_{n+1}} p_{n+1}(f_{n+1}). \end{aligned} \tag{2.49}$$

Below, we use the notation $\sum_\emptyset(\cdot) = 0$ and $\prod_\emptyset(\cdot) = 1$. The following inequality is straightforward when $m = k+1$, and readily implied by (2.48) taken with $n \in \{k+1, ..., m-1\}$ when $m \geqq k+2$. For each $m \geqq k+1$, we have

$$\begin{aligned} M_m(J,f) &\leqq M_{k+1}(J^*,f) \exp\Big\{\sum_{n=k+1}^{m-1} \varepsilon_n\Big\} \prod_{n=k+1}^{m-1} \Big(\frac{b_n q_n}{q_{n+1}} p_{n+1}(f_{n+1})\Big) \\ &\leqq M_{k+1}(J^*,f)\, \mathcal{E}\, \frac{q_{k+1}}{q_m} \Big\{\prod_{n=k+2}^{m} b_{n-1} p_n(f_n)\Big\}. \end{aligned} \tag{2.50}$$

Likewise, making use of (2.49), we obtain that, for $m \geqq n_0 + 1$

$$\begin{aligned} M_m(f) &\geqq M_{n_0}(f) \exp\Big\{-\sum_{n=n_0}^{m-1} \varepsilon_n\Big\} \prod_{n=n_0}^{m-1} \Big(\frac{b_n q_n}{q_{n+1}} p_{n+1}(f_{n+1})\Big) \\ &\geqq M_{n_0}(f)\, \mathcal{E}^{-1}\, \frac{q_{n_0}}{q_m} \Big\{\prod_{n=n_0+1}^{m} b_{n-1} p_n(f_n)\Big\} \\ &\geqq M_{n_0}(f) \Big\{\frac{\mathcal{E}^{-1}}{b_m}\Big\} \Big\{\frac{q_{n_0}^{1-|f|^2_{\mathbf{H}}}}{q_m^{\rho}}\Big\} \frac{1}{q_m^{1-\rho-|f|^2_{\mathbf{H}}}} \Big\{\prod_{n=n_0}^{m} b_n\Big\} \Big\{\prod_{n=n_0}^{m-1} (q_n^{|f|^2_{\mathbf{H}}}/2)\Big\}. \end{aligned} \tag{2.51}$$

Our assumptions imply that for $m \geqq n_0+1 \geqq 2$, $\mathcal{E}^{-1}/b_m \geqq 2\mathcal{E}^{-1} \geqq 1$. Moreover, we have $\rho/(1-|f|^2_{\mathbf{H}}) > \rho > \triangle > \triangle^2$. It follows therefore from (2.44)(i) that

$$q_{n_0}^{1-|f|^2_{\mathbf{H}}} q_m^{-\rho} \geqq \left\{ q_{n_0}\, q_{n_0+1}^{-\rho/(1-|f|^2_{\mathbf{H}})} \right\}^{1-|f|^2_{\mathbf{H}}} \geqq \left\{ q_{n_0}\, q_{n_0+1}^{-\triangle^2} \right\}^{1-|f|^2_{\mathbf{H}}} \geqq 1.$$

This, when combined with (2.49) and (2.44)(ii) entails that, for each $m \geqq n_0$,

(2.52) $$M_m(f) \geqq M_{n_0}(f).$$

We note for further use that (2.52) holds under (C.3) and (C.5) only. A consequence of this fact is that (C.1)–(C.3)–(C.5) jointly imply that $\hat{T}(f) \neq \emptyset$.

From now on, we assume that all five assumptions (C.1)–(C.5) hold. Here, (C.1) is used to ensure, via (2.52), that $M_m(f) \neq 0$ for $m \geqq k+1$. By combining (F.4) with (2.38), (2.50) and (2.51), we obtain that, for all $m \geqq n_0$,

(2.53) $$\begin{aligned} \frac{M_m(J,f)}{M_m(f)|J|^\rho} &\leqq \frac{\mathcal{E}^2 q_{k+1} M_{k+1}(J^*,f)}{M_{n_0}(f) q_{n_0} |J|^\rho} \left\{ \prod_{n=n_0+1}^{k+1} b_{n-1} p_n(f_n) \right\}^{-1} \\ &\leqq \frac{\mathcal{E}^2 q_{k+1} M_{k+1}(J^*,f)}{q_{n_0}|J|^\rho\, p_{k+1}(f_{k+1})} \left\{ \prod_{n=n_0}^{k} b_n \right\}^{-1} \left\{ \prod_{n=n_0+1}^{k} (q_n^{|f|^2_{\mathbf{H}}}/2) \right\}^{-1}. \end{aligned}$$

We complete (2.53) by an evaluation of $M_{k+1}(J^*,f)$. Towards this aim, we consider the following two cases.

Case 1. Assume that $|J| \in [q_{k+1}^{\triangle}, q_k)$. By (2.33)–(2.34), (2.36), (2.39) and (2.42),

$$\begin{aligned} M_{k+1}(J^*,f) &\leqq N_{q_{k+1}}(J^*_{k+1}, f_{k+1}) \leqq \exp\{\varepsilon_k\} \frac{|J^*_{k+1}|}{q_{k+1}} p_{k+1}(f_{k+1}) \\ &\leqq 3\exp\{\varepsilon_k\} \frac{|J|}{q_{k+1}} p_{k+1}(f_{k+1}) \leqq \frac{3\mathcal{E}|J|\, p_{k+1}(f_{k+1})}{q_{k+1}}. \end{aligned}$$

By combining this bound with (2.44)(ii) in (C.5), (2.53) and $|J| \leqq q_k$, we obtain that, for all $m \geqq k+1$

(2.54) $$\begin{aligned} \frac{M_m(J,f)}{M_m(f)|J|^\rho} &\leqq \frac{3\mathcal{E}^3 |J|^{1-\rho}}{q_{n_0}} \left\{ \prod_{n=n_0}^{k} b_n \right\}^{-1} \left\{ \prod_{n=n_0+1}^{k} (q_n^{|f|^2_{\mathbf{H}}}/2) \right\}^{-1} \\ &\leqq \frac{6\mathcal{E}^3}{q_{n_0}} q_k^{1-\rho-|f|^2_{\mathbf{H}}} \left\{ \prod_{n=n_0}^{k} b_n \right\}^{-1} \left\{ \prod_{n=n_0+1}^{k-1} (q_n^{|f|^2_{\mathbf{H}}}/2) \right\}^{-1} \leqq \frac{6\mathcal{E}^3}{q_{n_0}}. \end{aligned}$$

It follows in this case that (2.30) holds with $c = 6\mathcal{E}^3/q_{n_0}$.

Case 2. Assume that $|J| \in [q_{k+1}, q_{k+1}^{\triangle})$. In this case, we will need the assumption (C.4). By combining (F.4) with (2.33)–(2.34), (2.36), (2.39) and (2.42) in (C.4), we obtain that, for all $m \geqq k+1$,

$$\begin{aligned} M_{k+1}(J^*, f) &\leqq N_{q_{k+1}}(J^*_{k+1}, f_{k+1}) \\ &\leqq c_3 |J^*_{k+1}|^{1-|f|^2_{\mathbf{H}}-\triangle}\, \mathbf{E} N_{q_{k+1}}(f_{k+1}) \\ &\leqq 3c_3 |J|^{1-|f|^2_{\mathbf{H}}-\triangle} \left\{ \frac{p_{k+1}(f_{k+1})}{q_{k+1}} \right\}. \end{aligned}$$

Recall that $\rho = 1 - |f|^2_{\mathbf{H}} - 2\triangle$. The inequality above, when combined with (2.53) and (2.44)(ii), which entails that $|J|^{\triangle} \leqq q_{k+1}^{\triangle^2} \leqq q_k$, implies that, for all $m \geqq k+1$,

$$\begin{aligned} \frac{M_m(J,f)}{M_m(f)|J|^{\rho}} &\leqq \frac{3c_3\mathcal{E}^2|J|^{\triangle}}{q_{n_0}} \Big\{ \prod_{n=n_0}^{k} b_n \Big\}^{-1} \Big\{ \prod_{n=n_0+1}^{k} (q_n^{|f|^2_{\mathbf{H}}}/2) \Big\}^{-1} \\ &\leqq \frac{3c_3\mathcal{E}^2}{q_{n_0}} q_k \Big\{ \prod_{n=n_0}^{k} b_n \Big\}^{-1} \Big\{ \prod_{n=n_0+1}^{k} (q_n^{|f|^2_{\mathbf{H}}}/2) \Big\}^{-1} \\ &\leqq \frac{6c_3\mathcal{E}^2 q_k^{\rho}}{q_{n_0}} q_k^{1-\rho-|f|^2_{\mathbf{H}}} \Big\{ \prod_{n=n_0}^{k} b_n \Big\}^{-1} \Big\{ \prod_{n=n_0+1}^{k-1} (q_n^{|f|^2_{\mathbf{H}}}/2) \Big\}^{-1} \\ &\leqq \frac{6c_3\mathcal{E}^2}{q_{n_0}}. \end{aligned} \tag{2.55}$$

By combining (2.47), (2.54) and (2.55), we see that (2.30) holds with $c = 6(c_3 + \mathcal{E})\mathcal{E}^2/q_{n_0}$ and $\delta = q_{n_0+1}$. If follows from (2.46) and Fact 2 that $\dim(T(\nu,f)) \geqq \dim(\hat{T}(f)) \geqq \rho$. We note that this conclusion does not allow us yet to conclude (2.29) since the statements of (C.3), (C.4) and (C.5) depend not only upon the choice of $\rho < 1 - |f|^2_{\mathbf{H}}$, but also upon the construction of $\{f_n : n \geqq 1\}$ and $\{\mathcal{F}_n : n \geqq 1\}$. The next step will be needed to get rid of these restrictions.

STEP 2. We recall that, for our choice of $f \in \mathbf{K}$, $0 < |f|_{\mathbf{H}} < 1$. In this case it is always possible to define $\{f_n : n \geqq 1\}$ and $\{\mathcal{F}_n : n \geqq 1\}$ fulfilling (F.1)–(F.4) by setting $f_n = f$ and $\mathcal{F}_n = \{f\}$ for each $n \geqq 1$. Likewise, we may set $\triangle = \frac{1}{2}\{1 - |f|^2_{\mathbf{H}} - \rho\}$ for each specified choice of $\rho \in (1 - |f|^2_{\mathbf{H}} - \frac{1}{4}\min\{|f|^2_{\mathbf{H}}, 1 - |f|^2_{\mathbf{H}}\}, 1 - |f|^2_{\mathbf{H}})$. This allows us to make use of Step 1 to establish that, under (C.1)–(C.5), $\dim(T(\nu,f)) \geqq \rho$. We now turn to prove that, for each possible choice of $\rho \in (1 - |f|^2_{\mathbf{H}} - \frac{1}{4}\min\{|f|^2_{\mathbf{H}}, 1 - |f|^2_{\mathbf{H}}\}, 1 - |f|^2_{\mathbf{H}})$, there exists an event of probability 1 on which (C.1)–(C.5) hold. To infer from this fact that (2.29) holds with probability 1, we select a sequence $0 < \rho = \rho(n) \uparrow 1 - |f|^2_{\mathbf{H}}$. Since then $\dim(T(\nu,f)) \geqq \rho(n)$ with probability 1 for each n, the conclusion is straightforward.

(i) The Assumption (C.5) Since (C.5) is a collection of recurrent inequalities, the construction of a sequence $\{q_n : n \geqq n_1\}$ fulfilling (2.43) may be achieved by induction, given any choice of n_1 and of $q_{n_1} \geqq 1$.

(ii) The Assumption (C.2) By combining the Borel–Cantelli lemma with the triangle inequality and the definition (2.32) of $\mathcal{J}_{q_n}(f_n)$, we obtain readily that, independently of $f \in \mathbf{K}$ and $\{f_n : n \geqq 1\}$, (2.41) holds with probability 1 for all large n whenever

$$\begin{aligned} &\sum_{n=1}^{\infty} \mathbf{P}\Big(\sup_{0 \leqq i \leqq 1/q_n} \Big\{ \sup_{t \in [iq_n, (i+b_n)q_n]} \nu(Y_{t,(i+1)q_n - t} - Y_{iq_n, q_n}) \Big\} \geqq u_n \Big) \\ (2.56) \qquad &\leqq \sum_{n=1}^{\infty} 2q_n^{-1}\, \mathbf{P}\Big(\sup_{t \in [0, b_n q_n]} \nu(Y_{t, q_n - t} - Y_{0,q_n}) \geqq u_n \Big) =: \sum_{n=1}^{\infty} \mathcal{P}_n < \infty. \end{aligned}$$

Set, for convenience $g = q_n$, $b = b_n$, $\Gamma = 1 - b_n$ and $\varepsilon = 2u_n$. Keeping in mind that our assumptions imply that $g \downarrow 0$ and $\Gamma \uparrow 1$, we write

$$\begin{aligned} \mathcal{P}_n &= \mathbf{P}\Big(\sup_{t \in [0, b_n q_n]} \nu(Y_{t, q_n - t} - Y_{0,q_n}) \geqq u_n \Big) \\ &\leqq \mathbf{P}\Big(\sup_{t \in [0, bg]} \sup_{h \in [\Gamma g, g]} \nu(Y_{t,h} - Y_{0,g}) \geqq \varepsilon/2 \Big), \end{aligned}$$

which allows us to use the estimates (2.21), (2.22) and (2.28) with $\theta_0 = \max\{b, 1-\Gamma\} = b_n$, $1 - \Gamma = b_n$ and $\varepsilon = 2u_n$. We so obtain that

$$\begin{aligned} \mathcal{P}_n &\leqq \mathbf{P}\Big(\nu(W) \geqq L_g \varepsilon / \{12(1-\Gamma)\} \Big) + \mathbf{P}\Big(\|W\|_{b_n} \geqq L_g \varepsilon / 8 \Big) \\ (2.57) \qquad &= \mathbf{P}\Big(\nu(W) \geqq L_g u_n / \{6 b_n\} \Big) + \mathbf{P}\Big(\|W\|_{b_n} \geqq L_g u_n / 4 \Big). \end{aligned}$$

By (2.44)(iii), we have $L_g u_n = (2\log(1/q_n))^{1/2} u_n > 1/u_n \to \infty$ as $n \to \infty$. Let c_1 be chosen, via (2.13), in such a way that $\mathbf{P}(\nu(W) < c_1) > \Phi(1)$. This implies that $\beta = \beta(\nu, c_1, W) = c_1 / \Phi^{-1}(\mathbf{P}(W) < c_1) \leqq c_1$. We apply (2.17) with $R = L_g u_n / \{6b_n\}$ in combination with the fact that, ultimately, $(R - m)/\beta > (2/3)^{1/2} R \to \infty$. This shows that, for all large n,

$$\begin{aligned} \mathbf{P}\Big(\nu(W) \geqq L_g u_n / \{6b_n\} \Big) &\leqq 1 - \Phi\Big((2/3)^{1/2} L_g u_n / \{6 c_1 b_n\} \Big) \\ &\leqq \exp\Big(- L_g^2 \Big\{ \frac{u_n^2}{336 c_1^2 b_n^2} \Big\} \Big). \end{aligned}$$

Since the definition (2.40) of u_n implies that $u_n^2 / b_n^2 \geqq 1/b_n \to \infty$, for all large n,

$$\mathbf{P}\Big(\nu(W) \geqq L_g u_n / \{6b_n\} \Big) \leqq \exp(-L_g^2) = q_n^2. \tag{2.58}$$

Next, if c_2 is as in (2.24), and $\beta_n = \beta(\|\cdot\|_{b_n}, c_2+1, W)$, we infer from (2.17) and (2.27) that, whenever $R \geqq c_2+1$, we have

$$\mathbf{P}(\|W\|_{b_n} \geqq R) \leqq 1 - \Phi((R-c_2-1)/\beta_n).$$

By applying this inequality with $R = R_n := L_g u_n/4$, in combination with the fact that, ultimately, $R_n - c_2 - 1 \geqq (2/3)^{1/2} R_n \to \infty$, we obtain that

$$\mathbf{P}\Big(\|W\|_{b_n} \geqq L_g u_n/4\Big) \leqq 1 - \Phi\Big((2/3)^{1/2} R_n/\beta_n\Big) \leqq \exp\Big(-L_g^2\Big\{\frac{u_n^2}{316\beta_n^2}\Big\}\Big).$$

By combining the definition (2.40) of u_n, which implies that $u_n \geqq \beta_n^{1/2}$, with (2.45) we obtain that $u_n^2/\beta_n^2 \geqq 1/\beta_n \to \infty$. We have therefore, for all large n,

$$\mathbf{P}\Big(\|W\|_{b_n} \geqq L_g u_n/4\Big) \leqq \exp(-L_g^2) = q_n^2. \tag{2.59}$$

By combining (2.57) with (2.58) and (2.59), we see that, for all large n, $\mathcal{P}_n \leqq 2q_n^2$. Since (2.43)(ii) entails that $q_n = O(1/n!)$ as $n \to \infty$, we have (2.56).

(iii) The Assumption (C.1) We need show that, with probability 1 for all large n,

$$\min_{g \in \mathcal{F}_n} \#\mathcal{J}_{q_n}^{b_n}(g) \geqq 1.$$

Let $N = 1/q_n$. By (2.32), for each $n \geqq 1$ and $g \in \mathcal{F}_n$, $S_N := \#\mathcal{J}_{q_n}^{b_n}(g) = \#\mathcal{J}_{q_n}(g)$ follows a binomial distribution with parameters N and $p = p_n(g) = \mathbf{P}(\nu(Y_{q_n,0} - g) \leqq R/L_{q_n})$. Therefore, by (2.38) and the inequality $(1-u)^r \leqq \exp(-ru)$ for $r \geqq 0$ and $u \in [0,1)$,

$$\begin{aligned}\mathbf{P}(S_N = 0) &= \Big(1 - p_n(g)\Big)^N \leqq \Big(1 - \frac{1}{2} q_n^{|g|_{\mathbf{H}}^2}\Big)^{1/q_n} \leqq \exp\Big\{-\frac{1}{2} q_n^{-1+|g|_{\mathbf{H}}^2}\Big\} \\ &\leqq \exp\Big\{-\frac{1}{2} q_n^{-\triangle}\Big\} \leqq \frac{1}{q_n^2} \exp\Big\{-\frac{1}{8} q_n^{-\triangle}\Big\} =: \mathcal{R}_n.\end{aligned}$$

By (2.44)(iv), we obtain readily that $\sum_n \Theta_n \mathcal{R}_n$, which establishes our claim by combining the Bonferroni inequalities with the Borel–Cantelli lemma.

(iv) The Assumption (C.3) We start with the following fact concerning large deviation probabilities for binomial random variables. Let

$$\mathbf{h}(\lambda) = \begin{cases} \lambda \log \lambda - \lambda + 1 & \text{for } \lambda > 0, \\ 1 & \text{for } \lambda = 0. \end{cases} \tag{2.60}$$

FACT 4. *Let S_N follow a binomial distribution with parameters $N \geqq 1$ and $p \in [0,1]$. Then, for all $\lambda \in [1, 1/p]$,*

$$\mathbf{P}(S_N \geqq N\lambda p) \leqq \exp(-Np\mathbf{h}(\lambda)), \tag{2.61}$$

and for all $\lambda \in [0,1]$,

$$\mathbf{P}(S_N \leqq N\lambda p) \leqq \exp(-Np\mathbf{h}(\lambda)). \tag{2.62}$$

PROOF. See Lemma 3.8 in [8]. □

Since (C.5)(i) requires that $q_{n+1}^{\triangle} \leqq q_{n+1}^{\triangle^2} \leqq b_n q_n \leqq q_n$, we need check (2.42) for all intervals I of the form $I = [rq_n, sq_n]$ with $r, s \in \mathbf{N}$ and $0 \leqq r < s \leqq 1/q_n$. Moreover, by (2.36) we see that any two intervals I' and I'' within this class are such that $N_{q_{n+1}}(I' \cup I'', g) = N_{q_{n+1}}(I', g) + N_{q_{n+1}}(I'', g)$ whenever $|I' \cap I''| = 0$. In view of (2.39) our proof boils down to establish (2.42) for all intervals I of the form $I = I_r := [(r-1)q_n, rq_n]$ with $r \in \mathbf{N}$ and $1 \leqq r \leqq 1/q_n$. Now, for each $1 \leqq r \leqq 1/q_n$, $N_{q_{n+1}}(I_r, g)$ follows a binomial distribution with parameters $N = q_n/q_{n+1}$ and $p = p_{n+1}(g)$, as in (2.38). We note for further use that our assumptions entail that $\varepsilon_n \to 0$ as $n \to \infty$, whence, by (2.60), for all large n,

$$\frac{1}{4}\varepsilon_n^2 \leqq \mathbf{h}(\exp\{\pm\varepsilon_n\}) = \Big\{\frac{1+o(1)}{2}\Big\}\varepsilon_n^2. \tag{2.63}$$

Set, for convenience $\lambda^{\pm} = \exp\{\pm\varepsilon_n\}$. By (2.38), (2.44)(ii) and (2.61)-(2.63), we obtain readily that, for all large n and all $g \in \mathcal{F}_n$,

$$\begin{aligned}
&\mathbf{P}\Big(\bigcup_{r=1}^{1/q_n}\Big\{\pm\frac{N_{q_{n+1}}(I_r, g)}{\mathbf{E}N_{q_{n+1}}(I_r, g)} \geqq \pm\exp\{\pm\varepsilon_n\}\Big\}\Big) \leqq \frac{1}{q_n}\mathbf{P}\Big(\pm S_N \geqq \pm N\lambda^{\pm}p\Big) \\
&\leqq \frac{1}{q_n}\exp\Big\{-Np\mathbf{h}(\lambda^{\pm})\Big\} \leqq \frac{1}{q_n}\exp\Big\{-\frac{q_n}{q_{n+1}}\frac{1}{2}q_{n+1}^{|g|_{\mathbf{H}}^2}\mathbf{h}(\exp\{\pm\varepsilon_n\})\Big\} \\
&\leqq \frac{1}{q_n}\exp\Big\{-\frac{1}{8}q_n\varepsilon_n^2 q_{n+1}^{-2\triangle}\Big\} \leqq \frac{1}{q_{n+1}^2}\exp\Big\{\frac{1}{8}q_{n+1}^{-\triangle}\Big\} = \mathcal{R}_{n+1}.
\end{aligned}$$

Since (2.44)(iv) implies that $\sum_n \Theta_n \mathcal{R}_n < \infty$, the proof that (2.42) holds with probability 1 for all n_0 sufficiently large is completed by the Borel–Cantelli lemma.

(v) The Assumption (C.4) First, recall from (2.39) taken with $I = [-1, 1]$ that

$$\mathbf{E}N_{q_{n+1}}(g) = p_{n+1}(g)/q_{n+1}.$$

Next, observe that, for each $I = [rq_{n+1}, sq_{n+1}]$ with $0 \leqq r < s \leqq 1/q_{n+1}$, $N_{q_{n+1}}(I, g)$ follows a binomial distribution with parameters $N = s - r$ and $p = p_{n+1}(g)$. It follows from (2.62) that

$$
\begin{aligned}
&\mathbf{P}\Big(N_{q_{n+1}}(I,g) \geqq c|I|^{1-|g|^2_{\mathbf{H}}-\triangle}\,\mathbf{E}N_{q_{n+1}}(g)\Big) \\
&= \mathbf{P}\Big(N_{q_{n+1}}(I,g) \geqq c\{Nq_{n+1}\}^{1-|g|^2_{\mathbf{H}}-\triangle}\{p_{n+1}(g)/q_{n+1}\}\Big) \\
&= \mathbf{P}\Big(S_N \geqq N\Big\{c(Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle}\Big\}p\Big) \\
&\leqq \exp\Big\{-Np\mathbf{h}\Big(c(Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle}\Big)\Big\}.
\end{aligned} \tag{2.64}
$$

Observe that $c(Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle} \geqq c$. Since $\mathbf{h}(\lambda) = (1+o(1))\lambda\log\lambda > \lambda$ for all large λ, there exists a $c_3 \geqq 1$ such that

$$
\mathbf{h}(c_3(Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle}) \geqq c_3(Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle} \geqq (Nq_{n+1})^{-|g|^2_{\mathbf{H}}-\triangle}.
$$

This, when combined with (2.38), (2.64) and $N = s - r \geqq 1$, implies that for all large n

$$
\begin{aligned}
\mathbf{P}\Big(N_{q_{n+1}}(I,g) \geqq c|I|^{1-|g|^2_{\mathbf{H}}-\triangle}\,\mathbf{E}N_{q_{n+1}}\Big) &\leqq \exp\Big\{-\frac{1}{2}N^{1-|g|^2_{\mathbf{H}}-\triangle}q_{n+1}^{-\triangle}\Big\} \\
&\leqq \exp\Big\{-\frac{1}{2}q_{n+1}^{-\triangle}\Big\}.
\end{aligned} \tag{2.65}
$$

Denote for convenience by A_n the event that (2.43) does not hold. Since the total number of intervals $I = [rq_{n+1}, sq_{n+1}]$ with $0 \leqq r < s \leqq 1/q_{n+1}$ is bounded above by $1/q^2_{n+1}$, it follows from (2.65) in combination with the Bonferroni inequalities that

$$
\mathbf{P}(A_n) \leqq \frac{\Theta_{n+1}}{q^2_{n+1}}\exp\Big\{-\frac{1}{8}q_{n+1}^{-\triangle}\Big\} = \Theta_{n+1}\mathcal{R}_{n+1}.
$$

Since (2.44)(iv) implies that $\sum_n \Theta_n\mathcal{R}_n < \infty$, the proof that (2.43) holds with probability 1 for all n_0 sufficiently large is achieved via the Borel–Cantelli lemma.

STEP 3. We conclude by showing how to modify our just-given proof of (2.29) for a specified $f \in \mathbf{K}$ in order to assess its validity uniformly over all $f \in \mathbf{K}$. A crucial step in this argument is the fact, following from Theorem 2.1 in [6], that $\mathbf{K}$ is a compact subset of $(C_0[0,1], \nu)$ for each $\nu \in \mathcal{C}$. This implies in particular that, for each $g \in \mathbf{H}$

$$
\nu(g) \leqq |g|_{\mathbf{H}} \sup_{h\in\mathbf{K}} \nu(h) =: |g|_{\mathbf{H}} K_\nu, \tag{2.66}
$$

where $K_\nu < \infty$ is a finite constant.

For any $\triangle \in (0, \frac{1}{16})$, set

$$
\mathbf{K}^{\triangle} = \{f \in \mathbf{K} : 8\triangle \leqq |f|^2_{\mathbf{H}} \leqq 1 - 8\triangle\} \subseteq \mathbf{K}.
$$

Since $\mathbf{H}$ is separable, there exists a sequence $\mathcal{G}=\{g_n : n \geqq 1\} \subseteq \mathbf{K}$ fulfilling (i)–(ii)–(iii) below. Set $\mathcal{M}(\triangle, f)=\{n \geqq 1 : |g_n|_{\mathbf{H}} \leqq |f|_{\mathbf{H}},\ 2\triangle \leqq |g_n|_{\mathbf{H}} \leqq 1-2\triangle\}$ for $f \in \mathbf{K}^{\triangle}$. Then,

(2.67)
(i) $\mathcal{G}$ is dense in $(\mathbf{K}, |\cdot|_{\mathbf{H}})$;
(ii) For each $f \in \mathbf{K}^{\triangle}$, $\mathcal{M}(\triangle, f) \neq \emptyset$;
(iii) For each $f \in \mathbf{K}^{\triangle}$, $\lim_{n\to\infty} \Big\{ \min_{\substack{1 \leqq m \leqq n \\ m \in \mathcal{M}(\triangle, f)}} |f-g_m|_{\mathbf{H}} \Big\} = 0.$

It follows from (2.66) and (2.67) that $\mathcal{G}$ is dense in $(\mathbf{K}, \nu)$. Moreover, the compactness of $\mathbf{K}$ in $(C_0[0,1], \nu)$ entails that, for each sequence $\{\theta_n : n \geqq 1\}$ with $\theta_n \downarrow 0$, there exist integers $1 \leqq \Theta_1 \leqq \Theta_2 \leqq \ldots$, such that the following property holds. For each $n \geqq 1$ and $f \in \mathbf{K}_{\triangle}$,

$$(2.68) \qquad \mathcal{M}(\triangle, f) \cap \{1, \ldots, \Theta_n\} \neq \emptyset \quad \text{and} \quad \min_{\substack{1 \leqq i \leqq \Theta_n \\ i \in \mathcal{M}(\triangle, f)}} \nu(f-g_i) < \theta_n.$$

We now set $\mathcal{F}_n=\{g_i : 1 \leqq i \leqq \Theta_n\}$. For any $f \in \mathbf{K}^{\triangle}$, we define a sequence $\{f_n : n \geqq 1\}$ by setting, via (2.68), for each $n \geqq 1$, $f_n = g_i$ for some $g_i \in \mathcal{F}_n$ fulfilling $\nu(f-g_i) < \theta_n$ and $i \in \mathcal{M}(\triangle, f) \cap \{1, \ldots, \Theta_n\}$. It is obvious that $\{\mathcal{F}_n : n \geqq 1\}$ and $\{f_n : n \geqq 1\}$ fulfill (F.1)–(F.4). Moreover, it follows from Step 2 that we may construct with probability 1 a sequence $\{q_n : n \geqq n_0\}$ fulfilling (C.1)–(C.5), independently of $f \in \mathbf{K}^{\triangle}$ and $\{f_n : n \geqq 1\}$ as defined above. Therefore, the inequality $\dim(T(\nu, f)) \geqq 1-|f|^2_{\mathbf{H}}$ holds with probability 1 for all $f \in \mathbf{K}^{\triangle}$. By applying this result to a sequence $\triangle = \triangle_n \downarrow 0$, we obtain readily that the same statement holds with $\mathbf{K}^{\triangle}$ replaced by $\bigcup_n \mathbf{K}^{\triangle_n} = \mathbf{K} - \{0\}$. By treating separately the case of $f=0$, we conclude that (2.29) holds with probability 1 uniformly over all $f \in \mathbf{K}$. In view of the arguments of Sections 2.1 and 2.2, this last step completes the proof of Theorem 1.2.

3. Large deviation theorems for differences of norms

In the following Theorem 3.1, we evaluate upper tail probabilities for differences of lower semi-continuous norms on a Gaussian space. In spite of the fact that this result is very similar to analogue upper tail bounds which have been described in the literature for distributions of Lipshitz functionals, to our best knowledge it does not seem to be an immediate consequence of any classical estimate of this type. Therefore, it has interest in and of itself. We inherit the notation of the previous sections. The following facts will be useful.

FACT 5. *The space $\mathcal{X}^*_Z$ of measurable linear forms is separable in $L^1(\mathcal{X}, \mathcal{B}, \mathbf{P}_Z)$.*

PROOF. See Theorem 2, p. 86 in [15]. □

FACT 6. *A mapping* $\nu : \mathcal{X} \to [0, \infty]$ *is a lower semi-continuous semi-norm on a Hausdorff locally convex space* $\mathcal{X}$ *if and only if there exists a convex and symmetric subset* $\Pi \subseteqq \mathcal{X}^*$ *such that* $\nu(x) = \sup_{\pi \in \Pi} |\pi(x)|$ *for all* $x \in \mathcal{X}$.

PROOF. See e.g. Chapter II in Bourbaki [2]. □

FACT 7. *Let* Z *denote a centered Gaussian random variable with distribution given by a Radon measure* $\mathbf{P}_Z$ *on a Hausdorff locally convex space* $\mathcal{X}$. *Then, for any convex subset* $C \in \mathcal{B}_{\mathcal{X}}$ *such that* $C = -C$, *and each* $x \in \mathcal{X}$, *we have*

$$\mathbf{P}_Z(C) \geqq \mathbf{P}_Z(C + x). \tag{3.1}$$

PROOF. This is Anderson's inequality, see Theorem 9, p. 135 in [15]). □

We denote, as usual, by Φ the distribution function of a standard normal random variable.

THEOREM 3.1. *Let* $\|\cdot\|_1 = \sup_{\pi \in \Pi_1} |\pi(\cdot)|$ *and* $\|\cdot\|_2 = \sup_{\pi \in \Pi_2} |\pi(\cdot)|$ *denote two lower semi-continuous seminorms on a Hausdorff locally convex space* $\mathcal{X}$, *generated by the families of continuous linear forms* $\Pi_1 \subseteqq \mathcal{X}^*$ *and* $\Pi_2 \subseteqq \mathcal{X}^*$. *Let* Z *denote a centered* $\mathcal{X}$*-valued Gaussian vector with Radon distribution* $\mathbf{P}_Z$, *and denote by* $\mathbf{K}_Z$ *the unit ball of the RKHS of* $\mathbf{P}_Z$.

Let $M \in \mathbf{R}$ *be any constant so large that*

$$\mathbf{P}(\|Z\|_1 \leqq M) \geqq 1/2 \quad \textit{and} \quad \epsilon := \mathbf{E}\Big|\|Z\|_2 - \|Z\|_1\Big| \leqq M/12. \tag{3.2}$$

Let $\sigma_0 = 3M/\Phi^{-1}(1 - 2\epsilon/M)$. *Then, for each* $r \geqq 0$ *and* $R \geqq 3M + \sigma_0 r$,

$$\mathbf{P}\Big(\sup_{f \in r\mathbf{K}_Z} \Big\{\|Z + f\|_2 - \|Z + f\|_1\Big\} \geqq R\Big) \leqq 1 - \Phi\Big(\frac{R - 5M - r\sigma_0}{\sigma_0}\Big). \tag{3.3}$$

Moreover, if $\mathbf{P}(\|Z\|_1 \leqq M) \geqq 3/4$, *then for each* $r \geqq 0$ *and* $R \geqq 4M + \frac{4}{3}\sigma_0 r$,

$$\mathbf{P}\Big(\sup_{f \in r\mathbf{K}_Z} \Big|\|Z + f\|_2 - \|Z + f\|_1\Big| \geqq R\Big) \leqq 2\Big\{1 - \Phi\Big(\frac{3R/4 - 5M - r\sigma_0}{\sigma_0}\Big)\Big\}. \tag{3.4}$$

PROOF. We proceed in three steps.

STEP 1. *Estimate of the size of the set* Π_1.

Denote by $|\pi|_2 = \{\mathbf{E}|\pi(Z)|^2\}^{1/2}$ the L^2-norm induced by $\mathbf{P}_Z$ on $\mathcal{X}^* \subseteqq L^2(\mathcal{X}, \mathcal{B}, \mathbf{P}_Z)$ via the mapping $\mathrm{I}_Z : \pi \in \mathcal{X}^* \to \pi(Z) \in L^2$. And observe that, with this notation, $\pi(Z)$ follows a normal $N(0, |\pi|_2^2)$ distribution for each $\pi \in \mathcal{X}^*$. The definition of $\|\cdot\|_1$ implies that $|\pi(Z)| \leqq \|Z\|_1$ for each $\pi \in \Pi_1$. Thus, by (3.2), we have for each $\pi \in \Pi_1$

$$\mathbf{P}(\pi(Z) \leqq M) = \Phi(M/|\pi|_2) = 1 - \frac{1}{2}\mathbf{P}(|\pi(Z)| \geqq M) \geqq 1 - \frac{1}{2}\mathbf{P}(\|Z\|_1 \geqq M) \geqq \frac{3}{4}.$$

Since $\Phi^{-1}(3/4) = 0.67449... \geqq 0.66666... = 1/(1.5)$, we infer from this inequality that

$$\sup_{\pi \in \Pi_1} |\pi|_2 \leqq M/\{\Phi^{-1}(3/4)\} \leqq 1.5M. \tag{3.5}$$

STEP 2. *Estimate of the closeness of the sets* Π_1 *and* Π_2.

In view of Fact 6, without loss of generality, we may and do assume that the sets Π_1 and Π_2 are convex and symmetric in $\mathcal{X}^*$. We will prove in this case that

$$\sup_{\pi_2 \in \Pi_2} \left\{ \inf_{\pi_1 \in \Pi_1} |\pi_1 - \pi_2|_2 \right\} \leqq \sigma_0. \tag{3.6}$$

Fix an arbitrary $\pi_2 \in \Pi_2$ and assume that

$$\sigma := \sigma(\pi_2) := \inf_{\pi \in \Pi_1} |\pi_2 - \pi_1|_2 > 0.$$

Choose any $\delta \in (0, \sigma/2)$. Then, there always exists a $\hat{\pi} \in \Pi_1$ such that

$$\sigma \leqq |\pi_2 - \hat{\pi}|_2 \leqq \sigma + \delta < \frac{3}{2}\sigma. \tag{3.7}$$

Define a continuous linear form $g \in \mathcal{X}^*$ on $\mathcal{X}$ by setting

$$g = (\pi_2 - \hat{\pi})/|\pi_2 - \hat{\pi}|_2.$$

It is noteworthy that $g(Z)$ follows a standard normal $N(0,1)$ law. Moreover, since, for each $\pi \in \Pi_1$, the joint distribution of $\pi(Z)$ and $g(Z)$ is centered normal in $\mathbf{R}^2$, if we set $c(\pi, g) = \mathbf{E}(\pi(Z)g(Z))$ and define π_g via

$$\pi = \pi_g + c(\pi, g)g,$$

then $\pi_g(Z) := \pi(Z) - \mathbf{E}(\pi(Z)g(Z))Z$ is independent of $g(Z)$. Since Π_1 is symmetric, it follows that

$$\begin{aligned}
\|Z\|_2 - \|Z\|_1 &= \sup_{\pi \in \Pi_2} |\pi(Z)| - \sup_{\pi \in \Pi_1} |\pi(Z)| \geqq \pi_2(Z) - \sup_{\pi \in \Pi_1} \pi(Z) \\
&= \hat{\pi}(Z) + |\pi_2 - \hat{\pi}|_2 g(Z) - \sup_{\pi \in \Pi_1} \left\{ \pi_g(Z) + c(\pi, g)g(Z) \right\} \\
&\geqq \left\{ c(\hat{\pi}, g) + |\pi_2 - \hat{\pi}|_2 - \sup_{\pi \in \Pi_1} c(\pi, g) \right\} g(Z) - 2 \sup_{\pi \in \Pi_1} |\pi_g(Z)| \qquad (3.8) \\
&= \left\{ |\pi_2 - \hat{\pi}|_2 - \sup_{\pi \in \Pi_1} [c(\pi, g) - c(\hat{\pi}, g)] \right\} g(Z) - 2 \sup_{\pi \in \Pi_1} |\pi_g(Z)|.
\end{aligned}$$

By geometric arguments based upon the convexity of Π_1 we obtain that, for each $\pi \in \Pi_1$,

$$
\begin{aligned}
c(\pi,g) - c(\hat{\pi},g) &\leqq |\pi - \hat{\pi}|_2 \Big\{ |\pi_2 - \hat{\pi}|_2^2 - \sigma^2 \Big\}^{1/2} / \, |\pi_2 - \hat{\pi}|_2 \\
&\leqq \Big\{ |\pi|_2 + |\hat{\pi}|_2 \Big\} \Big\{ 2\sigma\delta + \delta^2 \Big\}^{1/2} / \sigma \\
&\leqq 2 \Big\{ \sup_{\pi \in \Pi_1} |\pi|_2 \Big\} \Big\{ \frac{2\delta}{\sigma} + \frac{\delta^2}{\sigma^2} \Big\}^{1/2}.
\end{aligned}
\tag{3.9}
$$

By (3.5), (3.7), (3.9), $0 < \delta < \sigma/2$ and $5/2 = 2.5 < (1.6)^2 = 2.56$, we obtain that

$$
c(\pi,g) - c(\hat{\pi},g) \leqq 2 \Big\{ 1.5M \Big\} \Big\{ \frac{5\delta}{2\sigma} \Big\}^{1/2} \leqq 3M 1.6 (\delta/\sigma)^{1/2} \leqq 5(\delta/\sigma)^{1/2}. \tag{3.10}
$$

By combining (3.7), (3.8) and (3.10), we see that the condition $g(Z) \geqq 0$ implies that

$$
\|Z\|_2 - \|Z\|_1 \geqq \Big\{ \sigma - 5(\delta/\sigma)^{1/2} \Big\} g(Z) - 2 \sup_{\pi \in \Pi_1} |\pi_g(Z)|.
$$

In particular, with the notation of the Theorem,

$$
\begin{aligned}
\epsilon M^{-1} = \mathbf{E}\Big(\Big| \|Z\|_1 - \|Z\|_2 \Big| \Big) M^{-1} &\geqq \mathbf{P}\Big(\|Z\|_2 - \|Z\|_1 \geqq M \Big) \\
&\geqq \mathbf{P}\Big(\Big\{ \sigma - 5(\delta/\sigma)^{1/2} \Big\} g(Z) \geqq 3M \Big) \ \mathbf{P}\Big(\sup_{\pi \in \Pi_1} |\pi_g(Z)| \leqq M \Big).
\end{aligned}
\tag{3.11}
$$

It follows from Anderson's inequality (3.1) that

$$
\mathbf{P}\Big(\sup_{\pi \in \Pi_1} |\pi_g(Z)| \leqq M \Big) \geqq \mathbf{P}\Big(\sup_{\pi \in \Pi_1} |\pi(Z)| \leqq M \Big) \geqq 1/2. \tag{3.12}
$$

Thus, by letting $\delta \downarrow 0$ in (3.11) we obtain that

$$
\mathbf{P}\Big(\sigma g(Z) \geqq 3M \Big) = 1 - \Phi(3M/\sigma) \leqq 2\epsilon M^{-1},
$$

and hence,

$$
\sigma \leqq 3M / \{ \Phi^{-1}(1 - 2\epsilon/M) \} = \sigma_0 .
$$

STEP 3. *Reduction to a convex functional.*

Let $\delta > 0, f \in r\mathbf{K}$. For each $\pi_2 \in \Pi_2$ we can find a $\hat{\pi}_2 \in \Pi_1$ such that $|\pi_2 - \hat{\pi}_2|_2 \leqq \sigma_0 + \delta$, the application $\pi_2 \to \pi_1$ being measurable. Set $\xi =$

$\sup_{\pi_2\in\Pi_2}\{\pi_2-\bar{\pi}_2\}(Z)$. We start with the inequalities

$$\begin{aligned}
\|Z+f\|_2-\|Z+f\|_1 &= \sup_{\pi_2\in\Pi_2}\pi_2(Z+f)-\sup_{\pi_1\in\Pi_1}\pi_1(Z+f)\\
&\leqq \sup_{\pi_2\in\Pi_2}\{\pi_2-\bar{\pi}_2\}(Z+f)\\
&\leqq \xi+\sup_{\pi_2\in\Pi_2}|\{\pi_2-\bar{\pi}_2\}(f)|\leqq \xi+\sup_{\pi_2\in\Pi_2}\{|\pi_2-\bar{\pi}_2|_2\}r\leqq \xi+(\sigma_0+\delta)r \qquad (3.13)\\
&\leqq \|Z\|_2+\|Z\|_1+(\sigma_0+\delta)r\leqq 2\|Z\|_1+\Big|\|Z\|_1-\|Z\|_2\Big|+(\sigma_0+\delta)r.
\end{aligned}$$

Note that our assumptions imply that $2M+4\epsilon\leqq(2+1/3)M=$ $=(2.333...)\times M\leqq 2.34M$. By applying (3.13) with $f=0$ and $r=0$ we see that the distribution of ξ obeys the bounds

$$\begin{aligned}
\mathbf{P}\Big(\xi\leqq 2.34M\Big)&\geqq \mathbf{P}\Big(\xi\leqq 2M+4\epsilon\Big)\\
&\geqq \mathbf{P}\Big(2\|Z\|_1\leqq 2M\Big)-\mathbf{P}\Big(\Big|\|Z\|_1-\|Z\|_2\Big|\geqq 4\epsilon\Big)\\
&\geqq 1/2-1/4=1/4.
\end{aligned}$$

Keeping in mind that $\Phi^{-1}(1/4)=-0.6744...\geqq-0.68$, we use the isoperimetric inequality (2.10), to obtain that, for each $R\geqq 2.34M$

$$\begin{aligned}
\mathbf{P}\Big(\xi\geqq R\Big)&\leqq 1-\Phi\Big(\Big\{R-2.34M\Big\}/\Big\{\sup_{\pi_2\in\Pi_2}|\pi_2-\bar{\pi}_2|_2\Big\}+\Phi^{-1}(1/4)\Big)\\
&\leqq 1-\Phi\Big(\frac{R-2.34M}{\sigma_0+\delta}-0.68\Big).
\end{aligned}$$

Since $\epsilon\leqq M/12$, we have $\Phi^{-1}(1-2\varepsilon/M)\geqq\Phi^{-1}(1-1/6)\geqq\Phi^{-1}(0.833)\geqq 0.96$. The definition of $\sigma_0=3M/\Phi^{-1}(1-2\epsilon/M)$ yields $0.68\,\sigma_0\leqq\{30.68/0.96\}M\leqq$ $2.13M$ and

$$\mathbf{P}\Big(\xi\geqq R\Big)\leqq 1-\Phi\Big(\frac{R-2.34M-0.68\,\sigma_0-\delta}{\sigma_0+\delta}\Big)\leqq 1-\Phi\Big(\frac{R-5M-\delta}{\sigma_0+\delta}\Big). \qquad (3.14)$$

In view of (3.13) we obtain (3.2) by letting $\delta\downarrow 0$ in (3.14).

To obtain (3.4), we note that, whenever $\mathbf{P}(\|Z\|_1\leqq M)\geqq 3/4$, the Chebyshev inequality yields

$$\begin{aligned}
\mathbf{P}(\|Z\|_2\leqq 4M/3)&\geqq \mathbf{P}(\|Z\|_1\leqq M)-\mathbf{P}(\|Z\|_2-\|Z\|_1\geqq M/3)\\
&\geqq \mathbf{P}(\|Z\|_1\leqq M)-\mathbf{P}(\|Z\|_2-\|Z\|_1\geqq 4\varepsilon)\geqq 3/4-1/4=1/2.
\end{aligned}$$

By applying (3.3) with the formal change of $\|\cdot\|_1$ and $\|\cdot\|_2$ into $\|\cdot\|_2$ and $\|\cdot\|_1$, and with the formal replacements of M and σ_0 by $M'=4M/3$ and $\sigma_0'=3M'/\Phi^{-1}(1-2\epsilon/M')\leqq 4\sigma_0/3$, respectively, we obtain readily (3.4). □

4. Norm-independent exceptional sets

In this section, we inherit the notations of Sections 1–3. In particular, W denotes a Wiener process, $\mathbf{K}$ stands for the Strassen set, and $T(\nu,\alpha)$ and $T(\nu,f)$ are the exceptional sets defined in (1.5). We aim to prove Theorem 1.3 by showing that these exceptional sets are essentially independent on the consistent norm $\nu\in\mathcal{C}$. Our proof will be decomposed into the following two steps.

STEP 1. Consider any countable family $\mathcal{N}$ of consistent norms dense in $\mathcal{C}$ with respect to the L^1-norm of the Wiener measure, and let $\mathbf{K}'$ denote any countable subset of $\mathbf{K}$ dense in $(\mathbf{K},|\cdot|_{\mathbf{H}})$. We will show that (1.12)–(1.13) and (1.14)–(1.15) hold with

$$T''_\alpha=\bigcup_{\nu\in\mathcal{N}}T(\nu,\alpha),\quad T_f=\bigcap_{\nu\in\mathcal{N}}T(\nu,f),\quad T'_\alpha=\bigcup_{f\in\mathbf{K}',|f|_{\mathbf{H}}>\alpha}T_f. \tag{4.1}$$

Since $\mathcal{N}$ is countable, it follows readily from Theorem 1.2, in combination with the σ-stability of Hausdorff dimension, that, with probability 1 for each $\alpha\in[0,1]$,

$$\dim(T''_\alpha)=1-\alpha^2. \tag{4.2}$$

Moreover, Theorem 1.2 also implies that, with probability 1 for each $f\in\mathbf{K}$,

$$\dim(T_f)=1-|f|^2_{\mathbf{H}}. \tag{4.3}$$

It follows readily from (4.2)–(4.3) that, with probability 1 for every $\alpha\in[0,1)$,

$$\dim(T'_\alpha)\geqq\sup_{f\in\mathbf{K}',|f|_{\mathbf{H}}>\alpha}\dim(T_f)=\sup_{f\in\mathbf{K}',|f|_{\mathbf{H}}>\alpha}\left(1-|f|^2_{\mathbf{H}}\right)=1-\alpha^2.$$

We note that this inequality also holds trivially when $\alpha=1$. On the other hand, the obvious inclusion $T'_\alpha\subset T''_\alpha$, when combined with (4.2), yields

$$\dim(T'_\alpha)\leqq\dim(T''_\alpha)=1-\alpha^2.$$

Thus, with probability 1, uniformly over all $\alpha\in[0,1]$

$$\dim(T'_\alpha)=\dim(T''_\alpha)=1-\alpha^2.$$

The proof of the "dimensional" part (1.15)–(1.16) of Theorem 1.3 is therefore completed. The following arguments are aimed towards proving the remaining inclusions (1.17)–(1.18).

We make use again of the sequences $\gamma_n=e^{-\sqrt{n}}$ and $t_{j,n}=jn^{-1}\gamma_n$ for $j\in\mathbf{N}$ and $n\geqq1$. For each $t\in[0,1)$ and $h\in(0,e^{-1})$ there exist unique integers

$n = n(t) \geqq 1$ and $i = i(t) \geqq 0$ such that $h \in [\gamma_{n+1}, \gamma_n)$ and $t \in [t_{i,n}, t_{i+1,n})$. For such choices of n and i, set

$$\hat{Y}_{t,h} = Y_{t_{i,n},\gamma_n}.$$

The random field $\hat{Y}_{t,h}$ is easier to handle than $Y_{t,h}$, being defined on a discrete set of indices. By applying the Borel–Cantelli lemma in combination with Lemma 2.2, we obtain readily that for each consistent norm ν with probability 1

$$\lim_{h\to 0} \Big\{ \sup_{t\in[0,1]} \nu(\hat{Y}_{t,h} - Y_{t,h}) \Big\} = 0. \tag{4.4}$$

We omit the details of this argument. In the sequel we will assume, without loss of generality that our random variables are defined on the event of probability 1 on which (4.4) holds for all $\nu \in \mathcal{N}$.

LEMMA 4.1. *Let $\eta \in \mathcal{C}$ and $\nu \in \mathcal{C}$ be any two consistent norms, and set $\varepsilon = \varepsilon(\eta, \nu) = \mathbf{E}|\eta(W) - \nu(W)|$. Assume that $M = M(\eta) > 0$ is so large that $\mathbf{P}(\eta(W) \leqq M) \geqq 3/4$ and $\varepsilon \leqq M/12$. Then, with probability* 1,

$$\lim_{h\downarrow 0} \Big(\sup_{t\in[0,1]} \Big\{ \sup_{f\in\mathbf{K}} |\nu(\hat{Y}_{t,h} - f) - \eta(\hat{Y}_{t,h} - f)| \Big\} \Big) \leqq \Delta, \tag{4.5}$$

where $\Delta := \Delta(\eta, \nu) = 9M/\Phi^{-1}(1 - 2\varepsilon/M)$.

PROOF. By definition of $\hat{Y}_{t,h}$ and making use of the distributional invariance by translation of the Wiener process increments, we obtain that, for any $n \geqq 1$

$$\begin{aligned}
&\mathbf{P}\Big(\sup_{h\in[\gamma_{n+1},\gamma_n)} \sup_{t\in[0,1]} \Big\{ \sup_{f\in\mathbf{K}} |\nu(\hat{Y}_{t,h} - f) - \eta(\hat{Y}_{t,h} - f)| \Big\} \geqq \Delta \Big) \\
&\leqq \sum_{0\leqq jn^{-1}\gamma_n\leqq 1} \mathbf{P}\Big(\sup_{f\in\mathbf{K}} |\nu(Y_{t_{j,n},\gamma_n} - f) - \eta(Y_{t_{j,n},\gamma_n} - f)| \geqq \Delta \Big) \\
&= \lfloor n\gamma_n^{-1} + 1 \rfloor\, \mathbf{P}\Big(\sup_{f\in\mathbf{K}} |\nu(Y_{0,\gamma_n} - f) - \eta(Y_{0,\gamma_n} - f)| \geqq \Delta \Big) \\
&\leqq 2n\gamma_n^{-1}\, \mathbf{P}\Big(\sup_{f\in L(\gamma_n)\mathbf{K}} |\nu(W - f) - \eta(W - f)| \geqq \Delta L(\gamma_n) \Big).
\end{aligned}$$

We now apply Theorem 3.1 to $\|\cdot\|_1 = \eta$, $\|\cdot\|_2 = \nu$, $r = L(\gamma_n)$, $R = \Delta L(\gamma_n)$ and $\sigma_0 = \Delta/3$. Since we have $5M + r\sigma_0 < 3\Delta L(\gamma_n)/8 = 3R/8$ for all large n, by (3.4), we get

$$\begin{aligned}
&\mathbf{P}\Big(\sup_{f\in L(\gamma_n)\mathbf{K}} |\nu(W - f) - \eta(W - f)| \geqq \Delta L(\gamma_n) \Big) \leqq 2[1 - \Phi(3R/8\sigma_0)] \\
&= 2[1 - \Phi(9L(\gamma_n)/8)] \leqq 2\exp\{-9^2 L(\gamma_n)^2/(2\times 8^2)\} \leqq \gamma_n^{9/8}.
\end{aligned}$$

The convergence of the series $\sum_n n\gamma_n^{1/8}$ implies therefore that

$$\sum_{n=1}^{\infty} \mathbf{P}\Big(\sup_{h\in[\gamma_{n+1},\gamma_n)} \Big(\sup_{t\in[0,1]} \Big\{ \sup_{f\in\mathbf{K}} |\nu(\hat{Y}_{t,h} - f) - \eta(\hat{Y}_{t,h} - f)| \Big\}\Big) \geqq \Delta \Big) < \infty.$$

The Borel–Cantelli lemma completes the proof of (4.5). □

STEP 2. For each $t\in[0,1]$, $\alpha\in[0,1]$, and any two norms $\eta\in\mathcal{C}$ and $\nu\in\mathcal{C}$, we have

$$\begin{aligned}
&\Big|\limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \nu(Y_{t,h}-f) - \limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \eta(Y_{t,h}-f)\Big| \\
&\leqq \limsup_{h\downarrow 0} \sup_{f\in\alpha\mathbf{K}} \Big|\nu(Y_{t,h}-f) - \eta(Y_{t,h}-f)\Big| \\
&\leqq \limsup_{h\downarrow 0} \sup_{f\in\alpha\mathbf{K}} \Big|\nu(Y_{t,h}-f) - \nu(\hat{Y}_{t,h}-f)\Big| \\
&\quad + \limsup_{h\downarrow 0} \sup_{f\in\alpha\mathbf{K}} \Big|\eta(Y_{t,h}-f) - \eta(\hat{Y}_{t,h}-f)\Big| \\
&\quad + \limsup_{h\downarrow 0} \sup_{f\in\alpha\mathbf{K}} \Big|\nu(\hat{Y}_{t,h}-f) - \eta(\hat{Y}_{t,h}-f)\Big| \\
&\leqq \limsup_{h\downarrow 0} \nu(Y_{t,h} - \hat{Y}_{t,h}) + \limsup_{h\downarrow 0} \eta(Y_{t,h} - \hat{Y}_{t,h}) \\
&\quad + \limsup_{h\downarrow 0} \sup_{f\in\alpha\mathbf{K}} \Big|\nu(\hat{Y}_{t,h}-f) - \eta(\hat{Y}_{t,h}-f)\Big| \\
&=: D_1(\alpha,t) + D_2(\alpha,t) + D_3(\alpha,t).
\end{aligned}$$

Assume now that $\nu\in\mathcal{N}$. It follows from (4.4) that, with probability 1, $D_1(\alpha,t) = D_2(\alpha,t) = 0$ for all $t\in[0,1]$. Moreover, by Lemma 4.1, for each fixed pair of norms $\eta,\nu\in\mathcal{C}$ with $\varepsilon(\eta,\nu)\leqq M(\eta)/12$, we have with probability 1, $D_3(\alpha,t)\leqq\Delta(\eta,\nu)$. Thus, for each $\eta\in\mathcal{C}$, we have with probability 1, for all $t\in[0,1]$, $\alpha\in[0,1)$ and $\nu\in\mathcal{N}$ such that $\varepsilon(\eta,\nu)\leqq M(\eta)/12$,

$$|\limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \nu(Y_{t,h}-f) - \limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \eta(Y_{t,h}-f)| \leqq \Delta(\eta,\nu). \tag{4.6}$$

The same arguments show that, on an event of probability 1, for all $t\in[0,1]$, $f\in\mathbf{K}$ and $\nu\in\mathcal{N}$ with $\varepsilon(\eta,\nu)\leqq M(\eta)/12$,

$$|\liminf_{h\downarrow 0} \nu(Y_{t,h}-f) - \liminf_{h\downarrow 0} \eta(Y_{t,h}-f)| \leqq \Delta(\eta,\nu). \tag{4.7}$$

Fix an arbitrary consistent norm $\eta\in\mathcal{C}$, and choose any $\alpha\in[0,1)$ and $t\in T(\eta,\alpha)$. By the definition (1.7) of the set $T(\eta,\alpha)$, we have

$$\limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \eta(Y_{t,h}-f) > 0.$$

Our assumption that $\mathcal{N}$ is dense in $\mathcal{C}$ with respect to the L^1-norm of the Wiener measure implies that there exists a norm $\nu \in \mathcal{N}$ such that $\varepsilon(\eta,\nu) \leqq M(\eta)/12$, and

$$\Delta(\eta,\nu) < \limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \eta(Y_{t,h}-f).$$

Moreover, (4.6) yields that, with probability 1,

$$\limsup_{h\to 0} \inf_{f\in\alpha\mathbf{K}} \nu(Y_{t,h}-f) > 0,$$

i.e. $t \in T(\nu,\alpha) \subset T''_\alpha$. Thus, by (4.1), the first inclusion in (1.17) is satisfied.

Let now $\alpha \in [0,1)$ and $t \in T'_\alpha$. By the definition (4.1) of the set T'_α there exists an $f' \in \mathbf{K}'$ such that $|f'|_{\mathbf{H}} > \alpha$ and $t \in T_{f'}$. In view of (1.7) and (4.1), this means that, for all $\nu \in \mathcal{N}$,

$$\liminf_{h\downarrow 0} \nu(Y_{t,h}-f') = 0, \tag{4.8}$$

whence, by (4.7),

$$\liminf_{h\downarrow 0} \eta(Y_{t,h}-f') \leqq \inf_{\nu\in\mathcal{N}:\varepsilon(\eta,\nu)\leqq M(\eta)/12} \Delta(\eta,\nu) = 0.$$

By (4.8) and the triangle inequality, this entails that, with probability 1,

$$\begin{aligned}\limsup_{h\to 0} \inf_{f\in\alpha\mathbf{K}} \eta(Y_{t,h}-f) &\geqq \limsup_{h\downarrow 0} \inf_{f\in\alpha\mathbf{K}} \{\eta(f'-f) - \eta(Y_{t,h}-f')\} \\ &\geqq \inf_{f\in\alpha\mathbf{K}} \{\eta(f'-f)\} - \liminf_{h\downarrow 0} \eta(Y_{t,h}-f') \\ &= \inf_{f\in\alpha\mathbf{K}} \{\eta(f'-f)\} > 0.\end{aligned}$$

Recalling (1.7), we see that this implies that $t \in T(\eta,\alpha)$. Since this holds with probability 1 uniformly over $t \in T'_\alpha$, we see that the second inclusion in (1.17) is satisfied.

Let now $f \in \mathbf{K}$ and $t \in T_f$. By the definition (4.1) of T_f, for all $\nu \in \mathcal{N}$,

$$\liminf_{h\downarrow 0} \nu(Y_{t,h}-f) = 0.$$

Hence, by (4.14)

$$\liminf_{h\to 0} \eta(Y_{t,h}-f) \leqq \inf_{\nu\in\mathcal{N}:\varepsilon(\eta,\nu)\leqq M(\eta)/12} \Delta(\eta,\nu) = 0.$$

It means that $t \in T(\eta,f)$. We have therefore completed the proof of (1.18). The proof of Theorem 1.3 is now complete.

Acknowledgement. We are greatly honored to be allowed to dedicate this paper to Endre Csáki, whose considerable achievements have remained a constant example and inspiration for our own research.

REFERENCES

[1] Borell, C., The Brunn–Minkowski inequality in Gauss space, *Invent. Math.* **30** (1975), 207–216. *MR* **53** #3246

[2] Bourbaki, N., *Éléments de mathématique*, Fasc. XV, Livre V: *Espaces vectoriels topologiques*, 2. édition, Actualités Sci. Industr., No. 1189, Hermann, Paris, 1966. *MR* **34** #3277

[3] Cameron, R. H. and Graves, R. E., Additive functionals on a space of continuous functions. I, *Trans. Amer. Math. Soc.* **70** (1951), 160–176. *MR* **12**, 718e

[4] Cameron, R. H. and Martin, W. T., Evaluation of various Wiener integrals by use of certain Sturm–Liouville differential equations, *Bull. Amer. Math. Soc.* **51** (1945), 73–90. *MR* **6**, 160c

[5] Deheuvels, P. and Lifshits, M. A., Strassen-type functional laws for strong topologies, *Probab. Theory Relat. Fields* **97** (1993), 151–167. *MR* **94i**:60048

[6] Deheuvels, P. and Lifshits, M. A., Necessary and sufficient conditions for the Strassen law of the iterated logarithm in nonuniform topologies, *Ann. of Probab.* **22** (1994), 1838–1856. *MR* **96d**:60051

[7] Deheuvels, P. and Mason, D. M., Random fractals generated by oscillations of processes with stationary and independent increments, *Probability in Banach spaces*, **9** (Sandjberg, 1993) (J. Hoffman-Jørgensen, J. Kuelbs, and M.B. Marcus, eds.) Progress in probability, **35**, Birkhäuser, Boston, 1994, 73–90. *MR* **96b**:60191

[8] Deheuvels, P. and Mason, D. M., On the fractal nature of empirical increments, *Ann. of Probab.* **23** (1995), 355–387. *MR* **96e**:60064

[9] Falconer, K. J., *The geometry of fractal sets*, Cambridge Tracts in Mathematics, **85**, Cambridge University Press, Cambridge, 1986. *MR* **88d**:28001

[10] Falconer, K. J., *Fractal geometry. Mathematical foundations and applications*, Wiley, Chichester–New York, 1990. *MR* **92j**:28008

[11] Jain, N. C., A zero-one law for Gaussian processes, *Proc. Amer. Math. Soc.* **29** (1971), 585–587. *MR* **43** #4099

[12] Kallianpur, G., Zero-one laws for Gaussian processes, *Trans. Amer. Math. Soc.* **149** (1970), 199–211. *MR* **42** #1200

[13] Ledoux, M. and Talagrand, M., *Probability in Banach spaces. Isoperimetry and processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **23**, Springer-Verlag, Berlin, 1991. *MR* **93c**:60001

[14] Lévy, P., *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937. *Zbl* **16**, 170

[15] Lifshits, M. A., *Gaussian random functions*, Kluwer, Dordrecht, 1995.

[16] Mueller, C., A unification of Strassen's law and Levy's modulus of continuity, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **56** (1981), 163–179. *MR* **82f**:60180

[17] Orey, S. and Taylor, J., How often on a Brownian path does the law of the iterated logarithm fail?, *Proc. London Math. Soc.* (3) **28** (1974), 174–192. *MR* **50** #11486

[18] Révész, P., A generalization of Strassen's functional law of iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **50** (1979), 257–264. *MR* **82c**:60057

[19] Robertson, A. P. and Robertson, W., *Topological vector spaces*, 2. edition, Cambridge tracts in mathematics and mathematical physics, No. 53,Cambridge University Press, London–New York, 1973. *MR* **50** #2854

[20] Stoyan, D. and Stoyan, H., *Fractals, random shapes and point fields*, Wiley series in probability and mathematical statistics: Applied probability and statistics, Wiley, Chichester–New York, 1994. *MR* **95h**:60016

[21] Strassen, V., An invariance principle for the law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **3** (1964), 211-226. *MR* **30** #5379

[22] SUDAKOV, V. N. and TSYRELSON, B. S., Extremal properties of half-spaces for spherically invariant measures, Problems in the theory of probability distributions, II, *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)* **41** (1974), 14–24. *MR* **51** #1932; *J. Soviet Math.* **9** (1974), 9–18.

(Received January 20, 1997)

UNIVERSITÉ PIERRE ET MARIE CURIE
LABORATOIRE DE STATISTIQUE THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
4, PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

pd@ccr.jussieu.fr

MANCOMTECH TRAINING CENTER
22-2-49 KOMENDANTSKII PROSPEKT
SAINT-PETERSBURG 197372
RUSSIA

mikhail@lifshits.spb.su

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 111–116

ON THE BEST APPROXIMATING ELLIPSE CONTAINING A PLANE CONVEX BODY

P. ERDŐS, E. MAKAI, Jr.* and I. VINCZE**

Dedicated to Endre Csáki on his sixtieth birthday

In the paper [EV] by P. Erdős and I. Vincze, § 6, the following statement is proved:

THEOREM 1. *Let K be a convex body in the plane and let us consider the set $\mathcal{E}$ of ellipses containing K along with the Blaschke distances of the elements of $\mathcal{E}$ from K. Then there exists a unique element of $\mathcal{E}$ having a minimal Blaschke distance from K.*

This theorem with its proof was published in Hungarian in [EV]. The authors give below the proof in English, applying a slight modification of the original version.

In our paper we turn also to a geometric characterization of this unique ellipse closest in Blaschke's sense to K. This question was raised as a problem in [EV].

THEOREM 2. *Under the hypotheses of Theorem 1, the boundary of the unique closest ellipse $E \in \mathcal{E}$ has at least three points in common with K, say, A_1, A_2, A_3, and also has at least three points having maximal distance from K, say, B_1, B_2, B_3, in such a way that they lie alternately on the boundary of E: A_1, B_1, A_2, B_2, A_3, B_3. Conversely, an ellipse $E \in \mathcal{E}$ having this property is identical with the unique ellipse in $\mathcal{E}$ having minimal distance from K.*

For basic facts on convex bodies we refer to [BF].

In the following we turn to the proofs of our statements. At the end of the paper we will point out how Theorem 1 follows also from Theorem 2.

PROOF OF THEOREM 1. An easy compactness argument shows the existence of an ellipse $E \in \mathcal{E}$ closest to K, which, of course, cannot be degenerate.

Suppose there exist more than one element of $\mathcal{E}$ having the same minimal (Blaschke) distance from K, which distance will be denoted by d. Consid-

---

1991 *Mathematics Subject Classification.* Primary 52A10; Secondary 52A27.

Key words and phrases. Planar convex body, ellipse, approximation, Blaschke-distance, alternation property.

*Research (partially) supported by Hungarian National Foundation for Scientific Research Grant No. T 016387.

**Research (partially) supported by Hungarian National Foundation for Scientific Research Grant No. T 016384.

ering two such ellipses, both of them are contained by the parallel body K_d and their convex hull is also contained by K_d. This convex hull also has the distance d from K.

Let C be a common interior point of the two ellipses, and let the radial functions of the two ellipses with respect to C be r_1, r_2. Then for the numbers i, j of the simple zeros and multiple zeros of $r_1 - r_2$ we have $i + 2j \leqq 4$. We will make our discussion according to the number of zeros and sign changes of $r_1 - r_2$. We speak of an *intersection* of the ellipses if $r_1 - r_2$ has a zero and there is a sign change there, a *simple intersection* if $r_1 - r_2$ has a single zero there, and a *non-intersectional common point* if $r_1 - r_2$ has a zero and there is no sign change there.

a) Let us consider the case when the two ellipses have four different points of simple intersection. In this case the body K is contained in their intersection. Each of the two ellipses has two parts outside their intersection. Let us draw a line through one of the four points, which lies outside the intersection and does not touch any of the two ellipses. We shall now consider that further ellipse whose boundary passes through the four common points and touches the mentioned line. This ellipse is determined uniquely, lies in the interior of the union of the two ellipses, except the four common points of the boundaries of the two ellipses. (Analytically, if the ellipses are given by the quadratic inequalities $f_1 \geqq 0$, $f_2 \geqq 0$, then we consider an ellipse given by $\lambda f_1 + (1-\lambda) f_2 \geqq 0$, for some $0 < \lambda < 1$.) In this way we have obtained an ellipse which contains K, but does not have a common point with the boundary of K_d. This ellipse has a smaller distance to K, which is a contradiction.

b) Let us consider that case when the two ellipses touch each other and have two other points of simple intersection. If the touching point A has a smaller distance from K than d, then our above procedure leads to the stated result: That ellipse E, which has a tangent at A, coinciding with the tangent l of the two ellipses at A, and whose boundary passes through the two other points and at one of the two points has as tangent a line as described in case a), contains K and has from it a distance smaller than d. (We can give this ellipse also analytically, like above.)

Now suppose A has a distance d from K. Then E lies in the interior of K_d, except for A, and l is tangent to E at A. Let l' be a translate of l, close to l, intersecting E in a small chord BC. Then for a small $\varepsilon > 0$ there is a homothetic copy E_ε of E, close to E, with ratio of homothety $1 + \varepsilon$, also having BC as a chord, such that the open large (small) arc BC of the boundary of E lies in the interior (exterior) of E_ε. (By affine invariance it suffices to show this for E a circle, where this is immediate.) In particular, $A \notin E_\varepsilon$ and, for l' sufficiently close to l, and ε sufficiently small, E_ε contains K and lies in the interior of K_d. Thus the distance of E_ε from K is smaller than d, a contradiction.

c) If the two ellipses have two points of intersection, B, C, say, then consider the analytically given ellipse E, like in a). If both at B and C the tangents of the two ellipses are different, then E lies in the interior of K_d, a contradiction. If at one of B and C the tangents are different, then we proceed with E as in the second case considered in b). If both at B and C the tangents of the two ellipses coincide, then a short calculation shows that the ellipses have non-intersectional common points, a contradiction. (By affine invariance, it suffices to perform this calculation for the cases $B=(-1,1)$, $C=(1,1)$, the tangents being $y=-x$, $y=x$, and $B=(-1,0)$, $C=(1,0)$, the tangents being $x=-1$, $x=1$, and in both cases considering the boundary points on the y-axis.)

d) If the two ellipses E_1, E_2 satisfy that say E_1 lies in the interior of E_2, then E_1 has a smaller distance from K than d, a contradiction.

e) If the two ellipses E_1, E_2 have one non-intersectional common point A, and, say, $E_1 \subset E_2$, then E_1 lies in the interior of K_d, except for A. Then we proceed like in the second case considered in b).

f) If the two ellipses E_1, E_2 have two non-intersectional common points A_1, A_2, and, say, $E_1 \subset E_2$, then E_1 lies in the interior of K_d, except for A_1, A_2. If one of A_1, A_2 lies in the interior of K_d, we have the cases considered in d), e). If both A_1, A_2 have a distance d to K, then consider the tangents l_1, l_2 of E_1 at A_1, A_2. Let l_1', l_2' be translates of l_1, l_2, close to l_1, l_2, intersecting E_1 in small chords B_1C_1, B_2C_2, such that $\overline{B_1C_1}/\overline{B_1'C_1'} = \overline{B_2C_2}/\overline{B_2'C_2'}$, where $B_1'C_1'$, $B_2'C_2'$ are the (affine) diameters of E_1, parallel to B_1C_1, B_2C_2. Now let $\varepsilon > 0$ be small, and A_1^*, A_2^* the points on the segments OA_1, OA_2 satisfying $\overline{OA_1^*}/\overline{OA_1} = \overline{OA_2^*}/\overline{OA_2} = 1-\varepsilon$, where O is the centre of E_1. Then there is an ellipse E^* whose boundary passes through A_1^*, B_1, C_1, B_2, C_2. Moreover, it passes through A_2^*, too. (By affine invariance, it suffices to show this for E_1 a circle, where it follows from a symmetry consideration, for a conic passing through B_1, C_1, B_2, C_2 and a point of the symmetry axis.) Then O lies in the interior of E^*, A_1, A_2 lie in the exterior of E^*, and for l_1', l_2' sufficiently close to l_1, l_2, and ε sufficiently small, E^* contains K and is contained in the interior of K_d, a contradiction.

Having checked all possible relative positions of our two ellipses (that have common interior points), the theorem is proved. □

In this way we have proved the uniqueness of the ellipse which contains a planar convex body and has minimal Blaschke distance from it. Now we turn to the geometric characterization of this ellipse, that is analogous to the Bonnesen characterization of the minimal circular ring of a closed convex plane curve, cf. [B], p. 487, [BF], pp. 54–55.

The Bonnesen characterization is the following. The boundary of a plane convex body K can be covered by a circular ring bounded by two concentric circles, of radii $R \geq r$, such that $R-r$ is minimal under the above conditions. This minimal circular ring is unique. If the common centre of the circles is O, there are four directed segments OA_1, OB_1, OA_2, OB_2, their directions following each other in the above order, such that

$$A_1, B_1, A_2, B_2 \in \operatorname{bd} K, \qquad \overline{OA_1} = \overline{OA_2} = r, \qquad \overline{OB_1} = \overline{OB_2} = R.$$

Conversely, this alternation property characterizes the minimal circular ring. Still we note that the outer circle, inner circle and mid-circle (of centre O, and radius $(r+R)/2$) of the minimal circular ring are the best outer, inner approximating circles of K, and the best approximating circle of K, respectively, in the sense of Blaschke distance.

PROOF OF THEOREM 2. The question will be handled by an analogue of the Chebyshev approximation method ([N], Ch. II, § 2, Ch. III, § 4).

(1) Let E be the ellipse containing K with the distance $d(K,E)$ minimal. We will show the alternation property of Theorem 2.

Let the support functions of E and K be $h_E \geqq h_K$. If for some $\alpha \in S^1$ (the unit circle in $\mathbb{R}^2$) we have $h_E(\alpha) = h_K(\alpha)$, then the boundary point A of E, having $\langle A, (\cos\alpha, \sin\alpha)\rangle$ maximal, is a common boundary point of K and E. Recall that the Blaschke distance $d(K_1, K_2)$ of two plane convex bodies K_1, K_2, with support functions h_1, h_2 equals $\max_\varphi |h_1(\varphi) - h_2(\varphi)|$. Thus for the above ellipse E we also have $h_E \leqq h_K + d(K,E)$. If for some $\beta \in S^1$ we have $h_E(\beta) = h_K(\beta) + d(K,E)$, then the boundary point B of E, having $\langle B, (\cos\beta, \sin\beta)\rangle$ maximal, satisfies that $B - (\cos\beta, \sin\beta)d(K,E) \in \operatorname{bd} K$, and B lies at a distance $d(K,E)$ to K. Therefore, in order to show the alternation property in the theorem, it suffices to prove that there are $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3 \in S^1$, following each other in this cyclic order, such that α_i is a minimum point of $h_E - h_K$, and β_i is a maximum point of $h_E - h_K$ $(i = 1, 2, 3)$.

To prove our statement it suffices to investigate the case $E \neq K$. Then we have

$$\max\,(h_E(\varphi) - h_K(\varphi)) > 0.$$

Now we consider the set of minimum and maximum points of the function $h_E - h_K$. If $\alpha \in S^1$ is a minimum point, then consider a maximal arc $[\alpha', \alpha''] \subset S^1$ such that $\alpha \in [\alpha', \alpha'']$, α', α'' are minimum points and $[\alpha', \alpha'']$ does not contain any maximum points. We admit the case that $[\alpha', \alpha'']$ reduces to the singleton $\{\alpha\}$. Dually, for β a maximum point we consider maximal arcs $[\beta', \beta''] \subset S^1$, with the dual properties. These arcs are disjoint. If $[\alpha', \alpha'']$ is such a maximal arc of the first type, then the extremum of $h_E - h_K$, greater than α'' (in the cyclic sense) and nearest to it, is a maximum. The dual statement is valid for the maximal arcs $[\beta', \beta'']$ of the second type. Hence the arcs of type $[\alpha', \alpha'']$ and those of type $[\beta', \beta'']$ follow each other alternately. If there follow three intervals of type $[\alpha', \alpha'']$ and three intervals of type $[\beta', \beta'']$ each other, we are done.

Now let us suppose that there are at most two intervals of type $[\alpha', \alpha'']$, and at most two intervals of type $[\beta', \beta'']$. In this case there are values $\varphi_1, \varphi_2, \varphi_3, \varphi_4 \in S^1$, following each other in this cyclic order, such that the intervals of type $[\alpha', \alpha'']$ lie in the open arcs (φ_1, φ_2) and (φ_3, φ_4), and those of type $[\beta', \beta'']$ lie in the open arcs (φ_2, φ_3) and (φ_4, φ_1). (Not each of these arcs $(\varphi_i, \varphi_{i+1})$ needs to contain an interval of type $[\alpha', \alpha'']$ or $[\beta', \beta'']$.)

Let us consider the polar E° of E with respect to an interior point O of E. Then E° is an ellipse, and the polars of the support lines of E with outer normals $(\cos\varphi_i, \sin\varphi_i)$ are points P_i of $\operatorname{bd} E^\circ$. Let us consider an ellipse E_1° whose boundary passes through $P_1, \ldots, P_4$ and through a point P close to a fixed point Q on the open arc P_1P_2 of E°, P being either inside, or outside E°. Then $\operatorname{bd} E^\circ$ and $\operatorname{bd} E_1^\circ$ have only these points P_i in common, since five points uniquely determine an ellipse. Moreover, the arcs (P_i, P_{i+1}) of $\operatorname{bd} E_1^\circ$ pass alternately inside or outside the arcs (P_i, P_{i+1}) of $\operatorname{bd} E^\circ$ (i.e., of E°). In fact, else, at some P_i, E° and E_1° would have a common tangent, implying $E^\circ = E_1^\circ$. Moreover, one can prescribe that the arc (P_1, P_2) of $\operatorname{bd} E_1^\circ$ should pass inside or outside of the arc (P_1, P_2) of $\operatorname{bd} E^\circ$.

Letting E_1 be the polar of E_1° with respect to O, E_1 is also an ellipse (for P sufficiently close to Q). For the support functions h_E and h_{E_1} of E and E_1 we have $h_{E_1}(\varphi_i) = h_E(\varphi_i)$, and $h_{E_1}(\varphi) - h_E(\varphi)$ is alternately positive and negative in the open intervals $(\varphi_i, \varphi_{i+1})$, and we can prescribe its sign in (φ_1, φ_2). So we may suppose $h_{E_1}(\varphi) - h_E(\varphi)$ positive in (φ_1, φ_2) and (φ_3, φ_4), and negative in (φ_2, φ_3) and (φ_4, φ_1).

Hence, recalling the choice of φ_i, we have, for some $\varepsilon_0 > 0$, $h_{E_1}(\varphi) \geqq h_K(\varphi) + \varepsilon_0$ in $[\varphi_1, \varphi_2]$ and $[\varphi_3, \varphi_4]$, and $h_{E_1}(\varphi) \leqq h_K(\varphi) + d(K, E) - \varepsilon_0$ in $[\varphi_2, \varphi_3]$ and $[\varphi_4, \varphi_1]$. For P sufficiently close to Q we have also that $h_{E_1}(\varphi) \leqq h_K(\varphi) + d(K, E) - \varepsilon_0$ in $[\varphi_1, \varphi_2]$ and $[\varphi_3, \varphi_4]$, and $h_{E_1}(\varphi) \geqq h_K(\varphi) + \varepsilon_0$ in $[\varphi_2, \varphi_3]$ and $[\varphi_4, \varphi_1]$. In conclusion, $h_K \leqq h_K + \varepsilon_0 \leqq h_{E_1} \leqq h_K + d(K, E) - \varepsilon_0$, implying $K \subset E_1$ and $d(K, E_1) = \max_\varphi (h_{E_1}(\varphi) - h_K(\varphi)) \leqq d(K, E) - \varepsilon_0$, contradicting the choice of E. Hence the best approximating ellipse E satisfies the alternation property.

(2) Now we show that the alternation property implies the best approximation property. So let the ellipse $E \supset K$ satisfy the alternation property, where again we may suppose $E \neq K$. We have to prove that there does not exist an ellipse $E' \supset K$, such that $d(K, E') < d(K, E)$.

Let α_i and β_i denote the angle of the outer normal of E at A_i and B_i. These angles have the cyclic order α_1, β_1, α_2, β_2, α_3, β_3. We have $h_E(\alpha_i) = h_K(\alpha_i)$. Let C_i be the point of K closest to B_i. Then K has a supporting line at C_i that is perpendicular to C_iB_i. Moreover we have $d(C_i, B_i) = d(K, E)$, and E is contained in the $d(K, E)$-neighbourhood of K, so the tangent of E at B_i is also perpendicular to B_iC_i. These imply $h_E(\beta_i) = h_K(\beta_i) + d(K, E)$.

Now suppose that there exists an ellipse $E' \supset K$ such that $d(K, E') < d(K, E)$. Let the support function of E' be $h_{E'}$. Then $h_K \leqq h_{E'} \leqq h_K + d(K, E') < h_K + d(K, E)$. By inflating E' a bit about its centre, we may even suppose that $h_K < h_{E'} < h_K + d(K, E)$. Also we have $h_K \leqq h_E \leqq h_K + d(K, E)$.

Consider the arc $[\alpha_1, \beta_1]$. We have

$$h_{E'}(\alpha_1) - h_E(\alpha_1) > 0, \qquad h_{E'}(\beta_1) - h_E(\beta_1) < 0.$$

Hence $h_{E'} - h_E$ has a zero in (α_1, β_1). Similarly we proceed for the other arcs

$(\beta_1, \alpha_2), \ldots, (\beta_3, \alpha_1)$. So $h_{E'} - h_E$ has six zeros. Choosing a suitable centre of polarity, the polars of E' and E will be ellipses having six common points, therefore they must coincide. Hence $E' = E$, $d(K, E') = d(K, E)$, contrary to the choice of E'. This proves that the alternation property implies the best approximation property, and thus finishes the proof of the theorem. □

REMARKS. 1. In fact the characterization of the best approximating ellipse $E \supset K$ in Theorem 2 implies its unicity, i.e., Theorem 1, analogously to [N], Ch. II, § 2, Theorems 3, 4, Ch. III, § 4, Theorems 3, 4. Namely, if $E_1, E_2 \supset K$ are both best approximating ellipses, and are given by quadratic inequalities $f_1 \geqq 0$, $f_2 \geqq 0$, then for $0 < \lambda < 1$ we can define an ellipse E by $\lambda f_1 + (1 - \lambda) f_2 \geqq 0$. We have $K \subset E \subset E_1 \cup E_2 \subset K_d$ for $d = d(K, E_1) = d(K, E_2)$, so E is also a best approximating ellipse containing K. So there exist A_1, B_1, A_2, B_2, A_3, B_3 satisfying the alternation property in Theorem 2. Moreover, E lies in $\operatorname{int} K_d$, except for the common points of $\operatorname{bd} E_1$ and $\operatorname{bd} E_2$. By $d(K, B_i) = d$ each point B_i is a common point of $\operatorname{bd} E_1$ and $\operatorname{bd} E_2$. Then $E_1, E_2 \subset K_d$ implies that at B_i both E_1 and E_2 have the same tangent, for $i = 1, 2, 3$. Hence $E_1 = E_2$.

2. An analogous property by three alternating minimum and maximum points $\alpha_1, \ldots, \beta_3$ could be proved for the best inner Hausdorff approximation by ellipses, and for the best Hausdorff approximation by ellipses, provided we knew that these are not degenerating to segments. Conversely, the respective alternation property implies the best inner, or best approximation property.

REFERENCES

[B] BONNESEN, T., Über das isoperimetrische Defizit ebener Figuren, *Math Ann.* **91** (1924), 252–268. *Jb. Fortschr. Math.* **50** (1924), 487

[BF] BONNESEN, T. and FENCHEL, W., *Theorie der konvexen Körper*, Berichtigter Reprint, Springer, Berlin–Heidelberg–New York, 1974. *MR* **49** #9736

[EV] ERDŐS, P. and VINCZE, I., On the approximation of closed convex plane curves, *Mat. Lapok* **9** (1958), 19–36 (in Hungarian, summaries in Russian and German). *MR* **20** #6070

[N] NATANSON, I. P., *Konstruktive Funktionentheorie*, Akademie-Verlag, Berlin, 1955. *MR* **16**-1100

(Received February 4, 1997)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

makai@math-inst.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 117–125

ON THE RADIUS OF THE LARGEST BALL LEFT EMPTY BY A WIENER PROCESS

P. ERDŐS and P. RÉVÉSZ

Dedicated to E. Csáki for his sixtieth birthday

1. Introduction

Let $\{W(t)=(W_1(t),W_2(t),\dots,W_d(t))\in\mathbb{R}^d,\ t\geqq 0\}$ be a Wiener process in the d-dimensional Euclidean space where $d\geqq 3$ and let

$$\mathcal{C}(x,r)=\{y:\ y\in\mathbb{R}^d,\ \|y-x\|\leqq r\}.$$

Consider a path $W(t,\omega)$ of $W(t)$. We say that $\mathcal{C}(x,r)$ is left empty by $W(t,\omega)$ if

$$\mathcal{D}(x,r)=\mathcal{D}(x,r,\omega)=\mathcal{C}(x,r)\cap\{W(t,\omega),t\geqq 0\}=\emptyset.$$

Let

$$\rho(R)=\rho(R,\omega)=$$
$$=\max\{r:\ \exists x\in\mathbb{R}^d \text{ such that } \mathcal{C}(x,r)\subset\mathcal{C}(0,R) \text{ and } \mathcal{D}(x,r)=\emptyset\}$$

be the radius of the largest empty ball in $\mathcal{C}(0,R)$. We are interested in studying the properties of the stochastic process $\{\rho(R),\ R\geqq 0\}$.

Since $W(0)=0$, clearly

$$\rho(R)\leqq\frac{R}{2}. \tag{1}$$

First we give a sharper upper bound than the trivial one of (1). In fact we prove

THEOREM 1. *For any* $\varepsilon>0$,

$$\rho(R)\leqq\frac{R}{2}-R^{1-\varepsilon}\qquad a.s.. \tag{2}$$

if R is big enough.

Our next Theorem tells us that the upper bound of (2) is not very far from the best possible result.

1991 *Mathematics Subject Classification.* Primary 60J65; Secondary 60G17.
Key words and phrases. Brownian motion, path properties, strong theorems.

0081-6906/97/$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

THEOREM 2.

$$\rho(R) \geqq \frac{R}{2} - \frac{R}{4\log R} \qquad i.o.\ a.s.. \tag{3}$$

Here and similarly in the sequel i.o. a.s. (infinitely often almost surely) means that for almost all $\omega \in \Omega$ there exists a sequence $0 < R_1 = R_1(\omega) < R_2 = R_2(\omega) < \dots$ such that

$$\lim_{n\to\infty} R_n(\omega) = \infty$$

and

$$\rho(R_n) \geqq \frac{R_n}{2} - \frac{R_n}{4\log R_n}.$$

Theorem 2 tells us that for some R the $\rho(R)$ will be very big. The next Theorem tells that for some R the $\rho(R)$ will be much smaller.

THEOREM 3. *For any $\varepsilon > 0$ we have*

$$\rho(R) \leqq \frac{R}{(\log\log R)^{1/d-\varepsilon}} \qquad i.o.\ a.s.. \tag{4}$$

Now we show that the upper bound of (4) is close to the best posssible result.

THEOREM 4. *For any $\varepsilon > 0$*

$$\rho(R) \geqq \frac{R}{(\log R)^{(1+\varepsilon)/(d-3)}} \qquad a.s. \tag{5}$$

if R is big enough and $d \geqq 4$. Further

$$\rho(R) \geqq R(\log R)^{-(1+\varepsilon)} \qquad a.s.$$

if R is big enough and $d = 3$.

The proof of Theorem 1 is based on a theorem that seems to be interesting in itself. In order to formulate it we introduce a few notations.

For any $x \in \mathbb{R}^d$ with $\|x\| = 1$ and $0 < \vartheta < 1$ define the cone $\mathcal{K}(x,\vartheta)$ as follows:

$$\mathcal{K}(x,\vartheta) = \left\{ y:\ y \in \mathbb{R}^d,\ \left(\frac{y}{\|y\|}, x\right) \geqq 1 - \vartheta \right\}.$$

Clearly for any $0 < \vartheta < 1$ there exists a positive integer $K = K(\vartheta)$ and a sequence $x_1, x_2, \dots, x_K$ such that

$$x_i \in \mathbb{R}^d, \qquad \|x_i\| = 1, \qquad (i = 1, 2, \dots, K)$$

$$\bigcup_{i=1}^{K} \mathcal{K}(x_i, \vartheta) = \mathbb{R}^d, \qquad K \leqq L(1-(1-\vartheta)^2)^{-(d-1)/2}$$

where L is an absolute positive constant.

Let

$$\mathcal{L}_i = \mathcal{L}_i(R) = \mathcal{L}_i(R, \varepsilon, \vartheta) = \{y : y \in \mathcal{K}(x_i, \vartheta),\ R^{\varepsilon} \leqq (y, x_i) \leqq R^{1-\varepsilon}\}$$

where

$$i = 1, 2, \ldots, K, \qquad 0 < \varepsilon < 1/2, \qquad R > 0.$$

Now we have

THEOREM 5. *For any* $0 < \varepsilon < 1/2,\ 1/2 < \vartheta < 1,$

$$\mathbf{P}\left\{\limsup_{R\to\infty} \bigcup_{i=1}^{K} \{\mathcal{L}_i \cap \{W(t,\omega),\ t \geqq 0\} = \emptyset\}\right\} = 0.$$

Note that Theorem 5 tells us that for any R big enough $W(t)$ meets all frustum of cones $\mathcal{L}_i(R)$ $(i = 1, 2, \ldots, K)$.

Remember that

$$\limsup_{R\to\infty} A_R = \bigcap_{R<\infty} \bigcup_{s>R} A_s.$$

2. Proofs of Theorems 1 and 5

LEMMA 1. *Let* $\{W(t) \in \mathbb{R}^1, t \geqq 0\}$ *be a Wiener process. Then*

$$\mathbf{P}\{\sup_{T^{\varepsilon} \leqq t \leqq T^{1-\varepsilon}} t^{-1/2} W(t) \leqq (2\delta \log\log T)^{1/2}\} \leqq \exp(-(\log T)^{1-2\delta})$$

for any $0 < \varepsilon < 1/2,\ 0 < \delta < 1/2,$ *if* T *is big enough.*

PROOF. Let

$$t_k = \exp(k \log k), \qquad (k = 1, 2, \ldots)$$
$$k_1 = k_1(\varepsilon, T) = \min\{k:\ t_k \geqq T^{\varepsilon}\},$$
$$k_2 = k_2(\varepsilon, T) = \max\{k:\ t_k \leqq T^{1-\varepsilon}\}.$$

Then

$$k_2 - k_1 \sim \frac{(1-2\varepsilon)\log T}{\log\log T},$$
$$\frac{t_k}{t_{k+1}} \sim \frac{1}{k},$$

$$\mathbf{P}\{t_{k+1}^{-1/2}(W(t_{k+1})-W(t_k))\geqq(2\delta\log\log T)^{1/2}\}\geqq(\log T)^{-\delta}(\log\log T)^{-1}$$

and

$$\mathbf{P}\{\max_{k_1\leqq k\leqq k_2} t_{k+1}^{-1/2}(W(t_{k+1})-W(t_k))<(2\delta\log\log T)^{1/2}\}\leqq$$

$$\leqq(1-(\log T)^{-\delta}(\log\log T)^{-1})^{k_2-k_1}\sim\exp\left(-\frac{(1-2\varepsilon)(\log T)^{1-\delta}}{(\log\log T)^2}\right).$$

Note that if

$$t_{k+1}^{-1/2}(W(t_{k+1})-W(t_k))>(2\delta\log\log T)^{1/2}$$

and

$$t_{k+1}^{-1/2}W(t_{k+1})<(\delta\log\log T)^{1/2}$$

then

$$(2\delta\log\log T)^{1/2}<t_{k+1}^{-1/2}W(t_{k+1})-t_k^{-1/2}\left(\frac{t_k}{t_{k+1}}\right)^{1/2}W(t_k)<$$

$$<(\delta\log\log T)^{1/2}-t_k^{-1/2}\left(\frac{t_k}{t_{k+1}}\right)^{1/2}W(t_k)$$

and

$$t_k^{-1/2}W(t_k)<-(2^{1/2}-1)\left(\frac{t_{k+1}}{t_k}\right)^{1/2}(\delta\log\log T)^{1/2}\sim$$

$$\sim-(2^{1/2}-1)k^{1/2}(\log\log T)^{1/2}.$$

Consequently

$$\{\forall k:\ k_1\leqq k\leqq k_2, t_{k+1}^{-1/2}(W(t_{k+1})-W(t_k))\leqq(2\delta\log\log T)^{1/2}\}\cup$$

$$\cup\{\exists k:\ k_1\leqq k\leqq k_2, t_k^{-1/2}W(t_k)\leqq-(2^{1/2}-1)k^{1/2}(\delta\log\log T)^{1/2}\}\supset$$

$$\supset\{\forall k:\ k_1\leqq k\leqq k_2, t_{k+1}^{-1/2}W(t_{k+1})\leqq(\delta\log\log T)^{1/2}\}.$$

Since

$$\mathbf{P}\{t_k^{-1/2}W(t_k)<-(2^{1/2}-1)k^{1/2}(\delta\log\log T)^{1/2}\}\leqq$$

$$\leqq\exp\left(-\frac{(2^{1/2}-1)^2}{2}k\delta\log\log T\right),$$

we have

$$\mathbf{P}\{\inf_{k_1\leqq k\leqq k_2} t_k^{-1/2}W(t_k)<-(2^{1/2}-1)k^{1/2}(\delta\log\log T)^{1/2}\}\leqq$$

$$\leqq(k_2-k_1)\exp\left(-\frac{(2^{1/2}-1)^2}{2}k_1\delta\log\log T\right)\sim$$

$$\sim\frac{(1-2\varepsilon)\log T}{\log\log T}\exp\left(-\frac{(2^{1/2}-1)^2}{2}\varepsilon\delta\log T\right)\leqq T^{-\alpha},$$

where

$$\alpha = \frac{\varepsilon\delta}{4}(2^{1/2}-1)^2$$

and

$$\begin{aligned}
&\mathbf{P}\{\sup_{T^\varepsilon \leqq t \leqq T^{1-\varepsilon}} t^{-1/2} W(t) \leqq (\delta \log\log T)^{1/2}\} \leqq \\
&\leqq \mathbf{P}\{\sup_{k_1 \leqq k \leqq k_2} t_{k+1}^{-1/2} W(t_{k+1}) \leqq (\delta \log\log T)^{1/2}\} \leqq \\
&\leqq \mathbf{P}\{\sup_{k_1 \leqq k \leqq k_2} t_{k+1}^{-1/2} (W(t_{k+1}) - W(t_k)) \leqq (2\delta \log\log T)^{1/2}\} + \\
&\qquad + \mathbf{P}\{\inf_{k_1 \leqq k \leqq k_2} t_k^{-1/2} W(t_k) < -(2^{1/2}-1) k^{1/2} (\delta \log\log T)^{1/2}\} \leqq \\
&\leqq \exp\left(-\frac{(1-2\varepsilon)(\log T)^{1-\delta}}{(\log\log T)^2}\right) + T^{-\alpha} \leqq \exp(-(\log T)^{1-2\delta}).
\end{aligned}$$

Hence we have Lemma 1.

LEMMA 2. *Let*

$$\{W(t) = (W_1(t), W_2(t), \ldots, W_d(t)) \in \mathbb{R}^d,\ t \geqq 0\}$$

be a Wiener process. Then for any

$$0 < \varepsilon < \frac{1}{2}, \qquad 1/2 < \vartheta < 1,$$

and R big enough we have

$$\mathbf{P}\{\exists t:\ t \geqq 0,\ W(t) \in \mathcal{L}(R)\} \geqq 1 - C(\log R)^{-3/2}$$

where

$$\begin{aligned}
\mathcal{L}(R) &= \{y = (y_1, y_2, \ldots, y_d):\ y \in \mathcal{K}(e_1, \vartheta),\ R^\varepsilon \leqq (y, e_1) \leqq R^{1-\varepsilon}\} = \\
&= \{y:\ R^\varepsilon \leqq y_1 \leqq R^{1-\varepsilon},\ y_1 \geqq \Theta \|(y_2, \ldots, y_d)\|\}, \\
e_1 &= (1, 0, 0, \ldots, 0) \in \mathbb{R}^d, \\
\Theta &= \frac{1-\vartheta}{(1-(1-\vartheta)^2)^{1/2}}
\end{aligned}$$

and C is a positive constant.

PROOF. Let

$$\begin{aligned}
&T = R^2, \\
&t_0 = t_0(T, \varepsilon, \delta) = \inf\{t:\ t \geqq T^\varepsilon,\ W_1(t) \geqq (2\delta t \log\log t)^{1/2} = r\}, \\
&r = (2\delta t_0 \log\log t_0)^{1/2}.
\end{aligned}$$

By Lemma 1

$$\mathbf{P}\{t_0 \leqq T^{1-\varepsilon}\} \geqq 1 - \exp(-(\log T)^{1-2\delta})$$

if $0 < \delta < 1/2$. Hence

$$\mathbf{P}\{R^{\varepsilon/2} \leqq r \leqq R^{1-\varepsilon/2}\} \geqq 1 - \exp(-(2\log R)^{1-2\delta}).$$

Observe that

$$\mathbf{P}\{\|(W_2(t_0),\dots,W_d(t_0))\| \geqq (3t_0 \log\log t_0)^{1/2} = \left(\frac{3}{2\delta}\right)^{1/2} r\} \leqq$$

$$\leqq \exp\left(-\frac{3}{2}\log\log t_0\right) \leqq \exp\left(-\frac{3}{2}\log\log T^{\varepsilon}\right) = (2\varepsilon\log R)^{-3/2},$$

$$\mathbf{P}\{\exists t:\ t \geqq 0,\ W(t) \in \mathcal{L}(R)\} =$$

$$= \mathbf{P}\{\exists t:\ t \geqq 0,\ R^{\varepsilon} \leqq W_1(t) \leqq R^{1-\varepsilon},\ \Theta\|(W_2(t),\dots,W_d(t))\| \leqq W_1(t)\} \geqq$$

$$\geqq (1-(2\varepsilon\log R)^{-3/2})(1-\exp(-2\log R)^{1-2\delta}))$$

if

$$\Theta^2 < \frac{2\delta}{3},$$

i.e. if $\vartheta > 1/2$ and $\delta < 1/2$ is close enough to $1/2$.

Hence we have Lemma 2.

LEMMA 3. *By the conditions of Theorem 3 we have*

$$\mathbf{P}\left\{\bigcup_{i=1}^{K}\{\mathcal{L}_i \cap \{W(t,\omega),\ t \geqq 0\} = \emptyset\}\right\} < C(\log R)^{-3/2},$$

where C is a positive constant depending only on ϑ.

PROOF. Lemma 3 is a trivial consequence of Lemma 2.

PROOF OF THEOREM 5. Let $R_k = e^k$ $(k = 1, 2, \dots)$. By Lemma 3 we have

$$\mathbf{P}\left\{\limsup_{k\to\infty}\bigcup_{i=1}^{K}\{\mathcal{L}_i(R_k) \cap \{W(t,\omega),\ t \geqq 0\} = \emptyset\}\right\} = 0.$$

Let $R_k \leqq R < R_{k+1}$. Then

$$R^{\varepsilon/2} \leqq R_k^{\varepsilon} \quad \text{and} \quad R^{1-\varepsilon/2} \geqq R_{k+1}^{1-\varepsilon}$$

which, in turn, implies Theorem 5.

PROOF OF THEOREM 1. It is a trivial consequence of Theorem 5.

3. Proof of Theorem 2

First we recall two known Theorems.

THEOREM OF HIRSCH ([1], p. 39).

$$\inf_{t\leqq T} W_1(t) \geqq -T^{1/2}(\log T)^{-1} \qquad i.o.\ a.s..$$

THEOREM A ([1], p. 242).

$$\mathbf{P}\{\exists t \geqq 0 : W(t) \in C(u,r)\} = \left(\frac{r}{R}\right)^{d-2},$$

where $\|u\| = R > r$.

PROOF OF THEOREM 2. By Theorem A we have

$$\mathbf{P}\{\{W(t), t \geqq R^2\} \cap C(0,R) = \emptyset, \|W(R^2)\| \geqq 2R\} \geqq p,$$

where $p > 0$ is an absolute constant. Even

$$\mathbf{P}\left\{\{W(t),\ t \geqq R^2\} \cap C(0,R) = \emptyset, \|W(R^2)\| \geqq 2R \,\middle|\, \inf_{t\leqq R^2} W_1(t) \geqq \frac{-R}{2\log R}\right\}$$
$$\geqq p.$$

Consequently,

$$\{W(t),\ t \geqq 0\} \cap \left\{x = (x_1, x_2, \ldots, x_d) :\ x_1 \leqq \frac{-R}{2\log R},\ \|x\| \leqq R\right\} =$$
$$= \emptyset \qquad \text{a.s. i.o.}.$$

Hence we have Theorem 2.

4. Proof of Theorem 3

Let $K = K(R)$ be the smallest positive integer for which

$$K \geqq 2\left(\frac{\log\log R}{\log 3}\right)^{1-\varepsilon d} \qquad (\varepsilon d < 1).$$

Then there exists a sequence $x_1, x_2, \ldots, x_K$ such that

$$\bigcup_{i=1}^{K} C_i = C(0,R),$$

where

$$C_i = C\left(x_i, \frac{R}{(\log\log R)^{1/d-\varepsilon}}\right).$$

Observe that if for some i $(i = 1, 2, \ldots, K)$

$$\{W(t,\omega), t \geqq 0\} \cap C_i \neq \emptyset,$$

then the probability, that a neighbouring ball of C_i will be visited by $W(\cdot)$, is larger than or equal to 3^{2-d}. Hence

$$\mathbf{P}\left\{\bigcap_{i=1}^{K}\{\{W(t,\omega), t \geqq 0\} \cap C_i \neq \emptyset\}\right\} \geqq 3^{-K(d-2)} \sim$$

$$\sim \exp\left(-(\log 3)2(d-2)\left(\frac{\log\log R}{\log 3}\right)^{1-\varepsilon d}\right) = \alpha_R$$

and

$$\sum_{k=1}^{\infty} \alpha_{R(k)} = \infty$$

if

$$R(k) = \left(\frac{\log k}{\log 3}\right)^{k/d}.$$

Observe also that

$$\frac{R(k+1)}{(\log\log R(k+1))^{1/d-\varepsilon}} > R(k).$$

Hence we have Theorem 3 by Borel–Cantelli lemma.

5. Proof of Theorem 4

Recall the following two Theorems.

THEOREM OF DVORETZKY–ERDŐS ([1], p. 195). *For any $\varepsilon > 0$ and $d > 3$ we have*

$$\|(W_2(t), W_3(t), \ldots, W_d(t))\| \geqq t^{1/2}(\log t)^{-(1+\varepsilon)/(d-3)} \qquad a.s.$$

if t is big enough.

THEOREM B ([1], p. 253). *Let $d = 3$ and*

$$\xi(R) = \lambda\{t : t \geqq 0, \|W(t)\| \leqq R\},$$

where λ is the Lebesgue measure. Then

$$\xi(R) \leqq R^2(\log R)^{1+\varepsilon} \qquad a.s.$$

if R is big enough.

LEMMA 4. *Consider the cone $\mathcal{K}(x,\vartheta)$ with $x=(1,0,\dots,0)$ and $\vartheta=\vartheta_R=(\log R)^{-(1+\varepsilon)/(d-3)}$. Then*

$$\mathcal{K}(x,\vartheta)\cap(\mathbb{R}^d-\mathcal{C}(0,R))\cap\{W(t,\omega),\ t\geqq 0\}=\emptyset \qquad a.s.$$

if R is big enough.

PROOF. Clearly if $W_1(t)\geqq r$ then

$$t\geqq\frac{r^2}{2\log\log r}.$$

Hence by the Theorem of Dvoretzky–Erdős

$$\|(W_2(t),\dots,W_d(t))\|\geqq\frac{r}{(2\log\log r)^{1/2}}(2\log r)^{-(1+\varepsilon)/(d-3)}\geqq$$

$$\geqq r(\log r)^{-(1+2\varepsilon)/(d-3)}.$$

Hence $(W_1(t),W_2(t),\dots,W_d(t))\notin K(x,\vartheta)$ if t is big enough and we have Lemma 4. Theorem 4 in case $d\geqq 4$ is a simple consequence of Lemma 4.

Theorem B implies Theorem 4 in case $d=3$.

In order to prove Theorem 4 in case $d=3$ observe that in $\mathcal{C}(0,R)$ one can find $(\log R)^{3+3\varepsilon}$ disjoint balls of radius

$$R^*=\frac{R}{(\log R)^{1+\varepsilon}}.$$

The visit of $(\log R)^{3+3\varepsilon}$ disjoint balls of radius R^* requires at least

$$\frac{(R^*)^2}{(\log\log R)^3}(\log R)^{3+3\varepsilon}\geqq R^2(\log R)^{1+\varepsilon/2}$$

time. Hence Theorem B implies Theorem 4 in case $d=3$.

REFERENCE

[1] RÉVÉSZ, P., *Random walk in random and non-random environments*, World Scientific Publ. Co., Teaneck, NJ, 1990. *MR* **92c**:60096

(Received January 22, 1996)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

TECHNISCHE UNIVERSITÄT
INSTITUT FÜR STATISTIK
WIEDNER HAUPTSTRASSE 8–10
A–1040 WIEN
AUSTRIA

revesz@ci.tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 127–151

ERDŐS–RÉNYI–SHEPP TYPE LAWS IN THE NON-I.I.D. CASE

A. FROLOV, A. MARTIKAINEN and J. STEINEBACH

Dedicated to Professor E. Csáki for his sixtieth birthday

1. Introduction

Consider a sequence $X_1, X_2, \dots$ of independent, but not necessarily identically distributed random variables (non-i.i.d. r.v.'s) with moment-generating functions $\phi_i(t) = \mathbf{E}\exp\{tX_i\}$ $(i=1,2,\dots)$. The following conditions will be assumed:

(A1) $\mathbf{E}X_i = 0$ $(i=1,2,\dots)$;

(A2) There exist positive constants H and $c_1, c_2, \dots$ such that $|L_i(z)| = |\log\phi_i(z)| \leqq c_i$ in the complex circle $|z| < H$. Here $\log z$ denotes the principal value of the natural logarithm of z;

(A3) $\limsup\limits_{n\to\infty} \sup\limits_j \dfrac{1}{B_{n,j}} \sum\limits_{i=n+1}^{n+j} (c_i^2 + c_i) < \infty$,

where

$$B_{n,j} = \sum_{i=n+1}^{n+j} \sigma_i^2, \quad \sigma_j^2 = \mathbf{E}X_j^2 \qquad (j=1,2,\dots);$$

(A4) There exist $\delta > 0$, j_0 such that $B_{n,j} > j\delta$ $\forall n$ $\forall j \geqq j_0$.

It is well known that, under (A2), $L_i(z)$ can be expanded into a convergent power series

$$L_i(z) = \sum_{k=1}^{\infty} \frac{\gamma_{k,i}}{k!} z^k, \qquad |z| < H, \tag{1.1}$$

where $\gamma_{k,i}$ is the k-th cumulant of X_i. We have $\gamma_{1,i} = 0$ and $\gamma_{2,i} = \sigma_i^2$ by (A1) and (A3). From Cauchy's inequality on the derivatives of analytic functions, it is obvious that (A2) is equivalent to the following condition:

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F10.

Key words and phrases. Erdős–Rényi–Shepp type laws, sums of independent random variables, large deviations, increments of sums, exact convergence rates.

(A2′) There exist positive constants H and $c_1, c_2, \dots$ such that $L_i(z)$ can be expanded into a power series (1.1) with

$$|\gamma_{k,i}| \leqq \frac{k!c_i}{H^k} \qquad \forall k, i.$$

For integer $n \geqq 0$ and $j \geqq 1$, put $S_0 = 0$, $\Phi_{n,0} = 1$,

$$S_n = \sum_{i=1}^{n} X_i, \quad \Phi_{n,j}(t) = \mathbf{E}\exp\{t(S_{n+j} - S_n)\} = \prod_{i=n+1}^{n+j} \phi_i(t),$$

$$(1.2) \qquad M_{n,j}(t) = \frac{\Phi'_{n,j}(t)}{\Phi_{n,j}(t)} = \sum_{i=n+1}^{n+j} m_i(t), \quad m_i(t) = (\log \phi_i(t))' \quad (i = 1, 2, \dots),$$

$$\rho_{n,j}(\alpha) = \inf_{t \in (0,H)} \{\Phi_{n,j}(t)\exp(-t\alpha B_{n,j})\}, \quad \alpha > 0.$$

For $c > 0$, let $\alpha_{n,j} = \alpha_{n,j}(c)$ denote the positive solution of the equation

$$\rho_{n,j}(\alpha) = \exp(-j/c),$$

provided this solution exists. In fact, since $S_{n+j} - S_n$ is nondegenerate, it follows from Lemma 2.1 in Deheuvels [9] that the function

$$\Psi_{n,j}(\alpha) = -\log\Big\{\inf_{t:\Phi_{n,j}(t)<\infty} \{\Phi_{n,j}(t)\exp(-t\alpha)\}\Big\},$$

satisfies

$$\lim_{\alpha\to\infty} \Psi_{n,j}(\alpha)/\alpha = H_{n,j} = \sup\{t\colon\ \Phi_{n,j}(t) < \infty\}$$
$$= \min_{n<i\leqq n+j} \{\sup\{t\colon\ \phi_i(t) < \infty\}\} = t_{n,j}.$$

The solution of the equation

$$\Psi_{n,j}(\alpha B_{n,j}) = j/c$$

has been discussed in detail by Deheuvels, Devroye and Lynch [11] and Deheuvels [9]. In particular, it has been shown there that this equation has a solution for $c \in (c_{n,j}, \infty)$, where

$$c_{n,j} = j \Big/ \int_0^{t_{n,j}} tM'_{n,j}(t)dt.$$

So, if c is sufficiently large, a solution exists. We assume that

$$H \leqq H_{n,j} \qquad \forall n, j,$$

and if the function $\Phi_{n,j}(t)\exp(-t\alpha B_{n,j})$ attains its minimum in $(0,H)$, let $t^*_{n,j}=t^*_{n,j}(c)$ denote the corresponding argument. Note that, by definition,

$$\rho_{n,j}(\alpha_{n,j})=\Phi_{n,j}(t^*_{n,j})\exp(-t^*_{n,j}\alpha_{n,j}B_{n,j})=\exp(-j/c), \tag{1.3}$$

$$M_{n,j}(t^*_{n,j})=\alpha_{n,j}B_{n,j}. \tag{1.4}$$

For integer sequences $K=K(N)$ consider the following Erdős–Rényi–Shepp type statistics:

$$U_N=\max_{0\leqq n\leqq N-K} t^*_{n,K}(S_{n+K}-S_n-\alpha_{n,K}B_{n,K}), \tag{1.5}$$

$$W_N=\max_{0\leqq n\leqq N-K}\max_{1\leqq k\leqq K} t^*_{n,k}(S_{n+k}-S_n-\alpha_{n,k}KB_{n,k}/k), \tag{1.6}$$

$$T_N=\max_{1\leqq n\leqq N} t^*_{n,K(n)}(S_{n+K(n)}-S_n-\alpha_{n,K(n)}KB_{n,K(n)}/K(n)). \tag{1.7}$$

In the case of independent, identically distributed random variables (i.i.d. r.v.'s) with $\mathbf{E}X_1=0$, $\mathbf{E}X_1^2=1$, Erdős and Rényi [13] studied the critical choice of $K=K(N)=[c\log N]$, where $[x]$ denotes (here and in the sequel) the integer part of x and $0\leqq c_0<c=c(\alpha)<\infty$, $0<\alpha<A$,

$$A=\lim_{t\uparrow H} m_1(t), \quad c_0=1\Big/\int_0^H tm_1'(t)dt, \quad H=\sup\{t>0\colon \Phi_1(t)<\infty\}.$$

They proved that

$$\lim_{N\to\infty}(U_N/K)=0 \quad \text{a.s.}, \tag{1.8}$$

i.e.,

$$\lim_{N\to\infty}\max_{0\leqq n\leqq N-K}(S_{n+K}-S_n)/K=\alpha \quad \text{a.s.}. \tag{1.9}$$

The definition of A and c_0 as given above has been introduced by Deheuvels, Devroye and Lynch [11], p. 211. It implies a necessary restriction on c which was not mentioned in explicit form in the original papers of Shepp [21], and Erdős and Rényi [13]. Moreover, several authors following these first papers made the oversight of stating their theorems as valid for all $c>0$ (see Remark 3, p. 212 in [11]). For a full form of the Erdős–Rényi law of large numbers covering also the case $0<c\leqq c_0$, and for a further discussion of c_0, confer [10].

Under the same assumptions, Shepp [21] earlier proved that

$$\lim_{N\to\infty}(T_N/K)=0 \quad \text{a.s.}, \tag{1.10}$$

i.e.

$$\lim_{N\to\infty} \max_{0\leqq n\leqq N-K} (S_{n+K(n)} - S_n)/K(n) = \alpha \quad \text{a.s..} \tag{1.11}$$

Csörgő and Steinebach [7] obtained a first convergence rate statement for (1.8)–(1.9) by showing that

$$\lim_{N\to\infty} (U_N/K_N^{1/2}) = 0 \quad \text{a.s.,} \tag{1.12}$$

i.e.,

$$\max_{0\leqq n\leqq N-K} (S_{n+K} - S_n)/K = \alpha + o(K^{-1/2}) \quad \text{a.s..} \tag{1.13}$$

An analogous assertion was also given for W_N. The exact convergence rates have been derived by Deheuvels, Devroye, Lynch [11] (for U_N, T_N) and Deheuvels and Devroye [10] (for W_N). They are summarized in the following theorem:

THEOREM A. *For $c_0 < c < \infty$, choose $K = K(N) = [c \log N]$. Then*

(i) $$\lim_{N\to\infty} (U_N/\log K) = -\frac{1}{2} \quad \textit{in probability;}$$

(ii) $$\limsup_{N\to\infty} (U_N/\log K) = \frac{1}{2} \quad \textit{a.s.;}$$

(iii) $$\liminf_{N\to\infty} (U_N/\log K) = -\frac{1}{2} \quad \textit{a.s..}$$

In statements (i)–(iii), *U_N can be replaced by W_N and T_N, respectively.*

While the Erdős–Rényi–Shepp laws of (1.8)–(1.11) retain if $K = [c \log N]$ is replaced by any integer sequence $K(N) \sim c \log N$ as $N \to \infty$, their convergence rate counterparts of Theorem A critically depend on the specific choice of $K(N)$. This is obvious from an extension of Theorem A due to Bacro [1] who proved:

THEOREM B. *For $c_0 < c < \infty$ and $\lambda \in R$, choose $K = K(N) = [c \log N + \lambda \log\log N]$. Then*

(i) $$\lim_{N\to\infty} (U_N/\log K) = -\frac{1}{2} - \frac{\lambda}{c} \quad \textit{in probability;}$$

(ii) $$\limsup_{N\to\infty} (U_N/\log K) = \frac{1}{2} - \frac{\lambda}{c} \quad \textit{a.s.;}$$

(iii) $$\liminf_{N\to\infty} (U_N/\log K) = -\frac{1}{2} - \frac{\lambda}{c} \quad \textit{a.s..}$$

In statements (i)–(iii), U_N *can be replaced by* T_N.

Some analogues of the above partial sum results have also been obtained for renewal processes (cf. e.g. Bacro, Deheuvels, Steinebach [2], Deheuvels, Steinebach [12] and Steinebach [22]). For a rather general methodology concerning the increments of stochastic processes, we also refer to the work of Csörgő [8], Steinebach [23], Csáki, Földes, Komlós [6], Csáki [4], Deheuvels [9], and Csáki, Csörgő [5].

Main aim of our present work is to extend the results of Theorems A and B to the non-i.i.d. case as follows:

THEOREM 1. *Let* $X_1, X_2, \ldots$ *be a non-i.i.d. sequence of r.v.'s satisfying* (A1)–(A4). *Choose* $K = K(N) = [c \log N + \lambda \log\log N]$, *and let* U_N, W_N, T_N *be as defined in* (1.5)–(1.7). *Then, there exists* $c_0 \geqq 0$ *such that for any* $c > c_0$ *and* $\lambda \in R$, *we have*

(i) $$\lim_{N\to\infty} (U_N / \log K) = -\frac{1}{2} - \frac{\lambda}{c} \quad \textit{in probability;}$$

(ii) $$\limsup_{N\to\infty} (U_N / \log K) = \frac{1}{2} - \frac{\lambda}{c} \quad \textit{a.s.;}$$

(iii) $$\liminf_{N\to\infty} (U_N / \log K) = -\frac{1}{2} - \frac{\lambda}{c} \quad \textit{a.s.}.$$

In statements (i)–(iii), U_N *can be replaced by* W_N *or* T_N, *respectively, if additionally* $\sigma_i^2 \geqq \sigma^2 > 0$.

For an earlier extension of the Erdős–Rényi law of large numbers to the non-i.i.d. case see also Lin [17]. His results correspond to the convergence rate statements of Csörgő and Steinebach [7] in the i.i.d. case which preceded Theorems A and B. Under stronger conditions, Frolov [14], [15], [16] obtained a first version of Theorem 1 in the case $K = [c \log N]$, i.e. $\lambda = 0$.

The proof of Theorem 1 is essentially based on the following extension of Petrov's [18] large deviation result for sums of (non-i.i.d.) random variables which is of independent interest. To formulate this result, we consider an array $\{X_{n,j}, n \geqq 1,\ j = 1, \ldots, k_n\}$ of row-wise independent random variables. Assume that

(A5) $\mathbf{E}X_{n,j} = 0 \quad \forall n, j$;

(A6) There exist positive constants H, $c_{n,1}, c_{n,2}, \ldots$ such that $|L_{n,j}(z)| \leqq c_{n,j}$ in the complex circle $|z| < H$, where $L_{n,j}(z) = \log \phi_{n,j}(z) = \log \mathbf{E}\exp(zX_{n,j})$;

(A7) $$W = \limsup_{n\to\infty} \frac{1}{B_n} \sum_{j=1}^{k_n} (c_{n,j}^2 + c_{n,j}) < \infty,$$

where $B_n = \sum_{j=1}^{k_n} \mathbf{E}X_{n,j}^2 \to \infty$ as $n \to \infty$.

Similarly to (1.1), $L_{n,j}(z)$ can be expanded into a convergent power series

$$L_{n,j}(z)=\sum_{k=1}^{\infty}\frac{\gamma_{k,n,j}}{k!}z^k, \qquad |z|<H. \tag{1.14}$$

$\gamma_{k,n,j}$ the k-th cumulant of $X_{n,j}$, and an equivalent condition to (A6) is given by

(A6′) There exist positive constants H, and $c_{n,1}, c_{n,2}, \dots$ such that $|L_{n,j}(z)| \leqq c_{n,j}$ can be expanded into a power series (1.14) with

$$|\gamma_{k,n,j}| \leqq \frac{k! c_{n,j}}{H^k} \qquad \forall k, n, j. \tag{1.15}$$

We have $\gamma_{1,n,j}=\mathbf{E}X_{n,j}=0$ by (A5), and $\gamma_{2,n,j}=\mathbf{E}X_{n,j}^2$. Set

$$F_n(x)=\mathbf{P}\Big(\sum_{j=1}^{k_n}X_{n,j}<x\sqrt{B_n}\Big), \qquad \Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt.$$

THEOREM 2. *Assume* (A5)–(A7). *Then*

(i) $$\frac{1-F_n(x)}{1-\Phi(x)}=\exp\Big\{\frac{x^3}{\sqrt{B_n}}\lambda_n\Big(\frac{x}{\sqrt{B_n}}\Big)\Big\}\Big[1+O\Big(\frac{x+1}{\sqrt{B_n}}\Big)\Big],$$

(ii) $$\frac{F_n(-x)}{\Phi(-x)}=\exp\Big\{-\frac{x^3}{\sqrt{B_n}}\lambda_n\Big(-\frac{x}{\sqrt{B_n}}\Big)\Big\}\Big[1+O\Big(\frac{x+1}{\sqrt{B_n}}\Big)\Big],$$

for all $x \geqq 0$ such that $x/\sqrt{B_n}$ remains sufficiently small. Here

$$\lambda_n(t)=\sum_{k=0}^{\infty}a_{k,n}t^k \tag{1.16}$$

denotes the generalized Cramér series (cf. Petrov [18], *Ch. VIII. 2) which, for n sufficiently large, is majorized by a power series with coefficients not depending on n, and convergent in some circle $|t|<t_0$, so that the series $\lambda_n(t)$ converges uniformly in n for $|t|<t_0$.*

For later use, put

$$\Gamma_{k,n}=\frac{1}{B_n}\sum_{j=1}^{k_n}\gamma_{k,n,j}.$$

It will be proved below that

$$\lambda_n(t)=\frac{1}{2t}-\frac{1}{t^3}\sum_{k=2}^{\infty}\frac{(k-1)\Gamma_{k,n}}{k!}z^k, \quad t>0,$$

where $z=z(t)$ is defined by (2.18) and (2.20), and that that $\lambda_n(t)\sim -\Gamma_{3,n}/3$ as $t\to 0$.

Hence the coefficients $a_{k,n}$ can be expressed in terms of the cumulants of $X_{n,1},\dots,X_{n,k_n}$ up to the order $k+3$. In particular,

$$a_{0,n}=\Gamma_{3,n}/6,\ a_{1,n}=(\Gamma_{4,n}-3\Gamma_{3,n}^2)/24,\ a_{2,n}=(\Gamma_{5,n}-10\Gamma_{4,n}\Gamma_{3,n}+15\Gamma_{3,n}^3)/120.$$

2. Large deviation results

In the sequel, all limits are supposed to be taken as $n\to\infty$ if not mentioned otherwise, and $C_1,C_2,\dots$ denote some positive constants.

Proof of Theorem 2. Via the Esscher transform, we first introduce an auxiliary array $\{\overline{X}_{n,j},n\geqq 1,j=1,\dots,k_n\}$ of row-wise independent r.v.'s with distribution functions

$$\overline{V}_{n,j}(x)=\exp(-L_{n,j}(z))\int_{-\infty}^{x}\exp(zy)dV_{n,j}(y),\qquad -H<z<H,$$

where $V_{n,j}(x)=\mathbf{P}(X_{n,j}<x)$. We write

$$\overline{m}_{n,j}=\mathbf{E}\overline{X}_{n,j},\quad \overline{\sigma}_{n,j}^2=\mathbf{E}(\overline{X}_{n,j}-\overline{m}_{n,j})^2,\quad \overline{M}_n=\sum_{j=1}^{k_n}\overline{m}_{n,j},\quad \overline{B}_n=\sum_{j=1}^{k_n}\overline{\sigma}_{n,j}^2$$

$$G_{n,j}(x)=\overline{V}_{n,j}(x+\overline{m}_{n,j}),\quad \overline{S}_n=\sum_{j=1}^{k_n}\overline{X}_{n,j},\quad \overline{F}_n(x)=\mathbf{P}\left(\overline{S}_n<\overline{M}_n+x\sqrt{\overline{B}_n}\right).$$

Note that $\overline{V}_{n,j},\overline{m}_{n,j}$ etc. critically depend on the choice of z. This will be important in the proof below. Direct calculation shows that the cumulant-generating function

$$\overline{L}_{n,j}(h)=\log\mathbf{E}\exp(h\overline{X}_{n,j})=-L_{n,j}(z)+L_{n,j}(h+z)$$

exists for h such that $|h+z|<H$. Clearly

$$\overline{\gamma}_{k,n,j}=\left[\frac{d^k\overline{L}_{n,j}(h)}{dh^k}\right]_{h=0}=\left[\frac{d^kL_{n,j}(t)}{dt^k}\right]_{t=z}.$$

For the first two cumulants, it follows now that

$$\text{(2.17)}\qquad \overline{m}_{n,j}=\overline{\gamma}_{1,n,j}\sum_{k=2}^{\infty}\frac{\overline{\gamma}_{k,n,j}}{(k-1)!}z^{k-1},\quad \overline{\sigma}_{n,j}^2=\overline{\gamma}_{2,n,j}=\sum_{k=2}^{\infty}\frac{\gamma_{k,n,j}}{(k-2)!}z^{k-2}.$$

Both series converge in the circle $|z| < H$.

We write

$$t = \frac{\overline{M}_n}{B_n} = z + \sum_{k=3}^{\infty} \frac{\Gamma_{k,n}}{(k-1)!} z^{k-1}. \tag{2.18}$$

This series converges for $|z| < H$, too. By (A6) and (A7), its coefficients can be estimated by

$$\frac{|\Gamma_{k,n}|}{(k-1)!} \leqq \frac{k}{B_n H^k} \sum_{j=1}^{k_n} c_{n,j} \leqq C_1 \frac{k}{H^k}, \tag{2.19}$$

so that the series in (2.18) is majorized by a power series with coefficients not depending on n and convergent for $|z| < H$. In any smaller circle $|z| < H_1$, $0 < H_1 < H$, this series converges uniformly in n and z.

For all sufficiently small $|t|$, (2.18) has the unique real root

$$z = t - \frac{\Gamma_{3,n}}{2} t^2 - \frac{\Gamma_{4,n} - 3\Gamma_{3,n}^2}{6} t^3 + \dots. \tag{2.20}$$

This root tends to zero as $t \to 0$. If t remains small as $n \to \infty$ (which is assumed further) z also remains small.

By (A6) and the theorem on the inversion of analytic functions (cf. Privalov [19], p. 258), there exists a circle, the same for all sufficiently large n, with center at $t = 0$, within which the series on the right-hand side of (2.20) converges, and the absolute value of its sum does not exceed H_1. Applying the Cauchy inequality to the coefficients of this series, we find that for all sufficiently large n the series itself is majorized by a series with coefficients independent of n and a positive radius of convergence.

The definition of $\overline{V}_{n,j}$ implies that $V_{n,j}(x) = e^{L_{n,j}(z)} \int_{-\infty}^{x} e^{-zy} d\overline{V}_{n,j}(y)$. Hence

$$\mathbf{P}(S_n \geqq x) = \exp\Big\{\sum_{j=1}^{k_n} L_{n,j}(z)\Big\} \int_x^{\infty} e^{-zu} d\mathbf{P}(\overline{S}_n < u).$$

Writing $y_1 = z\sqrt{B_n}$ and substituting $u = \overline{M}_n + y\sqrt{B_n}$ we get

$$1 - F_n(x) = \exp\Big\{-z\overline{M}_n + \sum_{j=1}^{k_n} L_{n,j}(z)\Big\} \int_{(x\sqrt{B_n} - \overline{M}_n)/\sqrt{B_n}}^{\infty} e^{-yy_1} d\overline{F}_n(y). \tag{2.21}$$

Put $r_n = \overline{F}_n(x) - \Phi(x)$. We will use the following estimate due to Rozovsky [20]:

$$\sup_x |r_n(x)| \leqq C_2 \overline{B}_n^{-3/2} \Big(|D_n| + \sum_{j=1}^{k_n} \mathbf{E}_{n,j} \Big),$$

where

$$D_n = \sum_{j=1}^{k_n} \int\limits_{|x| \leqq \sqrt{\overline{B}_n}} x^3 dG_{n,j}(x), \quad \mathbf{E}_{n,j} = \sup_{z \geqq 0} z \int\limits_{|x|>z} x^2 dG_{n,j}(x).$$

The well-known relationship between the cumulants and central moments yield the equalities

$$\mathbf{E}(\overline{X}_{n,j} - \overline{m}_{n,j})^3 = \overline{\gamma}_{3,n,j}, \quad \mathbf{E}(\overline{X}_{n,j} - \overline{m}_{n,j})^4 = \overline{\gamma}_{4,n,j} + 3\overline{\sigma}_{n,j}^4.$$

It is clear that

$$\begin{aligned} |D_n| &\leqq \sum_{j=1}^{k_n} \big| \mathbf{E}(\overline{X}_{n,j} - \overline{m}_{n,j})^3 \big| + \sum_{j=1}^{k_n} \int\limits_{|x|>\sqrt{\overline{B}_n}} |x|^3 dG_{n,j}(x) \\ &\leqq \sum_{j=1}^{k_n} |\overline{\gamma}_{3,n,j}| + \overline{B}_n^{-1/2} \sum_{j=1}^{k_n} \mathbf{E}(\overline{X}_{n,j} - \overline{m}_{n,j})^4 \\ &\leqq \sum_{j=1}^{k_n} |\overline{\gamma}_{3,n,j}| + \overline{B}_n^{-1/2} \sum_{j=1}^{k_n} |\overline{\gamma}_{4,n,j}| + 3 \sum_{j=1}^{k_n} \overline{\sigma}_{n,j}^4. \end{aligned}$$

Note that $\mathbf{E}_{n,j}$ can be estimated from above by $\mathbf{E}_{n,j} \leqq \overline{\sigma}_{n,j}^2$ if the supremum is achieved at $\hat{z} = \hat{z}(n,j) \leqq 1$, and by $\mathbf{E}_{n,j} \leqq \hat{z}^{-2} \mathbf{E}(\overline{X}_{n,j} - \overline{m}_{n,j})^4 \leqq |\overline{\gamma}_{4,n,j}| + 3\overline{\sigma}_{n,j}^4$, otherwise. Hence

$$\sum_{j=1}^{k_n} \mathbf{E}_{n,j} \leqq \overline{B}_n + \sum_{j=1}^{k_n} |\overline{\gamma}_{4,n,j}| + 3 \sum_{j=1}^{k_n} \overline{\sigma}_{n,j}^4,$$

and we arrive at the estimate

$$\sup_x |r_n(x)| \leqq C_2 \overline{B}_n^{-3/2} \Big(\overline{B}_n + \sum_{j=1}^{k_n} |\overline{\gamma}_{3,n,j}| + (1 + \overline{B}_n^{-1/2}) \sum_{j=1}^{k_n} |\overline{\gamma}_{4,n,j}| + 6 \sum_{j=1}^{k_n} \overline{\sigma}_{n,j}^4 \Big). \tag{2.22}$$

It follows from (2.17) that

$$\frac{\overline{B}_n}{B_n} = \frac{1}{B_n}\sum_{j=1}^{k_n}\overline{\sigma}_{n,j}^2 \leqq 1 + \sum_{k=3}^{\infty}\frac{\Gamma_{k,n}}{(k-2)!}z^{k-2}$$

and, like in (2.19), this series is again majorized by one with coefficients independent of n and a positive radius of convergence. Hence

$$C_3 B_n \leqq \overline{B}_n \leqq B_n C_4. \tag{2.23}$$

We find from (1.15) and (2.17) that

$$\overline{\sigma}_{n,j}^2 \leqq \sum_{k=2}^{\infty}\frac{\gamma_{k,n,j}}{(k-2)!}|z|^{k-2} \leqq \frac{c_{n,j}}{H^2}\sum_{k=2}^{\infty}k(k-1)\Big(\frac{|z|}{H}\Big)^{k-2} \leqq c_{n,j}C_5 \tag{2.24}$$

in the circle $|z| \leqq H_1$ for every positive $H_1 < H$. As in (2.17) we write for $m = 3, 4$,

$$\overline{\gamma}_{m,n,j} = \sum_{k=m}^{\infty}\frac{\gamma_{k,n,j}}{(k-m)!}z^{k-m},$$

and, as in (2.24), we find that

$$|\overline{\gamma}_{3,n,j}| \leqq c_{n,j}C_6, \qquad |\overline{\gamma}_{4,n,j}| \leqq c_{n,j}C_7. \tag{2.25}$$

Substituting (2.23), (2.24) and (2.25) in (2.22), we get

$$\sup_x |r_n(x)| \leqq C_8\overline{B}_n^{-3/2}\Big(\sum_{j=1}^{k_n}(c_{n,j}+c_{n,j}^2)+\overline{B}_n\Big) \leqq C_9\overline{B}_n^{-1/2}. \tag{2.26}$$

Here, we have also used (A7).

Now take

$$t = \frac{x}{\sqrt{\overline{B}_n}}. \tag{2.27}$$

Then t remains small, and by (2.18) we integrate in (2.21) from 0 to ∞.

By (2.26) we have

$$\begin{aligned}&\int_0^\infty \exp(-yy_1)d\overline{F}_n(y)\\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty \exp\Big\{-yy_1k - \frac{y^2}{2}\Big\}dy - r_n(0) + y_1\int_0^\infty r_n(y)e^{-yy_1}dy\end{aligned} \tag{2.28}$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty \exp\Big(-yy_1 - \frac{y^2}{2}\Big) dy + \alpha_n,$$

where

(2.29) $$|\alpha_n| \leqq \frac{2C_2}{\sqrt{B_n}}.$$

In the sequel, we confine ourselves to the case $x > 1$, because, for $0 \leqq x \leqq 1$, an application of Rozovsky's above estimate to $\beta_n = \sup_x |F_n(x) - \Phi(x)|$ gives, similarly to (2.26), $\beta_n = O(B_n^{-1/2})$, implying (i) and (ii) for $0 \leqq x \leqq 1$.

By (2.17) to (2.19), we conclude that

$$\overline{M}_n = B_n z + B_n \sum_{k=3}^\infty \frac{\Gamma_{k,n}}{(k-1)!} z^{k-1} = B_n z(1 + O(z))$$

and similarly

(2.30) $$\overline{B}_n = B_n + B_n \sum_{k=3}^\infty \frac{\Gamma_{k,n}}{(k-2)!} z^{k-2} = B_n(1 + O(z)).$$

Note that the constants in $O(\cdot)$ depend only on H and W so that $O(\cdot)$ does not exceed $1/2$ for all large n and small z. It follows from (2.20), (2.27) and (2.30) that

(2.31) $$y_1 = x(1 + O(z)).$$

Hence, $y_1 > 1/2$ for large n and small z. Set

$$I_1 = \int_0^\infty \exp\Big(-yy_1 - \frac{y^2}{2}\Big) dy, \qquad I_2 = \int_0^\infty \exp\Big(-yx - \frac{y^2}{2}\Big) dy.$$

Substituting $u = yy_1$ we find that

$$1 > y_1 I_1 = \int_0^\infty \exp\Big(-u - \frac{u^2}{2y_1^2}\Big) du > \int_0^\infty \exp(-u - 2u^2) du > 0.$$

By (2.30) we get

(2.32) $$C_{10} < z\sqrt{B_n} I_1 < C_{11}.$$

Recalling (2.28), (2.29), we find that for small z

$$\int_0^\infty \exp(-yy_1)d\overline{F}_n(y) = \frac{1}{\sqrt{2\pi}}I_1 + \alpha_n = \frac{1}{\sqrt{2\pi}}I_1(1+O(z)). \tag{2.33}$$

In the further arguments, we make use of Mill's ratio

$$\psi(y) = \frac{1-\Phi(y)}{\Phi'(y)} = e^{y^2/2}\int_y^\infty e^{-t^2/2}dt,$$

and write $I_1 = \psi(y_1)$, $I_2 = \psi(x)$. We have $\psi(x) - \psi(y_1) = (x-y_1)\psi'(u)$, where u is between y_1 and x. Furthermore $|\psi'(u)| = |u\psi(u) - 1| < u^{-2}$ for $u > 0$. By (2.31), $x - y_1 = O(zy_1)$, and consequently $\psi(x) - \psi(y_1) = O(z/y_1)$. So, in view of (2.23) we conclude that $I_2 - I_1 = O(B_n^{-1/2})$. It follows from (2.32) that $I_1 = I_2(1+O(z))$. Hence from (2.33) we find that

$$\int_0^\infty \exp(-yy_1)d\overline{F}_n(y) = \frac{1}{\sqrt{2\pi}}\psi(x)(1+O(z)).$$

This reduces (2.21) to the relation

$$1 - F_n(x) = \exp\left\{\frac{x^2}{2} - z\overline{M}_n + \sum_{j=1}^{k_n} L_{n,j}(z)\right\}[1-\Phi(x)](1+O(z)). \tag{2.34}$$

By (1.14) and (2.17),

$$z\overline{m}_{n,j} - L_{n,j}(z) = \sum_{k=2}^\infty \frac{k-1}{k!}\gamma_{k,n,j}z^k \quad (n = 1,2,\ldots;\ j = 1,2,\ldots,k_n).$$

This and (2.20) imply that

$$\frac{1}{B_n}\Big(z\overline{M}_n - \sum_{j=1}^{k_n} L_{n,j}(z)\Big)$$
$$= \sum_{k=2}^\infty \frac{(k-1)\Gamma_{k,n}}{k!}z^k = \frac{t^2}{2} - \frac{\Gamma_{3,n}}{6}t^3 - \frac{\Gamma_{4,n} - 3\Gamma_{3,n}^2}{24}t^4 + \ldots,$$

or

$$\frac{t^2}{2} - \frac{1}{B_n}\Big(z\overline{M}_n - \sum_{j=1}^{k_n} L_{n,j}(z)\Big) = t^3\lambda_n(t).$$

Hence

$$\frac{x^2}{2} - z\overline{M}_n + \sum_{j=1}^{k_n} L_{n,j}(z) = B_n t^3 \lambda_n(t).$$

Note that by (2.18), (2.20) and (2.27), $z = O(x/\sqrt{B_n})$. Then, by (2.34), we obtain (i).

Assertion (ii) can be proved in the same way.

REMARK 1. It follows from the conditions (A6) and (A7) that

$$\sup_n \frac{B_n}{k_n} < \infty.$$

PROOF. Put $A_n = \{j\colon c_{n,j} \geqq A,\ 1 \leqq j \leqq k_n\}$, $\overline{A}_n = \{1, \ldots, k_n\} \setminus A_n$. By Cauchy's inequality and (A7), for large n,

$$\sum_{j \in A_n} \mathbf{E}X_{n,j}^2 \leqq \frac{2}{H^2} \sum_{j \in A_n} c_{n,j} \leqq \frac{2}{AH^2} \sum_{j \in A_n} c_{n,j}^2 \leqq \frac{2(W+1)}{AH^2} B_n,$$

or

$$\sum_{j \in A_n} \mathbf{E}X_{n,j}^2 \leqq \frac{1}{2} B_n$$

if $A = 4(W+1)H^{-2}$. Hence

$$B_n \leqq 2 \sum_{j \in \overline{A}_n} \mathbf{E}X_{n,j}^2 \leqq \frac{4}{H^2} \sum_{j \in \overline{A}_n} c_{n,j} \leqq \frac{4A}{H^2} k_n.$$

LEMMA 1. *Let the assumptions of Theorem 2 hold, and*

$$\liminf \frac{B_n}{k_n} > 0. \tag{2.35}$$

Consider the function

$$u_n(t) = \frac{1}{k_n} \sup_{0<z<H} \Big\{ ztB_n - \sum_{j=1}^{k_n} L_{n,j}(z) \Big\}, \quad t > 0, \tag{2.36}$$

and its inverse $v_n = u_n^{-1}$. There exists $c' > 0$ such that if $0 < c \leqq c'$ then the sequences $t_n = v_n(c)$ and $z_n = z_n(t_n)$ are bounded and separated from 0: $0 < t' \leqq t_n \leqq t''$ and $0 < z' \leqq z_n \leqq z''$ for all large enough n.

PROOF. For any fixed n, the derivative

$$(ztB_n - L_n(z))'_z = tB_n - \overline{M}_n(z)$$

equals 0 at $z=z_n(t)$, which is a solution of equation (2.18). Moreover, it is positive as $z\in(0,z_n(t))$ and negative as $z>z_n(t)$.

The function $z_n(t)$ satisfies (2.20) for all sufficiently small $t>0$. The inverse function $t=t_n(z)$ satisfies (2.18) as $0\leqq z<H$. It follows from (2.19) that

$$\frac{t}{2}\leqq z_n(t)\leqq 2t \quad \text{as} \quad 0\leqq t\leqq \bar{t} \tag{2.37}$$

for large enough n and some $\bar{t}>0$, which does not depend on n. Hence $z_n(t)<H$ for some $\tilde{t}\in(0,\bar{t}]$, and for all $t\in(0,\tilde{t}]$,

$$u_n(t)=\frac{1}{k_n}\Big(z_n(t)tB_n-\sum_{j=1}^{k_n}L_{n,j}(z_n(t))\Big), \quad t\in(0,\tilde{t}]. \tag{2.38}$$

By (2.18) we have

$$k_nu_n'(t)=z_n'(t)tB_n+z_n(t)B_n-\overline{M}_n(z_n(t))z_n'(t)=z_n(t)B_n \tag{2.39}$$

and therefore

$$u_n(t)=\frac{B_n}{k_n}\int_0^t z_n(s)ds.$$

By (2.37), (2.35) and Remark 1, we have for some positive $C_{13}>C_{12}>0$ and all sufficiently large n

$$C_{12}\frac{t^2}{2}\leqq u_n(t)\leqq C_{13}t^2, \quad 0\leqq t\leqq \tilde{t}.$$

The functions $u_n(t)$ are continuous, increasing and $u_n(0)=0$. The inverse functions $v_n=u_n^{-1}$ have the same properties. Moreover,

$$\sqrt{\frac{c}{C_{12}}}\leqq v_n(c)\leqq\sqrt{\frac{c}{C_{13}}} \quad \text{as} \quad 0\leqq c\leqq\frac{1}{4}\tilde{t}^2.$$

In view of (2.37), the assertion of Lemma 1 follows.

LEMMA 2. *Let the assumptions of Lemma 1 hold. Then for any $c\in(0,c']$ and any ε,*

$$\mathbf{P}(S_n>s_n\sqrt{B_n})=e^{-ck_n}\frac{T_n}{t_nk_n^{1/2+\varepsilon}},$$

with $s_n=t_n+\varepsilon\log B_n/(B_nz_n)$, t_n,z_n as in Lemma 1. Here $C_{14}\leqq T_n\leqq C_{15}$ for some positive constants C_{14}, C_{15} and all sufficiently large n.

PROOF. Put $x=a_n\sqrt{B_n}$. Let $x\to\infty$, $a_n\leqq\tilde{t}$, $a_n>0$. By (2.34) and (2.38),

$$\begin{aligned}\mathbf{P}(S_n > a_n\sqrt{B_n}) &= \frac{1}{a_n\sqrt{B_n}}\exp\Big(-z_n(a_n)a_nB_n + \sum_{j=1}^{k_n} L_{n,j}(z_n(a_n))\Big)R_n \\ &= \frac{1}{a_n\sqrt{B_n}}e^{-k_nu_n(a_n)}R_n,\end{aligned} \tag{2.40}$$

where $|R_n - 1/\sqrt{2\pi}| \leqq 1/10$ for large enough n if z_n remains small.

The functions $u_n(t)$ and $z_n(t)$ are analytic in a small circle $|t| \leqq \bar{t}$. It is not difficult to check that the sequence $\{\sup\limits_{|t|\leqq t_0} |u_n''(t)|\}$ is bounded for sufficiently small $t_0 > 0$. Actually, by (2.18) we have

$$t_n'(z) = 1 + \sum_{k=1}^{\infty} \frac{\Gamma_{k+2,n}}{k!}z^k,$$

where $|\Gamma_{k+2,n}|/k! \leqq C(k+1)(k+2)H^{-k-2}$ in view of (2.19). Hence $t_n'(z) \geqq 1/2$ for all small enough z, and $z_n'(t) \leqq 2$ for all small enough t, $|t| \leqq t_0$. By (2.39), (2.35) and Remark 1, we obtain $\sup\limits_{|t|\leqq t_0} |u_n''(t)| \leqq C_{16}$.

Let $a_n = t_n + \varepsilon_n$, $|a_n| \leqq t_0$, $t_n \leqq t_0$. We have

$$u_n(a_n) = u_n(t_n) + u_n'(t_n)\varepsilon_n + W_n = c + z_n\varepsilon_n + W_n,$$

where $|W_n| \leqq C_{17}$. Put $\varepsilon_n = \varepsilon \log B_n/(B_nz_n)$. Then

$$u_n(s_n) = c + z_n\varepsilon\frac{\log B_n}{B_nz_n} + W_n.$$

The assertion of Lemma 2 follows from (2.40) and Remark 1.

3. Proof of Theorem 1

Once a suitable large deviation estimate like that of Lemma 2 has been established, the proof of Theorem 1 can be given adapting the methodology of the i.i.d. case (cf. Deheuvels, Devroye, Lynch [11], Deheuvels, Devroye [10], Bacro [1]). However, due to the fact of having independent, but not necessarily identically distributed summands, a number of modifications are necessary. For sake of readability of the paper, we outline the main steps of the proof.

Put $\overline{U}_N = U_N/\log K$. Applying Lemmas 1 and 2 to the sums $S_{n+j} - S_n$ we get the following results.

LEMMA 3. *There exists* $c_0 \geqq 0$ *having the following property. For any* $c > c_0$ *there exist sequences* $\{\alpha_{n,j}\}$ *and* $\{t^*_{n,j}\}$, *satisfying* (1.3) *and* (1.4). *Furthermore, there exist positive constants* α_1, α_2, t^*_1, t^*_2 *such that* $\alpha_1 \leqq \alpha_{n,j} \leqq \alpha_2$, $t^*_1 \leqq t^*_{n,j} \leqq t^*_2$ *for all* n *and* j *sufficiently large.*

LEMMA 4. *For any* ε, *there exist positive constants* C_{18}, C_{19} *such that*

$$\frac{C_{18}e^{-k/c}}{k^{1/2+(\pm 1/2+\varepsilon)}} \leqq \mathbf{P}\left(S_{n+k}-S_n \geqq B_{n,k}\alpha_{n,k} + \left(\pm\frac{1}{2}+\varepsilon\right)\frac{\log k}{t^*_{n,k}}\right) \leqq \frac{C_{19}e^{-k/c}}{k^{1/2+(\pm 1/2+\varepsilon)}}$$

for all n *and* $k \geqq K_1$.

The next result is well known (cf. Chung and Erdős [3]).

LEMMA 5. *For any events* $A_1, A_2, \dots, A_n$ *the following inequality holds:*

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) \geqq \frac{\left(\sum_{i=1}^{n} \mathbf{P}(A_i)\right)^2}{\sum_{i=1}^{n} \mathbf{P}(A_i) + \sum_{i\neq j} \mathbf{P}(A_i A_j)}.$$

LEMMA 6. *For any* $\varepsilon > 0$,

$$\mathbf{P}\left(\overline{U}_N \leqq -\frac{1}{2} - \frac{\lambda}{c} + \varepsilon\right) \to 1 \quad as \quad N \to \infty.$$

PROOF. By Lemma 4 we have

$$\begin{aligned}
&\mathbf{P}\left(\overline{U}_N \geqq -\frac{1}{2} - \frac{\lambda}{c} + \varepsilon\right)\\
&\leqq \sum_{n=0}^{N-K} \mathbf{P}\left(S_{n+K} - S_n \geqq \alpha_{n,K}B_{n,K} + \left(-\frac{1}{2} - \frac{\lambda}{c} + \varepsilon\right)\frac{\log K}{t^*_{n,K}}\right)\\
&\leqq \frac{C_{19}Ne^{-K/c}}{K^{\varepsilon-\lambda/c}} \leqq \frac{C_{20}}{K^{\varepsilon}} \to 0 \quad \text{as} \quad N \to \infty.
\end{aligned}$$

LEMMA 7. *For any* a, b, q, *and positive integers* $v < k$, *and for any positive* t, t_1, *the following inequality holds:*

$$\begin{aligned}
&\mathbf{P}(S_{i+k} - S_i \geqq a, S_{i+v+k} - S_{i+v} \geqq b)\\
&\leqq \Phi_{i+v,k-v}(t)e^{-tq} + \mathbf{P}(S_{i+k} - S_i \geqq a)\Phi_{i+k,v}(t_1)e^{-t_1(b-q)}.
\end{aligned}$$

PROOF. Confer Deheuvels, Devroye, Lynch [11], Lemma 4, in the i.i.d. case.

Now choose

$$t = t^*_{i+v,k}, \qquad b = \alpha_{i+v,k}B_{i+v,k} + \left(-u - \frac{\lambda}{c} - \varepsilon\right)\frac{\log k}{t},$$

$$q = \alpha_{i+v,k}B_{i+v,k} - \frac{1}{t}\log\Phi_{i+k,v}(t) + \frac{1}{t}\Big(s\log v - \frac{\lambda}{c}\log k\Big).$$

Note that, by Lemma 3, $t > \varepsilon_2 > 0$ for all i, v and large enough k. If t_1 is such that $t_1 < \varepsilon_2$, then the inequality $t_1 < t$ holds for large k.

LEMMA 8. *If* $\varepsilon > 0$, $\theta > 0$, $s > 0$, $u + s + \varepsilon \geqq 0$, *then*

$$\mathbf{P}(S_{i+k} - S_i \geqq a, S_{i+v+k} - S_{i+v} \geqq b) \leqq e^{-k/c}v^{-s}k^{\lambda/c} + \mathbf{P}(S_{i+k} - S_i \geqq a)k^{u+s+\varepsilon}v^{-\theta}$$

for any real a *and* $k \geqq v \geqq v_1 = v_1(\theta)$.

PROOF. Put $Q_{i,k} = \mathbf{P}(S_{i+k} - S_i \geqq a)$. By Lemma 7,

$$\mathbf{P}(S_{i+k} - S_i \geqq a, S_{i+v+k} - S_{i+v} \geqq b) \leqq \Phi_{i+v,k}(t)\exp\{-t\alpha_{i+v,k}B_{i+v,k}\}v^{-s}k^{\lambda/c} +$$
$$+ Q_{i,k}k^{t_1(u+\varepsilon)/t}v^{st_1/t}\exp\Big\{-t_1\Big(\frac{1}{t}\log\Phi_{i+k,v}(t) - \frac{1}{t_1}\log\Phi_{i+k,v}(t_1)\Big)\Big\}$$
$$\leqq e^{-k/c}v^{-s}k^{\lambda/c} + Q_{i,k}k^{u+s+\varepsilon}\exp\Big\{-t_1\sum_{r=i+k+1}^{i+k+v}\Big(\frac{1}{t}\log\phi_r(t) - \frac{1}{t_1}\log\phi_r(t_1)\Big)\Big\}.$$

Here we have used (1.3) and the inequalities $t_1 < t$, $v \leqq k$. Furthermore

$$\sum_{r=i+k+1}^{i+k+v}\Big(\frac{1}{t}\log\phi_r(t) - \frac{1}{t_1}\log\phi_r(t_1)\Big)$$
$$\geqq \frac{1}{2}(t_2 - t_1)B_{i+k,v} + \frac{1}{3!}(t_2^2 - t_1^2)\sum_{r=i+k+1}^{i+k+v}\gamma_{3,r} + \dots . \tag{3.41}$$

By (A2′) and (A3), for $l = 3, 4, \dots$,

$$\Big|\sum_{r=i+k+1}^{i+k+v}\frac{\gamma_{l,r}}{l!}\Big| \leqq C_{23}B_{i+k,v}\Big(\frac{1}{H}\Big)^l.$$

Moreover, for $0 < t_1 < t_2 < \varepsilon_2$,

$$t_2^{l-1} - t_1^{l-1} \leqq (t_2 - t_1)(l-1)\varepsilon_2^{l-2}.$$

Hence, if $0 < \varepsilon_2 < H$, the series in (3.41) can be bounded from below by $\tau B_{i+k,v}$ for some $\tau > 0$, which in view of (A4) can again be underestimated by $(\theta/t_1)\log v$ for any θ, provided $v \geqq v_1 = v_1(\theta)$.

LEMMA 9. *For any* $\varepsilon > 0$,

$$\mathbf{P}\Big(\overline{U}_N \geqq -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big) \to 1 \quad \textit{as} \quad N \to \infty.$$

PROOF. Writing

$$A_n=\left\{S_{n+K}-S_n\geqq \alpha_{n,K}B_{n,K}+\left(-\frac{1}{2}-\frac{\lambda}{c}-\varepsilon\right)\frac{\log K}{t^*_{n,K}}\right\},$$

we get by Lemma 5,

$$\mathbf{P}_1=\mathbf{P}\left(\overline{U}_N\geqq -\frac{1}{2}-\frac{\lambda}{c}-\varepsilon\right)=\mathbf{P}\left(\bigcup_{n=0}^{N-K}A_n\right)\geqq\frac{\mathbf{P}_2^2}{\mathbf{P}_2^2+\mathbf{P}_3} \tag{3.42}$$

with $\mathbf{P}_2=\sum_{n=0}^{N-K}\mathbf{P}(A_n)$ and $\mathbf{P}_3=\sum_{i\neq j}\mathbf{P}(A_iA_j)$.

By Lemma 4,

$$\mathbf{P}_2\leqq\frac{C_{19}Ne^{-K/c}}{K^{-\varepsilon-\lambda/c}}\leqq C_{21}K^{\varepsilon} \tag{3.43}$$

and

$$\mathbf{P}_2\geqq(N-K+1)C_{18}e^{-K/c}K^{\varepsilon+\lambda/c}\geqq C_{22}K^{\varepsilon} \tag{3.44}$$

for large enough N.

Via independence of the events A_i and A_j for $|i-j|>K$, we have

$$\begin{aligned}\mathbf{P}_3=&\left(\sum_{n=0}^{N-K}\mathbf{P}(A_n)\right)^2+\sum_{n=0}^{N-K}(\mathbf{P}(A_n)-(\mathbf{P}(A_n))^2)+\\&+\sum_{1\leqq|i-j|\leqq K}(\mathbf{P}(A_iA_j)-\mathbf{P}(A_i)\mathbf{P}(A_j))\leqq\mathbf{P}_2^2+\mathbf{P}_2+\mathbf{P}_5,\end{aligned} \tag{3.45}$$

where

$$\mathbf{P}_5=\sum_{i=0}^{N-K}\sum_{j=1}^{K}\mathbf{P}(A_iA_{i+j}).$$

Applying Lemma 8 with $k=K$, $v=j$, $u=1/2$, $s=2$, we get for j sufficiently large,

$$\mathbf{P}(A_iA_{i+j})\leqq e^{-K/c}j^{-2}K^{\lambda/c}+\mathbf{P}(A_i)K^{5/2+\varepsilon}j^{-\theta}.$$

Put $l=[K^{\varepsilon/2}]$. Recalling (3.43), we get

$$\mathbf{P}_5\leqq\sum_{i=0}^{N-K}\left(2l\mathbf{P}(A_i)+\sum_{j=l+1}^{K-l}\mathbf{P}(A_iA_{i+j})\right)$$

$$\leqq \sum_{i=0}^{N-K}\left(2l\mathbf{P}(A_i)+e^{-K/c}K^{\lambda/c}\sum_{j=1}^{\infty}j^{-2}+\mathbf{P}(A_i)K^{7/2+\varepsilon}l^{-\theta}\right)$$
$$\leqq C_{24}K^{3\varepsilon/2}+C_{25}+C_{26}K^{7/2+2\varepsilon}K^{-\theta\varepsilon/2}=O(K^{3\varepsilon/2}),$$

provided θ is chosen such that $(\theta+1)\varepsilon/2>7/2+2\varepsilon$.

By (3.42) and (3.45), $\mathbf{P}_1\geqq \mathbf{P}_2^2/(\mathbf{P}_2^2+\mathbf{P}_2+\mathbf{P}_5)$, where $\mathbf{P}_2=o(\mathbf{P}_2^2)$, and $\mathbf{P}_5=o(\mathbf{P}_2^2)$ as $N\to\infty$ by (3.44). This proves Lemma 9.

The assertion (i) of Theorem 1 follows from Lemmas 6 and 9.

LEMMA 10.

$$\limsup_{N\to\infty}\overline{U}_N\leqq \frac{1}{2}-\frac{\lambda}{c}\quad a.s..$$

PROOF. Choose $\varepsilon>0$. Taking Lemma 4 into account, we get

$$\mathbf{P}\left(\overline{U}_N>\frac{1}{2}-\frac{\lambda}{c}+\varepsilon\right)\leqq\sum_{n=0}^{N-K}\mathbf{P}\left(S_{n+K}-S_n\geqq\alpha_{n,K}B_{n,K}+\left(\frac{1}{2}-\frac{\lambda}{c}+\varepsilon\right)\frac{\log K}{t^*_{n,K}}\right)$$
$$\leqq\frac{C_{19}Ne^{-K/c}}{K^{1+\varepsilon-\lambda/c}}\leqq C_{27}K^{-(1+\varepsilon)}.$$

For any natural j put $N_j=\max\{N:[c\log N+\lambda\log\log N]=j\}$. We have

$$\mathbf{P}\left(\overline{U}_{N_j}>\frac{1}{2}-\frac{\lambda}{c}+\varepsilon\right)\leqq C_{27}j^{-(1+\varepsilon)}$$

and the series $\sum_{j=1}^{\infty}\mathbf{P}\left(\overline{U}_{N_j}>\frac{1}{2}-\frac{\lambda}{c}+\varepsilon\right)$ converges for any $\varepsilon>0$. By the Borel–Cantelli lemma and the inequalities $\overline{U}_N\leqq\overline{U}_{N_j}$ for $N_{j-1}<N\leqq N_j$, we complete the proof of Lemma 10.

LEMMA 11.

$$\liminf_{N\to\infty}\overline{U}_N\leqq -\frac{1}{2}-\frac{\lambda}{c}\quad a.s..$$

PROOF. It follows from Lemma 6.

LEMMA 12.

$$\limsup_{N\to\infty}\overline{U}_N\geqq \frac{1}{2}-\frac{\lambda}{c}\quad a.s..$$

PROOF. Choose $\varepsilon>0$. For $j=1,2,\dots$ we define $N_j=\min\{N:[c\log N+\lambda\log\log N]=j\}$ and put

$$R_j=\max_{N_{j-1}<n\leqq N_j-j}\frac{t^*_{n,j}(S_{n+j}-S_n-\alpha_{n,j}B_{n,j})}{\log j}.$$

Evidently,

$$\left\{R_j \geqq \frac{1}{2} - \frac{\lambda}{c} - \varepsilon\right\} = \bigcup_{i=1}^{M} A_i, \quad \text{where} \quad A_i = \{S_{m+i+j} - S_{m+i} \geqq x_{m+i}\},$$

$$m = N_{j-1} \quad x_n = \alpha_{n,j} B_{n,j} + \left(\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\right) \frac{\log j}{t^*_{n,j}}, \quad M = N_j - N_{j-1} - j.$$

Without loss of generality we can take $m = 0$. Note that

$$C_{28} e^{j/c} j^{-\lambda/c} \leqq M \leqq C_{29} e^{j/c} j^{-\lambda/c}$$

for all large j.

An application of Lemma 5 yields

$$\mathbf{P}\left(R_j \geqq \frac{1}{2} - \frac{\lambda}{c} - \varepsilon\right) = \mathbf{P}\left(\bigcup_{i=1}^{M} A_i\right) \geqq \frac{\mathbf{P}_6^2}{\mathbf{P}_6 + \mathbf{P}_7},$$

where $\mathbf{P}_6 = \sum_{i=1}^{M} \mathbf{P}(A_i)$, $\mathbf{P}_7 = \sum_{i \neq r} \mathbf{P}(A_i A_r)$.

By Lemma 4,

$$\mathbf{P}_6 \leqq C_{19} M e^{-j/c} j^{-1+\lambda/c+\varepsilon} \leqq C_{30} j^{-1+\varepsilon} \tag{3.46}$$

and

$$\mathbf{P}_6 \geqq M C_{18} e^{-j/c} j^{-1+\lambda/c+\varepsilon} \geqq C_{31} j^{-1+\varepsilon}. \tag{3.47}$$

As in (3.45) we have

$$\mathbf{P}_7 \leqq \mathbf{P}_6^2 + \mathbf{P}_6 + \sum_{i=1}^{M} \sum_{r=1}^{j} \mathbf{P}(A_i A_{i+r}). \tag{3.48}$$

An application of Lemma 8 with $k = j$, $v = r$, $u = -1/2$, $s = 1 + 2/\varepsilon$, gives

$$\mathbf{P}(A_i A_{i+r}) \leqq e^{-j/c} r^{-(1+2/\varepsilon)} j^{\lambda/c} + \mathbf{P}(A_i) j^{1/2+2/\varepsilon+\varepsilon} r^{-\theta},$$

for large enough r, where θ is an arbitrary positive constant.

Put $l = [j^{\varepsilon/2}]$. We have

$$\sum_{i=1}^{M} \sum_{r=1}^{j} \mathbf{P}(A_i A_{i+r}) = \sum_{i=1}^{M} \left(\sum_{r=1}^{l-1} \mathbf{P}(A_i A_{i+r}) + \sum_{r=l}^{j} \mathbf{P}(A_i A_{i+r})\right)$$

$$\leqq \sum_{i=1}^{M} \left(l \mathbf{P}(A_i) + e^{-j/c} j^{\lambda/c} \sum_{r=l}^{j} r^{-(1+2/\varepsilon)} + \mathbf{P}(A_i) j^{3/2+2/\varepsilon+\varepsilon} l^{-\theta}\right)$$

(3.49)

$$\leqq C_{30} j^{-1+3\varepsilon/2} + M e^{-j/c} l^{-2/\varepsilon} j^{\lambda/c} \sum_{r=1}^{j} r^{-1} + C_{32} j^{-1+\varepsilon} j^{3/2+2/\varepsilon+\varepsilon} j^{-\theta\varepsilon/2}$$
$$\leqq C_{30} j^{-1+3\varepsilon/2} + C_{29} j^{-1} \log j + C_{32} j^{1/2+2/\varepsilon+2\varepsilon} j^{-\theta\varepsilon/2} \leqq C_{33} j^{-1+3\varepsilon/2},$$

since θ can be chosen arbitrarily large. Here we used the definition of l, (3.46), (3.47) and the definition of M.

By (3.46)–(3.49)

$$\mathbf{P}_7 \leqq C_{30}^2 j^{-2+2\varepsilon} + C_{30} j^{-1+\varepsilon} + C_{33} j^{-1+3\varepsilon/2} \leqq C_{34} j^{-1+3\varepsilon/2} \tag{3.50}$$

if $\varepsilon < 2$ (without loss of generality).

From (3.46), (3.47) and (3.50), we obtain

$$\mathbf{P}\Big(R_j \geqq \frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big) = \frac{\mathbf{P}_6^2}{\mathbf{P}_7} \geqq C_{35} j^{-1+\varepsilon/2}$$

for large enough j.

Hence the series $\sum_{j=1}^{\infty} \mathbf{P}\Big(R_j \geqq \frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big)$ diverges for any $\varepsilon > 0$. Since the random variables R_j $(j = 1, 2, \dots)$ are independent, an application of the Borel–Cantelli lemma combined with $R_j \leqq \overline{U}_N$, for $N_j \leqq N < N_{j+1}$ and $j = K$, completes the proof of Lemma 12.

LEMMA 13.

$$\liminf_{N\to\infty} \overline{U}_N \geqq -\frac{1}{2} - \frac{\lambda}{c} \quad a.s..$$

PROOF. Take $\varepsilon > 0$. For any natural j, we put

$$N_j = \min\{N : [c \log N + \lambda \log\log n] = j\}.$$

Define $J_j = \{m : m = r[j^{\varepsilon/2}],\ r = 1, 2, \dots\}$, $L = \max\{l : (2l+1)j - 1 \leqq N_j - j\}$.

For any j, the random variables

$$Q_l = \max_{2lj \leqq i < (2l+1)j,\, i \in J_j} \frac{t^*_{i,j}(S_{i+j} - S_i - \alpha_{i,j} B_{i,j})}{\log j}, \qquad l = 0, 1, \dots, L,$$

are independent and $\overline{U}_{N_j} \geqq \max_{0 \leqq l \leqq L} Q_l$. Hence

$$\mathbf{P}_8 = \mathbf{P}\Big(\overline{U}_N < -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big) \leqq \prod_{l=0}^{L} \mathbf{P}\Big(Q_l < -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big).$$

Putting

$$A_i = \left\{ S_{i+j} - S_i \geqq \alpha_{i,j} B_{i,j} + \left(-\frac{1}{2} - \frac{\lambda}{c} - \varepsilon \right) \frac{\log j}{t^*_{i,j}} \right\},$$

we get

$$\mathbf{P}\left(Q_l \geqq -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon \right) = \mathbf{P}\left(\bigcup_{i \in I_l} A_i \right) \geqq \Sigma_1 - \Sigma_2,$$

where

$$\Sigma_1 = \sum_{i \in I_l} \mathbf{P}(A_i),$$

$$\Sigma_2 = \sum_{i \neq p; i,p \in I_l} \mathbf{P}(A_i A_p), \quad I_l = \{ i : 2lj \leqq i < (2l+1)j, i \in J_j \}.$$

We denote the cardinality of I_l by M_l. Then $M_l \sim j^{1-\varepsilon/2}$ as $j \to \infty$ for all l.

By Lemma 4,

$$\Sigma_1 \geqq M_l C_{18} e^{-j/c} j^{\lambda/c} j^{\varepsilon} \geqq C_{36} e^{-j/c} j^{\lambda/c} j^{1+\varepsilon/2}$$

for large enough j.

We have

$$\Sigma_2 \leqq \sum_{i \in I_l} \sum_{r=1}^{M_l - 1} \mathbf{P}(A_i A_{i+r[j^{\varepsilon/2}]}).$$

Applying Lemma 8 with $k = j$, $v = m$, $u = 1/2$, $s = 2$, and Lemma 4, we conclude

$$\begin{aligned} \mathbf{P}(A_i A_{i+m}) &\leqq e^{-j/c} m^{-2} j^{\lambda/c} + \mathbf{P}(A_i) j^{5/2+\varepsilon} m^{-\theta} \\ &\leqq C_{37} e^{-j/c} r^{-2} j^{\lambda/c} j^{-\varepsilon} + C_{38} e^{-j/c} j^{\lambda/c} j^{5/2+2\varepsilon} r^{-\theta} j^{-\theta\varepsilon/2} \end{aligned}$$

for large enough j, where θ is an arbitrary positive constant. Hence

$$\begin{aligned} \sum_{r=1}^{M_l - 1} \mathbf{P}(A_i A_{i+r[j^{\varepsilon/2}]}) &\leqq C_{39} e^{-j/c} j^{\lambda/c} j^{-\varepsilon} + C_{38} M_l e^{-j/c} j^{\lambda/c} j^{5/2+2\varepsilon} j^{-\theta\varepsilon/2} \\ &= o(j^{\lambda/c} e^{-j/c}) = o(j^{\lambda/c} j^{\varepsilon} e^{-j/c}) \end{aligned}$$

as $j \to \infty$, since θ can be chosen arbitrarily large.

Thus

$$\Sigma_2 = o(j^{\lambda/c} j^{1+\varepsilon/2} e^{-j/c}) = o(\Sigma_1) \quad \text{as} \quad j \to \infty,$$

and for large enough j,

$$\mathbf{P}\Big(Q_l \geqq -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big) \geqq C_{40} j^{1+\varepsilon/2} j^{\lambda/c} e^{-j/c}.$$

We have

$$\begin{aligned}\mathbf{P}_8 &\leqq (1 - C_{40} j^{1+\varepsilon/2} j^{\lambda/c} e^{-j/c})^{L+1} \\ &\leqq \exp\{-C_{40}(L+1) j^{1+\varepsilon/2} j^{\lambda/c} e^{-j/c}\} \leqq \exp\{-C_{41} j^{\varepsilon/2}\}\end{aligned}$$

for large enough j, where we used the inequality $1 - x \leqq \exp\{-x\}$ and

$$L \geqq \frac{N_j}{2j} - 2 \geqq C_{42} j^{-1} j^{-\lambda/c} e^{j/c},$$

for large enough j.

Hence the series $\sum_{j=1}^{\infty} \mathbf{P}\Big(\overline{U}_{N_j} < -\frac{1}{2} - \frac{\lambda}{c} - \varepsilon\Big)$ converges. By the Borel–Cantelli lemma and the inequalities $\overline{U}_N \geqq \overline{U}_{N_j}$ for $N_j \leqq N < N_{j+1}$, we get the conclusion of Lemma 13.

Assertions (ii) and (iii) of Theorem 1 follow from Lemmas 10, 11, 12, 13.

The last statement of Theorem 1 concerning W_N and T_N can be proved in a similar way.

Acknowledgements. This research was started while the third author was on sabbatical leave from Philipps-University Marburg and visiting the University of St. Petersburg. It was completed later, when the first two authors in turn visited Marburg.

All authors gratefully acknowledge the support by a special research grant from the University of Marburg (Hochschulsonderprogramm II) which made this exchange possible. The first two authors also acknowledge partial support of this research from ISF (grant JFD100).

REFERENCES

[1] Bacro, J.-N., Encadrements limites de lois de type Erdős–Rényi–Shepp, *Publ. Inst. Statist. Univ. Paris* **31** (1986), 1–13. *MR* **88k**:60056

[2] Bacro, J.-N., Deheuvels, P. and Steinebach, J., Exact convergence rates in Erdős–Rényi type theorems for renewal processes, *Ann. Inst. H. Poincaré Probab. Statist.* **23** (1987), 195–207. *MR* **88k**:60055

[3] Chung, K. L. and Erdős, P., On the application of the Borel–Cantelli lemma, *Trans. Amer. Math. Soc.* **72** (1952), 179–186. *MR* **13**, 567b

[4] Csáki, E., On the increments of additive functionals, *Studia Sci. Math. Hungar.* **26** (1991), 185–199. *MR* **93i**:60039

[5] Csáki, E. and Csörgő, M., Inequalities for increments of stochastic processes and moduli of continuity, *Ann. Probab.* **20** (1992), 1031–1052. *MR* **93c**:60055

[6] Csáki, E., Földes, A. and Komlós, J., Limit theorems for Erdős–Rényi type problems, *Studia Sci. Math. Hungar.* **22** (1987), 321–332. *MR* **89b**:60052

[7] CSÖRGŐ, M. and STEINEBACH, J., Improved Erdős–Rényi and strong approximation laws for increments of partial sums, *Ann. Probab.* **9** (1981), 988–996. *MR* **82k**:60063

[8] CSÖRGŐ, S., Erdős–Rényi laws, *Ann. Statist.* **7** (1979), 772–787. *MR* **80g**:60027

[9] DEHEUVELS, P., Functional Erdős–Rényi laws, *Studia Sci. Math. Hungar.* **26** (1991), 261–295. *MR* **93i**:60056

[10] DEHEUVELS, P. and DEVROYE, L., Limit laws of Erdős–Rényi–Shepp type, *Ann. Probab.* **15** (1987), 1363–1386. *MR* **88f**:60055

[11] DEHEUVELS, P., DEVROYE, L. and LYNCH, J., Exact convergence rate in the limit theorems of Erdős–Rényi and Shepp, *Ann. Probab.* **14** (1986), 209–223. *MR* **87d**:60032

[12] DEHEUVELS, P. and STEINEBACH, J., Sharp rates for increments of renewal processes, *Ann. Probab.* **17** (1989), 700–722. *MR* **90a**:60058

[13] ERDŐS, P. and RÉNYI, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42**#6907

[14] FROLOV, A. N., Exact convergence rate in the Erdős–Rényi law of large numbers for independent non-identically distributed random variables, *Vestnik St. Petersburg Univ., Ser. 1*, **26** (1993), No. 1, 60–61.

[15] FROLOV, A. N., Asymptotic behaviour of the increments of sums of independent random variables, *Vestnik St. Petersburg Univ., Ser. 1*, **26** (1993), No. 4, 52–56.

[16] FROLOV, A. N., Erdős–Rényi laws, Ph. D. Thesis, St. Petersburg University, 1993.

[17] LIN, Z. Y., The Erdős–Rényi laws of large numbers for nonidentically distributed random variables, *Chinese Ann. Math. Ser. B.* **11** (1990), 376–383. *MR* **92a**:60083

[18] PETROV, V. V., *Sums of independent random variables*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82, Springer-Verlag, Berlin–Heidelberg–New York, 1975. *MR* **52** #9335

[19] PRIVALOV, I. I., *Introduction to the theory of functions of a complex variable*, Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow–Leningrad, 1948, 1954 (in Russian). *MR* **13**, 334c; **16**, 121i

[20] ROZOVSKY, L. V., On the convergence rate in the Lindeberg–Feller theorem, *Vestnik Leningrad Univ. Math.* **7** (1979), 83–89.

[21] SHEPP, L. A., A limit law concerning moving averages, *Ann. Math. Statist.* **35** (1964), 424–428. *MR* **29** #4091

[22] STEINEBACH, J., Improved Erdős–Rényi and strong approximation laws for increments of renewal processes, *Ann. Probab.* **14** (1986), 547–559. *MR* **87k**:60090

[23] STEINEBACH, J., On general versions of Erdős–Rényi laws, *Z. Wahrsch. Verw. Gebiete* **56** (1981), 549–554. *MR* **82g**:60047

(Received April 23, 1996)

A. Frolov and A. Martikainen

DEPARTMENT OF MATHEMATICS AND MECHANICS
ST. PETERSBURG UNIVERSITY
BIBLIOTECHNAYA PL. 2
ST. PETERHOF
198904 SANKT PETERSBURG
RUSSIA

marti@marti.niimm.spb.su

J. Steinebach

FACHBEREICH MATHEMATIK
PHILIPPS-UNIVERSITÄT
HANS-MEERWEIN-STRASSE (LAHNBERGE)
D–35032 MARBURG
GERMANY

jost@mathematik.uni-marburg.de

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 153–166

EXTENSIONS OF BONFERRONI TYPE INEQUALITIES

B. R. HANDA and S. G. MOHANTY

Dedicated to Professor E. Csáki for his sixtieth birthday

Abstract

The classical Bonferroni type inequalities are given in a probabilistic framework in terms of joint occurrence of events. We provide two extensions of such type of inequalities, one being a multivariate case and the other being a general case, which go beyond the usual probabilistic interpretation.

1. Introduction

Consider n arbitrary events $A_1, A_2, \dots, A_n$ in the probability space $(S, \mathfrak{F}, \mathbf{P})$. Let $B_{n,t}$, $0 \leqq t \leqq n$ denote the event that exactly t among these occur. Let $\mathbf{P}(B_{n,t}) = \mathbf{P}_{[t]}$ and $\mathbf{P}_{(t)} = \sum\limits_{r=t}^{n} \mathbf{P}_{[r]}$. A well-known generalization of inclusion-exclusion principle leads to the following expressions:

$$\mathbf{P}_{[t]} = \sum_{j=t}^{n} (-1)^{j-t} \binom{j}{t} S_{j,n}, \tag{1}$$

$$\mathbf{P}_{(t)} = \sum_{j=t}^{n} (-1)^{j-t} \binom{j-1}{t-1} S_{j,n}, \tag{2}$$

where

$$S_{j,n} = \sum_{1 \leqq i_1 < i_2 < \cdots < i_j \leqq n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}), \qquad S_{0,n} = 1.$$

In view of the relations

$$S_{j,n} = \sum_{s=j}^{n} \binom{s}{j} \mathbf{P}_{[s]} \quad \text{and} \quad S_{j,n} = \sum_{s=j}^{n} \binom{s-1}{j-1} \mathbf{P}_{(s)},$$

1991 *Mathematics Subject Classification.* Primary 60C05; Secondary 05A20.

Key words and phrases. Inclusion-exclusion principle, Bonferroni inequalities, Pólya frequency sequence, inverse relation.

(1) and (2) follow from the inversion relation given below.

For sequences $\{y_k\}_{0\leqq k\leqq n}$ and $\{w_k\}_{0\leqq k\leqq n}$ of numbers,

$$y_k = \sum_{s=k}^{n} \binom{s}{k} w_s, \tag{3}$$

if and only if

$$w_k = \sum_{s=k}^{n} (-1)^{s-k} \binom{s}{k} y_s. \tag{4}$$

A number of statistical applications require upper and lower bounds for $\mathbf{P}_{[t]}$ and $\mathbf{P}_{(t)}$ in term of linear combinations of $S_{j,n}$. These bounds are called Bonferroni type inequalities. It is well known that if partial sums are taken in (1) and (2) the following inequalities will result:

$$\sum_{j=t}^{t+2u+1} (-1)^{j-t} \binom{j}{t} S_{j,n} \leqq \mathbf{P}_{[t]} \leqq \sum_{j=t}^{t+2u} (-1)^{j-t} \binom{j}{t} S_{j,n}, \tag{5}$$

and

$$\sum_{j=t}^{t+2u+1} (-1)^{j-t} \binom{j-1}{t-1} S_{j,n} \leqq \mathbf{P}_{(t)} \leqq \sum_{j=t}^{t+2u} (-1)^{j-t} \binom{j-1}{t-1} S_{j,n}. \tag{6}$$

These inequalities are extensions of classical Bonferroni inequalities (Gumbel [11], Takács [17]). The proofs of these inequalities as well as some of the bounds that improve upon these, depend on the relation of the type

$$\mathbf{P}_{[t]} = \sum_{j=t}^{t+a} (-1)^{j-t} \binom{j}{t} S_{j,n} + (-1)^{a+1} \delta(t,a,n), \tag{7}$$

and

$$\mathbf{P}_{(t)} = \sum_{j=t}^{t+a} (-1)^{j-t} \binom{j-1}{t-1} S_{j,n} + (-1)^{a+1} \delta^*(t,a,n), \tag{8}$$

for $a \leqq n-t$ where δ and δ^* ($\geqq 0$) are the remainder terms (see (1) and (2)). If a lower bound for remainder $\delta(t,a,n)$, say $\delta_1(t,a,n)$, where $\delta_1(t,a,n)$ depends on $S_{j,n}$, $j = t,\ t+1, \ldots, t+a+1$, can be obtained, then (7) provides inequalities

$$\begin{aligned} &\sum_{j=t}^{t+2u+1} (-1)^{j-t} \binom{j}{t} S_{j,n} + \delta_1(t, 2u+1, t) \leqq \mathbf{P}_{[t]} \\ &\leqq \sum_{j=t}^{t+2u} (-1)^{j-t} \binom{j}{t} S_{j,n} + \delta_1(t, 2u, n) \end{aligned} \tag{9}$$

which are sharper inequalities than (5). Similarly, improved bounds over (6) can be written for $\mathbf{P}_{(t)}$ by starting with (8).

Two different forms of remainder terms δ and δ^* are available in literature. The first one we shall refer to as Galambos' form of remainders (see Galambos [4]), which as observed in Recsei and Seneta [6] is applicable to general sequences $\{y_k\}$ and $\{w_k\}$ related through (3) and (4) and has the following form

$$w_t = \sum_{j=t}^{t+a}(-1)^{j-t}\binom{j}{t}y_j + (-1)^{a+1}\sum_{s=t+a+1}^{n}\binom{s-t-1}{a}\binom{s}{t}w_s. \tag{10}$$

Both (7) and (8) follow from (10) by choosing $\{w_t\}$ and $\{y_t\}$ appropriately. In fact, for (7) we set $w_t = P_{[t]}$ and $y_j = S_{j,n}$, while for (8) we set $w_t = P_{(t)}/t$ and $y_j = S_{j,n}/j$.

Following Galambos [4] we conclude that

$$\begin{aligned}\delta(t,a,n) &= \sum_{s=t+a+1}^{n}\binom{s-t-1}{a}\binom{s}{t}w_s \\ &\geq \frac{a+1}{n-t}\binom{a+t+1}{t}y_{a+t+1} \quad (=\delta_1(t,a,n)).\end{aligned} \tag{11}$$

Thus (9) is true with the value of $\delta_1(t,a,n)$ given by (11) where $y_j = S_{j,n}$. By similar arguments one can get improved inequality for $P_{(t)}$ using (10) and (11).

The second form of remainder terms is due to Hoppe [12] which arises out of a method of iteration proposed in Hoppe and Seneta [13]. Hoppe's form of remainders is in terms of multiple summation of probabilities of unions and intersections of events $A_1, A_2, \ldots, A_n$. The derivation being solely in probabilistic framework, does not lead to any apparent analogue for sequences $\{w_t\}$ and $\{y_t\}$ similar to (10).

In this article we shall concentrate on Galambos' form of remainder terms and his technique outlined above for constructing Bonferroni type inequalities. This essentially consists of first writing the numbers in the remainder form (10) and then looking for a suitable lower bound for the remainder term which when used in (10) gives improved bounds for the required numbers. This technique will be extended to multivariable case in Section 2 to obtain Bonferroni type bounds for probability of at least some given number of events from each of several classes of events in the probability space. As is apparent from our earlier description the probabilistic setting is of no significance after the numbers involved have been shown to follow the relation of the type (3). In Section 3, we show that the Galambos form of remainders can be extended to write Bonferroni type bounds for w_t when binomial coefficients in (3) are replaced by numbers forming a Pólya frequency sequence.

2. Multivariate extensions of Bonferroni type inequalities

Let $\{A_1, A_2, \ldots, A_M\}$ and $\{B_1, B_2, \ldots, B_N\}$ be two arbitrary classes of events in the probability space $(S, \mathfrak{F}, \mathbf{P})$. Let $B_{M,N;m,n}$, $0 \leqq m \leqq M$, $0 \leqq n \leqq N$ be the event that exactly m among M A_is and n among N B_is occur. Let $\mathbf{P}_{[m,n]} = \mathbf{P}(B_{M,N;m,n})$ and $\mathbf{P}_{(m,n)} = \sum_{r=m}^{M} \sum_{s=n}^{m} \mathbf{P}_{[r,s]}$.

Let

$$S_{m,n;M,N} = \sum_{\substack{1 \leqq i_1 < i_2 < \cdots < i_m \leqq M \\ 1 \leqq j_i < j_2 < \cdots < j_n \leqq N}} \mathbf{P}\Big(\bigcap_{t=1}^{m} A_{i_t}, \bigcap_{p=1}^{n} B_{j_p}\Big).$$

Starting with relation of Fréchet [3] given by

$$\mathbf{P}_{[m,n]} = \sum_{t=m+n}^{M+N} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i}{m} \binom{j}{n} S_{i,j;M,N} \tag{12}$$

and the inverse relation

$$S_{m,n;M,N} = \sum_{t=m+n}^{M+N} \sum_{i+j=t} \binom{i}{m} \binom{j}{n} \mathbf{P}_{[i,j]}. \tag{13}$$

Meyer [15] obtained the bivariate form of Bonferroni inequalities generalizing (5) as below.

$$\begin{aligned} &\sum_{t=m+n}^{m+n+2k+1} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i}{m} \binom{j}{n} S_{i,j;M,N} \leqq P_{[m,n]} \\ &\leqq \sum_{t=m+n}^{m+n+2k} (-1)^{t-(m+n)} \binom{i}{m} \binom{j}{n} S_{i,j;M,N} \end{aligned} \tag{14}$$

where $k \leqq (M' + N' - 1)/2$ is a non-negative integer and $M' = M - m$, $N' = N - n$. Analogous results for $\mathbf{P}_{(m,n)}$ are

$$\mathbf{P}_{(m,n)} = \sum_{t=m+n}^{M+N} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i-1}{m-1} \binom{i-1}{n-1} S_{i,j;M,N}, \tag{15}$$

$$S_{i,j;M,N} = \sum_{t=m+n}^{M+N} \sum_{i+j=t} \binom{i-1}{m-1} \binom{j-1}{n-1} \mathbf{P}_{(m,n)}, \tag{16}$$

and

(17)
$$\sum_{t=m+n}^{m+n+2k+1} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i-1}{m-1}\binom{j-1}{n-1} S_{i,j;M,N} \leqq \mathbf{P}_{(m,n)}$$
$$\leqq \sum_{t=m+n}^{m+n+2k} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i-1}{m-1}\binom{j-1}{n-1} S_{i,j;M,N}.$$

It is now possible to improve upon these bounds by deriving the Galambos form of remainders for bivariate case analogous to (7) and (8) or more specifically (9) of one variate case. First, note that the following inversion relations hold: For any double sequences $\{y_{m,n}\}$, $\{w_{m,n}\}$, $0 \leqq m \leqq M$, $0 \leqq n \leqq N$,

(18)
$$y_{m,n} = \sum_{i=m}^{M} \sum_{j=n}^{N} \binom{i}{m}\binom{j}{n} w_{i,j}$$

if and only if

(19)
$$w_{m,n} = \sum_{i=m}^{M} \sum_{j=n}^{N} (-1)^{i+j-(m+n)} \binom{i}{m}\binom{j}{n} y_{i,j}.$$

Next for any $0 \leqq a \leqq M' + N'$, we can write

(20)
$$w_{m,n} = \sum_{t=m+n}^{m+n+a} \sum_{\substack{i+j=t \\ m \leqq i \leqq M \\ n \leqq j \leqq N}} (-1)^{t-(m+n)} \binom{i}{m}\binom{j}{n} y_{i,j} + (-1)^{a+1} \delta(m,n;a;M,N)$$

where

(21)
$$\delta(m,n;a;M,N) = \sum_{t=m+n+a+1}^{M+N} \sum_{\substack{i+j=t \\ m \leqq i \leqq M \\ n \leqq j \leqq N}} (-1)^{t-(m+n)-a-1} \binom{i}{m}\binom{j}{n} y_{i,j}.$$

Substituting for y_{ij} from (18) into (21) we have

$$\delta(m,n;a;M,N)$$
$$= \sum_{u=a+1}^{M'+N'} \sum_{\substack{x+y=u \\ 0 \leqq x \leqq M' \\ 0 \leqq y \leqq N'}} (-1)^{u-a-1} \binom{x+m}{m}\binom{y+n}{n} \sum_{p=x}^{M'} \sum_{q=y}^{N'} \binom{p+m}{x+m}\binom{q+n}{y+n} w_{p+m,q+n}$$

(22)

$$= \sum_{p=0}^{M'} \sum_{q=0}^{N'} \binom{p+m}{m} \binom{q+n}{n} w_{p+m,q+n} (-1)^{a+1} \sum\sum_{\substack{a+1 \leqq x+y \leqq p+q \\ 0 \leqq x \leqq p \\ 0 \leqq y \leqq q}} (-1)^{x+y} \binom{p}{x} \binom{q}{y}.$$

Now for any positive integer $u = x + y$ we have the relations (see Galambos [4])

$$\sum_{x=0}^{u} (-1)^x \binom{p}{x} = - \sum_{x=u+1}^{p} (-1)^x \binom{p}{x} = (-1)^u \binom{p-1}{u}.$$

Using these, it follows that

$$\sum\sum_{\substack{a+1 \leqq x+y \leqq p+q \\ 0 \leqq x \leqq p \\ 0 \leqq y \leqq q}} (-1)^{x+y} \binom{p}{x} \binom{q}{y} = (-1)^{a+1} \sum_{y=a+1-p}^{q} \binom{q}{y} \binom{p-1}{a-y}$$

$$= (-1)^{a+1} \binom{p+q-1}{a}.$$

Substituting the last expression into (22) we finally get the remainder term as

$$\delta(m, n; a; M, N)$$

(23)
$$= \sum\sum_{\substack{a+1 \leqq p+q \leqq M'+N' \\ 0 \leqq p \leqq M' \\ 0 \leqq q \leqq N'}} \binom{p+m}{m} \binom{q+n}{n} \binom{p+q-1}{a} w_{p+m,q+n}$$

which is clearly non-negative if $w_{i,j}$'s are assumed to be non-negative.

Thus we have the Galambos form of remainder representation as given by (20) and (23). It gives rise to a bivariate extension of Bonferroni type of inequalities.

THEOREM 1.

(24)
$$\sum_{t=m+n}^{m+n+2k+1} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i}{m} \binom{j}{n} y_{i,j} \leqq w_{m,n}$$

$$\leqq \sum_{t=m+n}^{m+n+2k} \sum_{i+j=t} (-1)^{t-(m+n)} \binom{i}{m} \binom{j}{n} y_{i,j}$$

for $1 \leqq m \leqq M$, $1 \leqq n \leqq N$ *and* $k \leqq \frac{1}{2}(M' + N' - 1)$.

We obtain (12), (13) and (14) by setting $w_{i,j} = \mathbf{P}_{[i,j]}$ and $y_{i,j} = S_{i,j;M,N}$ in (18), (19) and (24), whereas (15), (16) and (17) follow by setting $w_{ij} = \mathbf{P}_{(i,j)}/ij$ and $y_{ij} = S_{i,j;M,N}/ij$, $i,j \geqq 1$ in (18), (19) and (24).

These inequalities can be further improved by following Galambos' approach. For this, let

$$\delta^*(m,n;a;M,N) = \sum_{\substack{x+y=a+1 \\ 0\leqq x\leqq M' \\ 0\leqq y\leqq N'}} \binom{x+m}{m}\binom{y+n}{n} y_{x+m,y+n} \tag{25}$$

where $a+1 \leqq M'+N'$. We show that

$$\delta(m,n;a;M,N) \geqq \frac{a+1}{M'+N'}\delta^*(m,n;a;M,N). \tag{26}$$

For proving (26) we note that carrying out manipulations on (25) similar to that which was done on (21) to bring it to the form (23), we will get

$$\begin{aligned}&\delta^*(m,n;a;M,N)\\ &=\sum_{\substack{a+1\leqq p+q\leqq M'+N' \\ 0\leqq p\leqq M' \\ 0\leqq q\leqq N'}} \binom{p+m}{m}\binom{q+n}{n}\frac{p+q}{a+1}\binom{p+q-1}{a} w_{p+m,q+n}\end{aligned}$$

which is $\leqq \dfrac{M'+N'}{a+1}\delta(m,n;a;M,N)$. Hence (26) is true.

Finally, using lower bound for δ given by (26) in relation (20) yields the improved Bonferroni type inequalities given below.

THEOREM 2.

$$\begin{aligned}&\sum_{t=m+n}^{m+n+2k+1}\sum_{i+j=t}(-1)^{t-(m+n)}\binom{i}{m}\binom{j}{n}y_{i,j} + \frac{2k+2}{M'+N'}\delta^*(m,n;k+1;M,N)\\ &\leqq w_{m,n} \leqq \sum_{t=m+n}^{m+n+2k}\sum_{i+j=t}(-1)^{t-(m+n)}\binom{i}{m}\binom{j}{n}y_{i,j} - \frac{2k+2}{M'+N'}\delta^*(m,n;2k;M,N)\end{aligned} \tag{27}$$

for $1\leqq m\leqq M$, $1\leqq n\leqq N$, $k\leqq \frac{1}{2}(M'+N'-1)$ *and* δ^* *given by* (25).

The improved Bonferroni bounds on $\mathbf{P}_{[m,n]}$ and $\mathbf{P}_{(m,n)}$ follows from (27) by particular choices of $w_{m,n}$ and $y_{i,j}$ as discussed before.

Another type of extension of Bonferroni type inequalities is possible in the bivariate situation, by considering an alternative of type (20) with a different remainder term. The derivations follow argument similar to those used for obtaining (20) with remainder in the form (23) and therefore we state the result below without derivation:

$$\begin{aligned} w_{m,n} = &\sum_{x=0}^{a}\sum_{y=0}^{b}(-1)^{x+y}\binom{x+m}{m}\binom{y+n}{n}y_{x+m,y+n} \\ &+(-1)^{a+b+1}\delta_1(m,n;a,b;M,N)+(-1)^{a+1}\delta_2(m,n;a;M,N) \\ &+(-1)^{b+1}\delta_3(m,n;b;M,N) \end{aligned} \tag{28}$$

where

$$\delta_1(m,n;a,b;M,N) = \sum_{p=a+1}^{M'}\sum_{q=b+1}^{N'}\binom{p+m}{m}\binom{q+n}{n}\binom{p-1}{a}\binom{q-1}{b}w_{p+m,q+n},$$

$$\delta_2(m,n;a;M,N) = \sum_{p=a+1}^{M'}\binom{p+m}{m}\binom{p-1}{a}w_{p+m,N},$$

$$\delta_3(m,n;b;M,N) = \sum_{q=b+1}^{N'}\binom{q+n}{n}\binom{q-1}{b}w_{M,q+n}.$$

Now starting with equation (28) and choosing a, b suitably as odd or even, it is possible to construct new Bonferroni inequalities by using the following bounds for δ_1, δ_2 and δ_3:

$$\begin{aligned} &\frac{a+1}{M'}\binom{a+m+1}{m}\frac{b+1}{N'}\binom{b+n+1}{n}y_{a+m+1,b+n+1} \leqq \delta_1(m,n;a,b;M,N) \\ &\leqq \binom{a+m+1}{m}\binom{b+m+1}{n}y_{a+m+1,b+n+1}, \end{aligned}$$

$$\delta_2(m,n;a;M,N) \geqq \frac{a+1}{M'}\binom{a+m+1}{m}y_{a+m+1,N},$$

and

$$\delta_3(m,n;b;M,N) \geqq \frac{b+1}{N'}\binom{b+n+1}{n}y_{M,b+n+1}.$$

As an illustration, we will arrive at the following inequality by choosing both a and b as even.

THEOREM 3.

$$
\begin{aligned}
w_{m,n} \leqq \sum_{i=m}^{m+2t} \sum_{j=n}^{n+2u} (-1)^{i+j-(m+n)} \binom{i}{m} \binom{j}{n} y_{i,j} \\
- \frac{2t+1}{M'} \binom{m+2t+1}{m} y_{m+2t+1,N} \\
- \frac{2u+1}{N'} \binom{n+2u+1}{n} y_{M,n+2u+1} \\
- \frac{2t+1}{M'} \binom{m+2t+1}{m} \frac{2u+1}{N'} \binom{n+2u+1}{n} y_{m+2t+1,n+2u+1}
\end{aligned} \tag{29}
$$

for $1 < m \leqq M$, $1 \leqq n \leqq N$, $t < \frac{1}{2}(M'-1)$, $u \leqq \frac{1}{2}(N'-1)$.

All the derivations in this section have been carried out for bivariate situation for simplicity of presentation. The extension to more than two classes of events is now obvious and hence is not discussed here.

At the end of this section we remark that it may be of interest to examine which form of the remainders (20) or (28) can yield sharper bounds. However, from a practical point of view upper bound in (27) may be more useful than the bound given by (29). In (29) terms $y_{m+2t+1,N}$ and $y_{M,n+2u+1}$ are present which involve $S_{m+2t-1,N;M,N}$ and $S_{M,n+2u+1;M,N}$ when bounds for $\mathbf{P}_{[m,n]}$ or $\mathbf{P}_{(m,n)}$ are required. In this case, the difficulty arises if probabilities of events $\bigcap_{i=1}^{M} A_i$ or $\bigcap_{i=1}^{N} B_i$ are not available.

For further references on bivariate and multivariate Bonferroni inequalities one may see Chen and Seneta [2], Galambos and Lee [5], [6], Galambos and Simonelli [7], Galambos and Xu [8]–[10], and Lee [14].

3. Bonferroni type bounds for a class of sequences

In this section we provide an extension of Galambos' technique from completely non-probabilistic point of view. As remarked earlier the crux of the technique lies in the relations (3) and (4) between non-negative sequences of numbers $\{w_k\}$ and $\{y_k\}$. In this section, we deal with the extension of Bonferroni type bounds for any sequence $\{w_k\}$ related to a sequence $\{y_k\}$ through the relation (3) in which binomial coefficients have been replaced by a general set of numbers. We shall obtain a set of sufficient conditions on the numbers to provide Bonferroni type of bounds for any w_k. First, a lemma on a generalized inverse relation is given.

LEMMA. *Let a_i, $i=0,1,\dots,n$, $a_0 \neq 0$ be a given set of numbers. The following inversion relation hold for any two sequences $\{w_k\}_{0\leq k\leq n}$ and $\{y_k\}_{0\leq k\leq n}$:*

$$y_k = \sum_{s=k}^{n} b_{s-k} w_s, \tag{30}$$

if and only if

$$w_t = \sum_{s=t}^{n} (-1)^{s-t} a_{s-t} y_s, \tag{31}$$

where $b_0 = 1/a_0$, and for $r \geqq 1$

$$b_r = \frac{1}{a_0^{r+1}} \begin{vmatrix} a_1 & a_2 & \dots & a_{r-1} & a_r \\ a_0 & a_1 & \dots & a_{r-2} & a_{r-1} \\ 0 & a_0 & \dots & a_{r-3} & a_{r-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_0 & a_1 \end{vmatrix}. \tag{32}$$

PROOF. Substitute for y_s from (30) in (31) and rearrange. Then we have

$$w_t = \sum_{u=t}^{n} w_u \sum_{s=t}^{n} (-1)^{s-t} a_{s-t} b_{u-s}.$$

It can be checked that the second summation is 1 if $u=t$ and is zero if $u>t$ since it is equal to

$$\frac{1}{a_0^{u-t+1}} \begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_{u-t-1} & a_{u-t} \\ a_0 & a_1 & a_2 & \dots & a_{u-t-1} & a_{u-t} \\ 0 & a_0 & a_1 & \dots & a_{u-t-2} & a_{u-t-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & a_0 & a_1 \end{vmatrix}$$

when expanded by its first row. Similarly,

$$\begin{aligned} y_k &= \sum_{u=k}^{n} y_u \sum_{s=k}^{u} (-1)^{u-s} a_{u-s} b_{s-k} \\ &= \sum_{u=k}^{n} y_u \sum_{v=k}^{u} (-1)^{v-k} a_{v-k} b_{u-v} \end{aligned}$$

by substituting $v=u-s+k$. The r.h.s. is of the same form as the r.h.s. of the previous expression. This proves the lemma. □

A sequence of numbers $\{a_n\}_{n\geqq 0}$ is called a Pólya frequency sequence if all minors of infinite matrix $(a_{j-i})_{i,j\geqq 0}$ have non-negative determinants (see Brenti [1], p. 9). Next we give a sufficient set of conditions for Bonferroni type bounds to hold for w_t.

THEOREM 4. *Let $n \geqq 0$ be a fixed integer. Let $\{w_k\}_{0 \leq k \leq n}$ be a sequence of non-negative numbers. Let another sequence $\{y_k\}_{0 \leq k \leq n}$ be defined by relation* (30). *If $\{a_k\}_{k \geq 0}$ is a Pólya frequency sequence with $a_0 > 0$, for $t = 0, 1, \ldots, n$ and $u \geqq 0$, then the following inequalities hold:*

$$\sum_{j=t}^{t+2u+1} (-1)^{j-t} a_{j-t} y_j \leqq w_t \leqq \sum_{j=t}^{t+2u} (-1)^{j-t} a_{j-t} y_j. \tag{33}$$

PROOF. We rewrite (31) as

$$w_t = \sum_{j=t}^{t+p-1} (-1)^{j-t} a_{j-t} y_j + (-1)^p \sum_{j=t+p}^{n} (-1)^{j-t-p} a_{j-t} y_j.$$

Denote by $(-1)^p \delta(t, p-1, n)$ the second term on the r.h.s. of the last expression. Then by substituting for y_j from (30), we have

$$\begin{aligned} \delta(t, p-1, n) &= \sum_{j=t+b}^{n} (-1)^{j-t-p} a_{j-t} \sum_{l=j}^{n} b_{l-j} w_l \\ &= \sum_{y=p}^{n-t} w_{y+t} \sum_{x=p}^{y} (-1)^{x-p} a_x b_{y-x} \\ &= \sum_{y=p}^{n-t} w_{y+t} \frac{M_{y-b+1,p-1}}{a_0^{y-p+1}}, \end{aligned} \tag{34}$$

where

$$M_{r,s} = \begin{vmatrix} a_{s+1} & a_{s+2} & \ldots & a_{s+r-1} & a_{s+r} \\ a_0 & a_1 & \ldots & a_{r-2} & a_{r-1} \\ 0 & a_0 & \ldots & a_{r-3} & a_{r-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & a_0 & a_1 \end{vmatrix}. \tag{35}$$

Since $\{a_k\}_{k \geq 0}$ is a Pólya frequency and $M_{y-p+1,p-1}$ are minors of the matrix $(a_{j-i})_{i,j \geq 0}$, we have $M_{y-p+1,p-1} \geqq 0$ and consequently $\delta(t, p-1, n) \geqq 0$. Hence the inequalities follow. □

Interestingly, the inequalities can be sharpened by finding a lower bound for remainder $\delta(t, p-1, n)$. For this purpose we require a preliminary result, which is as follows:

If $\{a_k\}_{k\geqq 0}$ is a Pólya frequency sequence, then $\left\{\dfrac{M_{p,t}}{M_{p,0}}\right\}_{p\geqq 1}$ is a non-decreasing sequence. To prove this, we observe that

$$\begin{aligned} & M_{p,0}M_{p+1,t} - M_{p+1,0}M_{p,t} \\ (36) \qquad & = a_0^p \begin{vmatrix} a_{t+1} & a_{t-2} & \dots & a_{t+p-1} & a_{t+p} & a_{t+p+1} \\ a_1 & a_2 & \dots & a_{p-1} & a_p & a_{p+1} \\ a_0 & a_1 & \dots & a_{p-2} & a_{p-1} & a_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_0 & a_1 & a_2 \end{vmatrix} \end{aligned}$$

which can be established by comparing the coefficient of a_{t+i} $(i=1,2,\ldots,p+1)$ on both sides, by actual expansion of determinants involved. Now since $\{a_i\}_{i\geqq 0}$ is a Pólya frequency sequence, it implies that the determinant on r.h.s. of (36) is $\geqq 0$. Hence $M_{p+1,t}/M_{p+1,0} \geqq M_{p,t}/M_{p,0}$ for $p\geqq 1$, implying thereby that $\left\{\dfrac{M_{p,t}}{M_{p,0}}\right\}_{p\geqq 1}$ is non-decreasing.

Next we give the improved form of Bonferroni bounds.

THEOREM 5. *If the non-negative sequences of numbers $\{w_t\}_{t\geqq 0}$ and $\{y_t\}_{t\geqq 0}$ satisfy* (30), *where $\{a_k\}_{k\geqq 0}$, $a_0>0$ is a Pólya frequency sequence, then the following inequalities hold:*

$$\begin{aligned} (37) \qquad & \sum_{j=t}^{t+2u+1} (-1)^{j-t} a_{j-t} y_j + \left(a_{2u+2} - a_0 \frac{M_{n-t-2u-2,2u+2}}{M_{n-t-2u-2,0}}\right) y_{t+2u+2} \\ & \leqq w_t \leqq \sum_{j=t}^{t+2u} (-1)^{j-t} a_{j-t} y_j - \left(a_{2u+1} - a_0 \frac{M_{n-t-2u-1,2u+1}}{M_{n-t-2u-1,0}}\right) y_{t+2u+1} \end{aligned}$$

where $t=0,1,\ldots,n$, $u\geqq 0$.

PROOF. Expanding the determinant $M_{y-p+1,p-1}$ by its first column we get $M_{y-p-1,p-1} = a_p M_{y-p,0} - a_0 M_{y-p,p}$. Thus

$$\frac{M_{y-p+1,p-1}}{M_{y-p,0}} = a_p - a_0 \frac{M_{y-p,p}}{M_{y-p,0}} \geqq a_p - a_0 \frac{M_{n-t-p,p}}{M_{n-t-p,0}}$$

as $\left\{\dfrac{M_{p,t}}{M_{p,0}}\right\}_{p\geqq 0}$ is nondecreasing. Using this bound in $\delta(t,p-1,n)$ given by (34) we get

$$\begin{aligned} \delta(t,p-1,n) &\geqq \left(a_p - a_0 \frac{M_{n-t-p,p}}{M_{n-t-p,0}}\right) \sum_{y=p}^{n-t} \frac{M_{y-p,p}}{a_0^{y-p+1}} w_{y+t} \\ &= \left(a_p - a_0 \frac{M_{n-t-p,p}}{M_{n-t-p,0}}\right) y_{p+t} \quad \text{(from (30))}. \end{aligned}$$

Using this lower bound for remainder term, we get the desired inequalities (37). □

In conclusion, it is remarked that the Bonferroni type inequalities can be extended beyond its original probabilistic framework.

REFERENCES

[1] Brenti, F., Unimodal, log-concave and Pólya frequency sequences in combinatorics, *Mem. Amer. Math. Soc.* **81** (1989), no. 413, 106 pp. *MR* **90d**:05014

[2] Chen, T. and Seneta, E., Multivariate Bonferroni-type bounds, *J. Appl. Probab.* **33** (1996), 729–740.

[3] Fréchet, M., *Les probabilités associées à un système d'événements compatibles et dépendants*, II, Cas particuliers et applications, Actualités sci. ind., no. 942, Exposés d'analyse général, Hermann, Paris, 1943, 81–211. *MR* **7**, 456g

[4] Galambos, J., Bonferroni inequalities, *Ann. Probability* **5** (1977), 577–581. *MR* **56** #6784

[5] Galambos, J. and Lee, M. Y., Extensions of some univariate Bonferroni-type inequalities to multivariate setting, *Probability theory and applications*, Math. Appl., 80, Kluwer, Dordrecht, 1992, 143–154. *MR* **94c**: 60025

[6] Galambos, J. and Lee, M. Y., Further studies of bivariate Bonferroni-type inequalities, *J. Appl. Probab.* **31A** (1994), 63–69. *MR* **95h**:60028

[7] Galambos, J. and Simonelli, I., *Bonferroni-type inequalities with applications*, Springer-Verlag, New York, N. Y., 1996.

[8] Galambos, J. and Xu, Y., Univariate and multivariate Bonferroni-type inequalities: methods for proof and questions of optimality, *Probability theory and applications*, Math. Appl. 80, Kluwer, Dordrecht, 1992, 155–164. See *MR* **93i**:60005

[9] Galambos, J. and Xu, Y., Some optimal bivariate Bonferroni-type bounds, *Proc. Amer. Math. Soc.* **117** (1993), 523–528. *MR* **93d**:60026

[10] Galambos, J. and Xu, Y., Bivariate extension of the method of polynomials for Bonferroni-type inequalities, *J. Multivariate Anal.* **52** (1995), 131–139. *Zbl* **817**.60011

[11] Gumbel, E. J., Gli eventi compatibili, *Giorn. Ist. Ital. Attuari* **9** (1938), 58–93. *Zbl* **19**, 128

[12] Hoppe, F. M., Beyond inclusion-exclusion: Natural identities, *Internat. Statist. Rev.* **61** (1993), 435–446.

[13] Hoppe, F. M., and Seneta, E., A Bonferroni identity and permutation bounds, *Internat. Statist. Rev.* **58** (1990), 253–261. See *MR* **94h**:60016

[14] Lee, M. Y., Bivariate Bonferroni inequalities, *Aequationes Math.* **44** (1992), 220–225. *MR* **93h**:60028

[15] Meyer, R. M., Note on a multivariate form of Bonferroni's inequalities, *Ann. Math. Statist.* **40** (1969), 692–693. *MR* **38** #6643

[16] Recsei, E. and Seneta, E., Bonferroni-type inequalities, *Adv. in Appl. Probab.* **19** (1987), 508–511. *MR* **88g**:60029

[17] Takács, L., On a general probability theorem and its applications in the theory of the stochastic processes, *Proc. Cambridge Philos. Soc.* **54** (1958), 219–224. *Zbl* **80**, 352

(Received May 22, 1996)

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY
HAUZ KHAS
IND–110016 NEW DELHI
INDIA

brhanda@netearth.ernet.in

DEPARTMENT OF MATHEMATICS AND
STATISTICS
McMASTER UNIVERSITY
HAMILTON, ONTARIO
L8S 4K1
CANADA

mohanty@mcmaster.ca

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 167–175

ALMOST SURE LIMIT THEOREMS FOR DEPENDENT RANDOM VARIABLES

G. HURELBAATAR

Dedicated to Endre Csáki on the occasion of his sixtieth birthday

Abstract

For partial sums S_k of strongly mixing and associated random variables we prove that

$$(1/\log n)\sum_{k\leq n}(1/k)\mathbf{I}\{S_k/a_k\in\cdot\}\to G(\cdot)$$

with probability 1 if and only if

$$(1/\log n)\sum_{k\leq n}(1/k)\mathbf{P}(S_k/a_k\in\cdot)\to G(\cdot)$$

under the same moment condition as assumed for independent random variables.

1. Introduction

One of the extensions of classical probability limit theorems is the so-called almost sure limit theorem. The basic result and starting point of these investigations is the almost sure central limit theorem, discovered by Brosamler [2] and Schatte [8] for i.i.d. random variables having finite $(2+\delta)$th moment and later proved by Fisher [3] and Lacey and Philipp [4] to hold under assuming only finite variance:

THEOREM. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and set $S_n = X_1 + \cdots + X_n$. Then*

$$\lim_{n\to\infty}\frac{1}{\log n}\sum_{k\leq n}\frac{1}{k}\mathbf{I}\left\{\frac{S_k}{\sqrt{k}}\in A\right\} = (2\pi)^{-1/2}\int_A e^{-t^2/2}dt \quad a.s.$$

for any Borel-set $A\subset\mathbb{R}$ with $\lambda(\partial A)=0$; moreover, the exceptional set of probability zero can be chosen to be independent of A. Here $\mathbf{I}$ denotes indicator function and λ denotes the Lebesgue measure.

Later Berkes and Dehling [1] proved a more general version of the almost sure central limit theorem and its functional version for independent, not necessarily identically distributed random variables.

1991 *Mathematics Subject Classification.* Primary 60F15, 60F05.

Key words and phrases. Strongly mixing, associated random variables, almost sure central limit theorem.

THEOREM A (Theorem 1 of Berkes and Dehling [1]). *Let $X_1, X_2, \dots$ be independent random variables and let $a_n > 0$, b_n be numerical sequences such that*

$$\mathbf{E}f\left(\left|\frac{S_n - b_n}{a_n}\right|\right) \leqq (\log\log n)^{-1-\varepsilon} f(e^{(\log n)^{1-\varepsilon}}) \qquad (n \geqq n_0) \tag{1.1}$$

for some $\varepsilon > 0$, where $f(x) \geqq 0$ is a real function such that $f(x)$ and $x/f(x)$ are nondecreasing and the right-hand side of (1.1) *is nondecreasing for $n \geqq n_0$. Assume that*

$$\frac{a_l}{a_k} \geqq C\left(\frac{l}{k}\right)^{\gamma} \qquad (l \geqq k \geqq n_0) \tag{1.2}$$

for some constants $C > 0$ and $\gamma > 0$. Then for any distribution function G the following statements are equivalent:

(a) *For any Borel set $A \in \mathbb{R}$ with $G(\partial A) = 0$ we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbf{I}\left\{\frac{S_k - b_k}{a_k} \in A\right\} = G(A) \qquad a.s. \tag{1.3}$$

where the exceptional set of probability zero is independent of A.

(b) *For any Borel set $A \in \mathbb{R}$ with $G(\partial A) = 0$ we have*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbf{P}\left(\frac{S_k - b_k}{a_k} \in A\right) = G(A). \tag{1.4}$$

In this paper we prove Theorem A for some dependent random variables under the assumptions (1.1) and (1.2). An interesting result of Berkes and Dehling [1] is that weak and strong laws of large numbers are equivalent on a set of log-density 1. We also prove this fact for strongly mixing and associated random variables.

2. Almost sure limit theorems

DEFINITION 2.1. Let $X_1, X_2, \dots$ be a sequence of random variables on some probability space $(\Omega, \mathcal{F}, P)$ and let σ_a^b be the σ-algebra generated by the random variables $X_a, X_{a+1}, \dots, X_b$. For any two σ-algebras $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$ define

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbf{P}(AB) - \mathbf{P}(A)\mathbf{P}(B)|\,;\ A \in \mathcal{A},\ B \in \mathcal{B}\}$$

and put

$$\alpha(n) = \sup_{k\geq 1} \alpha(\sigma_1^k, \sigma_{k+n}^\infty).$$

The sequence $X_1, X_2, \dots$ is called strongly mixing if $\alpha(n) \to \infty$ as $n \to \infty$.

DEFINITION 2.2. The sequence of random variables $X_1, X_2, \dots, X_n$ is called associated if

$$\mathbf{Cov}(f(X_1, X_2, \dots, X_n), g(X_1, X_2, \dots, X_n)) \geqq 0$$

for every $n \geqq 2$, whenever f, g: $\mathbb{R}^n \to \mathbb{R}^1$ are coordinatewise increasing. The following results will be used in our proofs.

THEOREM 2.1 (Theorem 1.1 in Rio, E. [7]). *Let X and Y be two integrable real-valued random variables. Let $\alpha = \alpha(\sigma(X), \sigma(Y)) \leqq 1/4$. Let $Q_X(u) = \inf\{t : \mathbf{P}(|X| > t) \leqq u\}$ denote the quantile function of $|X|$. Assume that $Q_X Q_Y$ is integrable on $[0,1]$. Then*

$$|\mathbf{Cov}(X, Y)| \leqq 2 \int_0^{2\alpha} Q_X(u) Q_Y(u) du.$$

This theorem implies immediately that if $|X| \leqq K$ and $|Y| \leqq K$ then

$$|\mathbf{Cov}(X, Y)| \leqq 4K^2 \alpha(\sigma(X), \sigma(Y)). \tag{2.1}$$

HÖFFDING'S EQUALITY. *If the covariance of X and Y exists, then*

$$\mathbf{Cov}(X, Y) = \int\int (\mathbf{P}(X > x, Y > y) - \mathbf{P}(X > x)\mathbf{P}(Y > y)) dx dy.$$

THEOREM 2.2. *Let $X_1, X_2, \dots$ be strongly mixing random variables and let $a_n > 0$, b_n be numerical sequences satisfying* (1.1) *and* (1.2). *Assume that*

$$\alpha(k) = O((\log \log k)^{-1-\delta}) \tag{2.2}$$

for some $\delta > 0$. Then the statements (a) *and* (b) *are equivalent.*

PROOF. Without loss of generality we can assume $b_n = 0$. Let $g(x)$ be a bounded Lipschitz function, e.g. $|g(x)| \leqq K$ and $|g(x) - g(y)| \leqq K|x - y|$ for some $K > 0$ and for all x, y. It suffices to show that

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{i=1}^{n} \frac{1}{i} \xi_i = 0 \quad \text{a.s.,} \tag{2.3}$$

where $\xi_k = g(\frac{S_k}{a_k}) - \mathbf{E} g(\frac{S_k}{a_k})$ (cf. [1]). Assume that $l > 2k$. Then we have

$$\begin{aligned} |\mathbf{E}(\xi_k \xi_l)| &= \left|\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l}{a_l}\right)\right)\right| \leqq \\ &\leqq \left|\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l}{a_l}\right) - g\left(\frac{S_l - S_{2k}}{a_l}\right)\right)\right| + \\ &\quad + \left|\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l - S_{2k}}{a_l}\right)\right)\right|. \end{aligned} \tag{2.4}$$

By (2.1) it follows that

$$\left|\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l - S_{2k}}{a_l}\right)\right)\right| \leqq 4K^2\alpha(k), \tag{2.5}$$

since g is bounded.

Setting $\lambda = a_l/a_{2k}$ and using $x/y \leqq f(x)/f(y)$ for $x \leqq y$ we get

$$\begin{aligned}\left|\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l}{a_l}\right) - g\left(\frac{S_l - S_{2k}}{a_l}\right)\right)\right| &\leqq C\mathbf{E}\left(\frac{|S_{2k}|}{a_l} \wedge 1\right) = \\ = \mathbf{E}\left(\frac{C}{\lambda}\left(\frac{|S_{2k}|}{a_{2k}} \wedge \lambda\right)\right) &\leqq \frac{C}{f(\lambda)}\mathbf{E}f\left(\frac{|S_{2k}|}{a_{2k}}\right).\end{aligned} \tag{2.6}$$

Now

$$\begin{aligned}&\mathbf{E}\left(\sum_{k=1}^{n}\frac{1}{k}\xi_k\right)^2 \leqq \\ &\leqq \sum_{k=1}^{n}\frac{1}{k^2}\mathbf{E}|\xi_k|^2 + 2\sum_{\substack{1\leqq k<l\leqq n\\ 2k\geqq l}}\frac{|\mathbf{E}(\xi_k\xi_l)|}{kl} + 2\sum_{\substack{1\leqq k<l\leqq n\\ 2k<l}}\frac{|\mathbf{E}(\xi_k\xi_l)|}{kl} = \\ &= \sum\nolimits_{1} + \sum\nolimits_{2} + \sum\nolimits_{3}.\end{aligned} \tag{2.7}$$

By trivial estimation $|\mathbf{E}(\xi_k\xi_l)| \leqq 4K^2$ we get that

$$\sum\nolimits_{1} \leqq 4K^2\sum_{k=1}^{\infty}\frac{1}{k^2} < \infty \tag{2.8}$$

$$\sum\nolimits_{2} \leqq 4K^2\sum_{k=1}^{n}\sum_{l=k+1}^{2k}\frac{1}{kl} = O(\log n). \tag{2.9}$$

By (2.4)–(2.6) we have

$$\begin{aligned}\sum\nolimits_{3} &\leqq 4K^2\sum_{\substack{1\leqq k<l\leqq n\\ 2k<l}}\frac{\alpha(k)}{kl} + C\sum_{\substack{1\leqq k<l\leqq n\\ 2k<l}}\frac{1}{kl}\frac{1}{f(\lambda)}\mathbf{E}f\left(\frac{|S_{2k}|}{a_{2k}}\right) = \\ &= \sum\nolimits_{31} + \sum\nolimits_{32}.\end{aligned} \tag{2.10}$$

(2.2) implies that

$$\begin{aligned}\sum\nolimits_{31} &\leqq C_1\sum_{l=2}^{n}\frac{1}{l}\sum_{k=1}^{l-1}\frac{\alpha(k)}{k} \leqq \\ &\leqq C_2\sum_{l=1}^{n}\frac{\log l}{l(\log\log l)^{1+\delta}} \leqq C(\log\log n)^{-1-\delta}\log^2 n.\end{aligned} \tag{2.11}$$

If $l/2k \geqq \exp(\log^{1-\varepsilon/2} n)$ then by (1.1) we have

$$\sum\nolimits_{32} \leqq C_1(\log\log n)^{-1-\varepsilon} \frac{f(e^{\log^{1-\varepsilon} n})}{f(Ce^{\gamma \log^{1-\varepsilon/2} n})} \sum_{1\leqq k<l\leqq n} \frac{1}{kl} \leqq \tag{2.12}$$
$$\leqq C(\log\log n)^{-1-\varepsilon} \log^2 n.$$

If $l/2k \leqq \exp(\log^{1-\varepsilon/2} n)$ then using $|\mathbf{E}(\xi_k\xi_l)| \leqq 4K^2$ we get that

$$\sum\nolimits_{32} \leqq C_3 \sum_{k\leqq l\leqq k\exp(\log^{1-\varepsilon/2} n)} \frac{1}{kl} \leqq \tag{2.13}$$
$$\leqq C_4 \sum_{k=1}^{n} \frac{1}{k} \log^{1-\frac{\varepsilon}{2}} n \leqq C \log^{2-\frac{\varepsilon}{2}} n.$$

Setting $T_n = \log^{-1} n \sum_{k\leqq n} k^{-1}\xi_k$, we have

$$\mathbf{E}T_n^2 \leqq C_1(\log\log n)^{-1-\min(\varepsilon,\delta)}.$$

Choosing $n_k = \exp(\exp(k^{1-\varepsilon/2}))$ we have

$$\sum_{j=1}^{\infty} \mathbf{E}T_{n_k}^2 < \infty.$$

Thus $T_{n_k} \to 0$ a.s.. Now for $n_k \leqq n \leqq n_{k+1}$ we have

$$|T_n - T_{n_k}| \leqq \left(1 - \frac{\log n_k}{\log n}\right).$$

Since $\log n_k / \log n \to 1$, it follows $T_n \to 0$ a.s.. This completes the proof of Theorem 2.2.

THEOREM 2.3. *Let $X_1, X_2, \dots$ be associated random variables and let $a_n > 0$, b_n be numerical sequences satisfying* (1.1) *and* (1.2) *with $\gamma \geqq 1/2$. Assume that*

$$u(n) = \sup_{k\geqq 1} \sum_{j:|k-j|\geqq n} \mathbf{Cov}(X_k, X_j) < C \tag{2.14}$$

for all $n \geqq 1$. Then the statements (a) *and* (b) *are equivalent.*

PROOF. The proof is similar to that of Theorem 2.2. The difference is only the estimation of $\mathbf{Cov}\,(g(S_k/a_k), g((S_l - S_{2k})/a_l))$ and consequently to show that

$$\sum\nolimits_{31} = O(\log n).$$

From the definition it follows that S_k and $S_l - S_{2k}$ are also associated. Therefore

$$H(x,y) = \mathbf{P}(S_k > xa_k, S_l - S_{2k} > ya_l) - \mathbf{P}(S_k > xa_k)\mathbf{P}(S_l - S_{2k} > ya_l)$$

is nonnegative for every $x, y \in \mathbb{R}$.

Using Höffding's equality and the absolute continuity of g we have

$$\mathbf{Cov}\left(g\left(\frac{S_k}{a_k}\right), g\left(\frac{S_l - S_{2k}}{a_l}\right)\right) = \int\int g'(x)g'(y)H(x,y)dxdy \leqq$$
$$\leqq K^2 \mathbf{Cov}\left(\frac{S_k}{a_k}, \frac{S_l - S_{2k}}{a_l}\right) = O\left(\frac{k}{a_k a_l}\right),$$

since $\mathbf{Cov}(S_k, S_l - S_{2k}) \leqq C_1 k u(k)$.

Now we put $O(k/a_k a_l)$ in (2.11) instead of $\alpha(k)$ and we get that

$$\sum\nolimits_{31} \leqq C_2 \sum_{l=2}^{n} \frac{1}{l^{3/2}} \sum_{k=1}^{l-1} \frac{1}{k^{1/2}} = O(\log n)$$

as desired.

3. Laws of Large Numbers

Definition 3.1. For a set $A \subset \mathbb{N}$ of positive integers, the log density $\mu_L(A)$ of A is defined by

$$\mu_L(A) = \lim_{n \to \infty} \frac{1}{\log n} \sum_{k \leqq n, k \in A} \frac{1}{k}$$

provided that the limit exists.

Definition 3.2. Let $\xi_1, \xi_2, \ldots$ and ξ be r.v.'s. We say that

$$\xi_n \xrightarrow{P} \xi \quad (\log)$$

if there exists a set $H \subset \mathbb{N}$ of log density 1 such that $\xi_n \xrightarrow{P} \xi$ as $n \to \infty$, $n \in H$. We say that

$$\xi_n \to \xi \quad \text{a.s.} \quad (\log)$$

if for a.e. ω there exists a set $H_\omega \subset \mathbb{N}$ of log density 1 such that $\xi_n(\omega) \to \xi(\omega)$ as $n \to \infty$, $n \in H_\omega$.

Theorem B (Theorem 3 of [1]). *Let $X_1, X_2, \ldots$ be independent random variables and let $a_n > 0$ be a numerical sequence satisfying* (1.1) *and* (1.2) *with $b_n = 0$. Then the following statements are equivalent:*

$$S_n / a_n \longrightarrow 0 \quad a.s.\ (\log) \tag{3.1}$$

and

$$S_n/a_n \xrightarrow{P} 0 \quad (\log). \tag{3.2}$$

Moreover, if $X_n/a_n \xrightarrow{P} 0$ *also holds then a third equivalent condition is* (3.3)-(3.4) *as follows:*

$$\frac{1}{\log n}\sum_{k\le n}\frac{1}{k}\{G_k(a_k)\wedge 1\} \longrightarrow 0 \quad as \quad n\to\infty \tag{3.3}$$

and

$$\frac{1}{\log n}\sum_{k\le n}\frac{1}{k}\{H_k(a_k)\wedge 1\} \longrightarrow 0 \quad as \quad n\to\infty, \tag{3.4}$$

where

$$G_k(\lambda)=\sum_{j=1}^{k}\Bigg[\int\limits_{|x|\ge\lambda} dF_j(x)+\lambda^{-2}\Bigg\{\int\limits_{|x|<\lambda} x^2 dF_j(x)-\Bigg(\int\limits_{|x|<\lambda} x\,dF_j(x)\Bigg)^2\Bigg\}\Bigg],$$

$$H_k(\lambda)=\frac{1}{\lambda}\Bigg|\sum_{j=1}^{k}\int\limits_{|x|<\lambda} x\,dF_j(x)\Bigg|$$

and F_j *is the distribution function of* X_j.

Theorem B was first proved by Berkes and Dehling [1] under the assumptions of Theorem A. Now we show the equivalence of (3.1) and (3.2) under the assumptions of Theorems 2.2 and 2.3.

THEOREM 3.1. *Let* $X_1, X_2, \ldots$ *be strongly mixing random variables with* $\mathbf{E}X_i=0$ *and let* $a_n>0$ *be a numerical sequence satisfying* (1.1) *and* (1.2) *with* $b_n=0$. *Assume that*

$$\alpha(k)=O((\log\log k)^{-1-\delta}).$$

Then the statements (3.1) *and* (3.2) *are equivalent.*

THEOREM 3.2. *Let* $X_1, X_2, \ldots$ *be associated random variables with* $\mathbf{E}X_i=0$ *and let* $a_n>0$ *be a numerical sequence such that* (1.1) *and* (1.2) *hold with* $b_n=0$. *Assume that*

$$u(n)=\sup_{k\ge1}\sum_{j:|k-j|\ge n}\mathbf{Cov}(X_k,X_j)<C.$$

Then the statements (3.1) *and* (3.2) *are equivalent.*

In the proofs of Theorem 3.1 and 3.2 we use the following lemma.

LEMMA 3.1 (Lemma 2 of [1]). *Let $x_1, x_2, \dots$ be a numerical sequence. Then the following statements are equivalent:*

(i) *There exists a subset $H \subset \mathbb{N}$ of* log *density* 0 *such that $x_n \to 0$ as $n \to \infty$, $n \notin H$.*

(ii) *For all $\varepsilon > 0$, the set $A(\varepsilon) = \{n; |x_n| > \varepsilon\}$ has* log *density* 0, *that is,*

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k \leqq n} \frac{1}{k} \mathbf{I}\{|x_k| > \varepsilon\} = 0.$$

Moreover, if x_n is bounded then (i) *and* (ii) *are equivalent to*

$$\text{(iii)} \qquad \lim_{n\to\infty} \frac{1}{\log n} \sum_{k \leqq n} \frac{1}{k} |x_k| = 0.$$

PROOF OF THEOREM 3.1. Clearly, in the case when $b_n = 0$ and G is the distribution concentrated at the origin, statements (a) and (b) in Theorem A reduce to

$$\text{(3.5)} \qquad \lim_{n\to\infty} \frac{1}{\log n} \sum_{k \leqq n} \frac{1}{k} \mathbf{I}\{|S_k/a_k| > \varepsilon\} = 0 \quad \text{a.s. for any} \quad \varepsilon > 0$$

and

$$\text{(3.6)} \qquad \lim_{n\to\infty} \frac{1}{\log N} \sum_{k \leqq N} \frac{1}{k} \mathbf{P}\{|S_k/a_k| > \varepsilon\} = 0 \quad \text{for any} \quad \varepsilon > 0,$$

respectively. By Lemma 3.1, (3.5) is equivalent to (3.1) while (3.6) can be written equivalently as

$$\text{(3.7)} \qquad \mu_L\{n: \mathbf{P}(|S_n/a_n| > \varepsilon) > \delta\} = 0 \quad \text{for any} \quad \varepsilon > 0,\ \delta > 0.$$

Setting

$$x_n = \inf\{\varrho > 0: \mathbf{P}(|S_n/a_n| > \varrho) \leqq \varrho\}$$

(3.7) implies

$$\mu_L\{n: x_n > \varepsilon\} \leqq \mu_L\{n: \mathbf{P}(|S_n/a_n| > \varepsilon) > \varepsilon\} = 0 \quad \text{for all} \quad \varepsilon > 0,$$

whence we get, using Lemma 3.1, that $x_n \to 0$ along a sequence $H \subset \mathbb{N}$ of log density 1. We can easily see that the following two statements are equivalent:

(i) $\mathbf{P}(|S_n/a_n| > \varepsilon) \leqq \varepsilon$ for any $\varepsilon > 0$.

(ii) $\mathbf{P}(|S_n/a_n| > \varepsilon) \to 0$ as $n \to \infty$.

Thus we get $S_n/a_n \xrightarrow{P} 0$ as $n \to \infty, n \in H$ that is, (3.2) holds. Conversely, (3.2) trivially implies (3.7) and thus (3.6).

PROOF OF THEOREM 3.3. Our proof immediately follows from Theorem 2.3 applying the same procedure as in the proof of Theorem 3.1.

ACKNOWLEDGEMENTS. I am much indebted to E. Csáki and I. Berkes for their fruitful conversations leading to more general formulation of the Theorems. Research supported by Hungarian National Foundation for Scientific Research Grant No. T 016384.

REFERENCES

[1] BERKES, I. and DEHLING, H., Some limit theorems in log density, *Ann. Probab.* **21** (1993), 1640–1670. *MR* **94h**:60026

[2] BROSAMLER, G. A., An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.* **104** (1988), 561–574. *MR* **89i**:60045

[3] FISHER, A., A pathwise central limit theorem for random walks, 1989 (preprint). (See *MR* **95g**:60041)

[4] LACEY, M. T. and PHILIPP, W., A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9** (1990), 201–205. *MR* **91e**:60100

[5] LEHMANN, E. L., Some concepts of dependence, *Ann. Math. Statist.* **37** (1966), 1137–1153. *MR* **34** #2101

[6] LOÈVE, M., *Probability theory*. Foundations. Random sequences, D. Van Nostrand Co., New York–Toronto–London, 1955. *MR* **16**–598f

[7] RIO, E., Covariance inequalities for strongly mixing processes, *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1993), 587–597. *MR* **94j**:60045

[8] SCHATTE, P., On strong versions of the central limit theorem, *Math. Nachr.* **137** (1988), 249–256. *MR* **89i**:60070

(Received April 4, 1995)

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS AT URBANA–CHAMPAIGN
URBANA, IL 61820
U.S.A.

hurlee@stat.uiuc.edu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 177–183

ESCAPE RATES FOR LÉVY PROCESSES*

D. KHOSHNEVISAN

Dedicated to Professor E. Csáki on the occasion of his sixtieth birthday

Abstract

We prove a space-time estimate for a Lévy process to hit a small set. As an application, we present escape rates for Lévy processes with strictly stable components.

§ 1. Introduction

Let X denote a d-dimensional Lévy process. It is a classical fact that a Borel set $A \subset \mathbb{R}^d$ is polar for X if and only if A has positive X-capacity; cf. Blumenthal and Getoor [BG]. A sharper variant of the aforementioned fact is the consequence of more recent investigations such as those of Benjamini et al. [BPP], Fitzsimmons and Salisbury [FS], Peres [Pe] and Salisbury [Sa]. Roughly speaking, these results provide in a variety of different contexts, quantitative estimates of the type: $\mathbf{P}^0(X_t \in A,\ \text{for some } t > 0) \asymp e^{-1}(A)$, where $f \asymp g$ implies the existence of some universal $C > 1$, such that $C^{-1}g \leqq f \leqq Cf$ pointwise, and $e(A)$ is the X-energy integral associated with A. One of the many uses of such an estimate is that one can often approximate the chance that X ever hits a small set. Wishing to study escape rates, we present a different sort of a quantitative estimate below. Our notation is more or less that of Markov process theory.

THEOREM 1.1. *Suppose X is a d-dimensional Lévy process. For any $b > a > 0$ and $\varepsilon > 0$,*

$$\frac{1}{2}\frac{\int_a^b \mathbf{P}^0(|X_r| \leqq \varepsilon)dr}{\int_0^b \mathbf{P}^0(|X_r| \leqq \varepsilon)} \leqq \mathbf{P}^0(|X_r| \leqq \varepsilon,\ \text{for some}\ \ a \leqq t \leqq b) \leqq$$

1991 *Mathematics Subject Classification.* Primary 60J30; Secondary 60J45, 60J75.
Key words and phrases. Lévy processes, hitting estimates.

*Research partially supported by NSF Grant DMS–95–03290.

$$\leq \frac{\int_a^{2b-a} \mathbf{P}^0(|X_r| \leq 2\varepsilon)dr}{\int_0^{b-a} \mathbf{P}^0(|X_r| \leq \varepsilon)dr},$$

whenever the integrals exist and are nonzero.

The above extends the estimates of Perkins and Taylor [PT], Takeuchi [T1, T2] and Takeuchi and Watanabe [TW], to cite a few examples. To illustrate the use of such a general inequality, let us restrict our attention to the class of processes described in Hendricks [H1, H2, H3]. Namely, we consider the case where X is a d-dimensional Lévy process with strictly stable components. In other words, there exists $\upsilon, \chi \in \mathbb{R}_+^d$ and $\alpha \in (0,2]^d$, such that for all $t > 0$ and all $\zeta \in \mathbb{R}^d$,

$$\mathbf{P}^0 \exp(i\zeta' X_t) = \exp\left(-t\sum_{j=1}^d |\zeta_j \chi_j|^{1/\alpha_j} - i\sum_{j=1}^d \upsilon_j \operatorname{sgn}(\zeta_j)\right). \tag{1.2}$$

Throughout, we shall assume that the coordinate processes are not completely asymmetric, i.e.,

$$\left|\frac{\upsilon_j}{\chi_j}\right| < \tan(\pi\alpha_j/2), \qquad |\chi_j| > 0, \quad \text{for all } \ j = 1, \ldots, d. \tag{1.3}$$

Viewed coordinate by coordinate, such processes scale, albeit differently in each. Define,

$$\beta = \sum_{j=1}^d \frac{1}{\alpha_j}. \tag{1.4}$$

Our intended application of Theorem 1.1 is the following:

THEOREM 1.5. *Suppose X is a Lévy process with stable components with parameters given by (1.2)–(1.4). When $\beta < 1$, X hits points. When $\beta = 1$, singletons are polar, but X is neighbourhood recurrent. When $\beta > 1$, X is transient. For $\beta \geq 1$, let φ: $\mathbb{R}_+^1 \mapsto \mathbb{R}_+^1$ be a decreasing function and define*

$$\mathfrak{I}(\varphi) = \begin{cases} \int_1^\infty t^{-1}\varphi^{\beta-1}(t)dt, & \text{if } \ \beta > 1 \\ \int_1^\infty t^{-1}|\ln\varphi(t)|^{-1}dt, & \text{if } \ \beta = 1. \end{cases}$$

When $\beta \geq 1$, $\mathbf{P}^0$-almost surely,

$$\liminf_{\substack{t\to\infty \\ (t\to 0^+)}} \frac{\max_{1\leq j\leq d} |X_t^j|^{\alpha_j}}{t\varphi(t)} = \begin{cases} \infty, & \text{if } \ \mathfrak{I}(\varphi) < \infty \\ 0, & \text{if } \ \mathfrak{I}(\varphi) = \infty. \end{cases}$$

When α is a constant vector, the above appears to various degrees of generality in Dvoretzky and Erdős [DE], Spitzer [Sp], Takeuchi [T1,T2] and Takeuchi and Watanabe [TW]. When α is not a constant vector, a different but equivalent formulation can be found in Hendricks [H1] with a longer proof. Our formulation has two distinct advantages over the latter: (1) the large-time results and the small-time results are the same; (2) ours incorporates all the known results as one. Note that the critical case (i.e., $\beta=1$) only applies to two cases: $d=1$ and $\alpha=1$ (Cauchy process on $\mathbb{R}^1$) or $d=2$ and $\alpha_1=\alpha_2=2$ (planar Brownian motion).

Above and throughout, we have used the notation: $\ln x \triangleq \log_e(x\vee 1)$, $x\geqq 0$.

§ 2. Proof of Theorem 1.1

Fix $0<a<b$ and define $T\triangleq \inf(s>0:|X_s|\leqq\varepsilon)$. Apply the strong Markov property at time T to see that

$$\begin{aligned}\mathbf{P}^0\int_a^{2b-a} 1(|X_r|\leqq 2\varepsilon)dr &\geqq \mathbf{P}^0\Big(\int_a^{2b-a} 1(|X_r|\leqq 2\varepsilon)dr \,\big|\, T\leqq b\Big)\mathbf{P}^0(T\leqq b)\\ &\geqq \inf_{|x|\leqq\varepsilon}\mathbf{P}^x\int_0^{b-a} 1(|X_r|\leqq 2\varepsilon)dr\,\mathbf{P}^0(T\leqq b)\\ &\geqq \mathbf{P}^0\int_0^{b-a} 1(|X_r|\leqq \varepsilon)dr\,\mathbf{P}^0(T\leqq b).\end{aligned}$$

Divide to obtain the upper bound. The lower bound follows along similar lines. Consider,

$$\begin{aligned}\mathbf{P}^0\Big(\int_a^b 1(|X_r|\leqq\varepsilon)dr\Big)^2 &= 2\mathbf{P}^0\int_a^b\int_a^r 1(|X_r|\leqq\varepsilon,|X_s|\leqq\varepsilon)dsdr\\ &\leqq 2\mathbf{P}^0\int_a^b\int_a^r 1(|X_s|\leqq\varepsilon)1(|X_r-X_s|\leqq 2\varepsilon)dsdr \qquad (2.1)\\ &= 2\int_a^b\int_s^b \mathbf{P}^0(|X_s|\leqq\varepsilon)\mathbf{P}^0(|X_{r-s}|\leqq 2\varepsilon)drds\end{aligned}$$

$$\leq 2\mathbf{P}^0 \int_a^b 1(|X_r| \leq \varepsilon)dr \int_0^b \mathbf{P}^0(|X_r| \leq 2\varepsilon)dr.$$

By the Cauchy–Schwartz inequality,

$$\mathbf{P}^0 \int_a^b 1(|X_r| \leq \varepsilon)dr = \mathbf{P}^0\Big(\int_a^b 1(|X_r| \leq \varepsilon)dr; T \leq b\Big)$$
$$\leq \sqrt{\mathbf{P}^0\Big(\int_a^b 1(|X_r| \leq \varepsilon)dr\Big)^2} \cdot \sqrt{\mathbf{P}^0(T \leq b)}.$$

Use (2.1) and solve to obtain the desired lower bound. □

REMARK 2.2. An inspection of the proof shows that for any $x \in \mathbb{R}^d$,

$$\mathbf{P}^x(|X_t| \leq \varepsilon, \text{ for some } a \leq t \leq b) \leq \frac{\int_a^{2b-a} \mathbf{P}^x(|X_r| \leq 2\varepsilon)dr}{\int_0^{b-a} \mathbf{P}^0(|X_r| \leq \varepsilon)dr}.$$

§ 3. Proof of Theorem 1.5

Throughout this section, X denotes a Lévy process with strictly stable components given by (1.2)–(1.4). Let us start with a technical lemma.

LEMMA 3.1. *The random variable $|X_1|$ has a bounded $\mathbf{P}^a$-density, uniformly over all $a \in \mathbb{R}^d$. When $a = 0$, this density g is positive on some neighbourhood of 0. Moreover, $\sup_x g(x)/g(0) < \infty$.*

PROOF. By properties of convolutions, it suffices to show that each component of X has the given properties. The lemma follows from the inversion theorem for Fourier transforms. □

For the rest of this section, define

$$S(x) \triangleq \max_{1 \leq j \leq d} |x_j|^{\alpha_j}, \qquad x \in \mathbb{R}^d.$$

LEMMA 3.2. *For all $r, a > 0$,*

$$\mathbf{P}^0(S(X_r) \leq a) \asymp (r^{-1}a \wedge 1)^{\beta}.$$

PROOF. Since the components X^j are independent α_j-stable processes, by Lemma 3.1,

$$\mathbf{P}^0(S(X_r) \leqq \varepsilon) = \prod_{j=1}^{d} \mathbf{P}(|X_1^j| \leqq \varepsilon^{1/\alpha_j} r^{-1/\alpha_j})$$
$$\asymp \left(\frac{\varepsilon}{r}\right)^{\beta} \wedge 1.$$

This proves the lemma. □

Since $\alpha_j \leqq 2$ for all j, it is not hard to show that for all $x, y \in \mathbb{R}^d$, $S(x+y) \leqq 4(S(x)+S(y))$. As such, $S(x)$ behaves much like $|x|$. Going through the proof of Theorem 1.1 and using Lemma 3.2, the following estimate emerges:

COROLLARY 3.3. *For all* $0 < a < b$ *and all* $\varepsilon > 0$ *small,*

$$\mathbf{P}^0(S(X_t) \leqq \varepsilon, \text{ for some } a \leqq t \leqq b) \asymp h(\varepsilon),$$

where

$$h(\varepsilon) = \begin{cases} 1 & \text{if } \beta < 1, \\ (\ln(1/\varepsilon))^{-1} & \text{if } \beta = 1, \\ \varepsilon^{\beta-1} & \text{if } \beta > 1. \end{cases}$$

We are now ready to prove Theorem 1.5. We shall do so for $t \to \infty$. The case $t \to 0^+$ is done similarly. The case $\beta < 1$ follows immediately from Corollary 3.3. Let us restrict our attention to the case $\beta \geqq 1$. We shall assume without loss of generality that $\varphi(x) \downarrow 0$ as $x \to \infty$. (When $\inf_x \varphi(x) > 0$, the result is simpler and also follows from the proof given below.)

Define $t_n \triangleq 2^n$, $\varphi_n \triangleq \varphi(t_n)$ and

$$E_n \triangleq \left\{ \inf_{t_n \leqq t \leqq t_{n+1}} S(X_t) \leqq 2^n \varphi_n \right\}. \tag{3.4}$$

Note that from Corollary 3.3,

$$\mathbf{P}^0(E_n) \asymp h(\varphi_n). \tag{3.5}$$

From the definition of $\mathfrak{J}(\varphi)$ given in (1.5), it follows that $\sum_n \mathbf{P}^0(E_n) < \infty$ if and only if $\mathfrak{J}(\psi) < \infty$, when $\psi(x) \triangleq A\varphi(Bx)$, for any $A, B > 0$.

Suppose $\mathfrak{J}(\varphi) < \infty$. The previous paragraph shows that for any $c > 0$,

$$\sum_n \mathbf{P}^0\Big(\inf_{t_n \leqq t \leqq t_{n+1}} |X_t| \leqq c 2^n \varphi(2^{n-1}) \Big) < \infty.$$

Since c is arbitrary and φ is decreasing, by the Borel–Cantelli lemma,

$$\liminf_{t\to\infty} \frac{|X_t|}{t\varphi(t)} = \infty,$$

$\mathbf{P}^0$-a.s.. Now suppose $\mathfrak{J}(\varphi) = \infty$. Note that Remark 2.2 holds with $|X.|$ replaced by $S(X.)$ everywhere. Using the Markov property and Lemma 3.1, for all n large enough,

$$\mathbf{P}^0(E_n \cap E_{n+k}) \leqq \mathbf{P}^0(E_n) \sup_x \mathbf{P}^x(E_{n+k}) \asymp \mathbf{P}^0(E_n)\mathbf{P}^0(E_{n+k}).$$

Theorem 1.5 follows from (3.4),(3.5), Kolmogorov's 0–1 law and the Kochen–Stone lemma ([KS]). □

REFERENCES

[BPP] Benjamini, I., Peres, Y. and Pemantle, R., Martin capacity for Markov chains, *Ann. Probab.* **23** (1995), 1332–1346.

[BG] Blumenthal, R. and Getoor, R. K., *Markov processes and potential theory*, Pure and Applied Mathematics, Vol. 29, Academic Press, New York, 1968. *MR* **41** #9348

[DE] Dvoretzky, A. and Erdős, P., Some problems on random walk in space, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950, University of California Press, Berkeley and Los Angeles, 1951, 353–367. *MR* **13**–852

[FS] Fitzsimmons, P. J. and Salisbury, T. S., Capacity and energy for multiparameter Markov processes, *Ann. Inst H. Poincaré Probab. Statist.* **25** (1989), 325–350. *MR* **91d**:60180

[H1] Hendricks, W. J., Hausdorff dimension in a process with stable components — an interesting counterexample, *Ann. Math. Statist.* **43** (1972), 690–694. *MR* **54** #14113

[H2] Hendricks, W.J., Multiple points for a process in $\mathbb{R}^2$ with stable components, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **28** (1973/74), 113–128. *MR* **50** #14950

[H3] Hendricks, W.J., Multiple points for a process in $\mathbb{R}^2$ with stable components, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **28** (1974), 113–128. *MR* **50** #14950

[KS] Kochen, S. B. and Stone, C. J., A note on the Borel–Cantelli lemma, *Illinois J. Math.* **8** (1964), 248–251. *MR* **28** #4562

[Pe] Peres, Y., Intersection-equivalence of Brownian paths and certain branching processes, *Comm. Math. Phys.* (to appear).

[PT] Perkins, E. A. and Taylor, S. J., Uniform measure results for the image of subsets under Brownian motion, *Probab. Theory Related Fields* **76** (1987), 257–289. *MR* **88m**:60122

[Sa] Salisbury, T. S., Energy, and intersection of Markov chains, *Proceedings of the IMS Workshop on Random Discrete Structures* (to appear).

[Sp] Spitzer, F., Some theorems concerning 2-dimensional Brownian motion, *Trans. Amer. Math. Soc.* **87** (1958), 187–197. *MR* **21** #3051

[T1] Takeuchi, J., On the sample paths of the symmetric stable processes in spaces, *J. Math. Soc. Japan* **16** (1964), 109–127. *MR* **29** #6537

[T2] Takeuchi, J., A local asymptotic law for the transient stable process, *Proc. Japan Acad.* **40** (1964), 141–144. *MR* **29** #6538

[TW] Takeuchi, J. and Watanabe, S., Spitzer's test for the Cauchy process on the line, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** (1964), 204–210. *MR* **30** #2545

(Received October 26, 1995)

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF UTAH
SALT LAKE CITY, UT 84112
U.S.A.

davar@math.utah.edu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 185–207

ON THE NUMBER OF COMPARISONS IN HOARE'S ALGORITHM "FIND"

B. KODAJ and T. F. MÓRI

Dedicated to Professor E. Csáki on the occasion of his sixtieth birthday

Abstract

In 1961 Hoare gave an extremely simple algorithm for finding the median from a list of size n. That algorithm was later investigated by Knuth, who derived a closed form expression for the expected number of comparisons. In the present paper we show that the (random) number of comparisons, divided by n, has a limit distribution as $n \to \infty$, and the rate of convergence measured in Wasserstein metric is $\Theta(\log n/n)$, while using other probabilistic distances, such as Φ-average compound distances with convex Young functions Φ, the rate of convergence is $\Theta(1/n)$.

1. Introduction

One basic problem in numerical data processing is to find, as quickly as possible, the k-th smallest element, say the median, out of n different numbers. This can be done without sorting the whole set of numbers by the help of a large variety of algorithms. If speed of algorithms is measured by the number of comparisons, the quickest algorithms only require $\mathcal{O}(n)$ comparisons even in the worst case. But if we are also satisfied with an algorithm of $\mathcal{O}(n)$ comparisons *on the average*, the most simple one, no doubt, is Hoare's algorithm FIND [7]: In order to select the k-th smallest of n numbers $a_1, a_2, \dots, a_n$, let us take the first one and compare it with all the other numbers. In that way the remaining numbers $a_2, a_3, \dots, a_n$ are divided into two groups according as they are less or greater than a_1. Then we can decide which group contains the element we are looking for, and what is its rank in that group, or — if we are lucky — it may turn out that a_1 is just the number we need. In the latter case we can stop immediately, but if not, we are given a new, smaller list to select from. Repeating subsequently the step above the process finally terminates with the desired element.

Knuth in [8] examined the average performance of Hoare's algorithm. Average is meant over all the $n!$ possible orderings, or, if one prefers probabilistic terms, it is the expected number of comparisons when the underlying distribution is uniform on the set of n-permutations. He was able to compute explicitly the expectation for every k, n $(1 \leqq k \leqq n)$. From his result it follows

1991 *Mathematics Subject Classification.* Primary 60F05; Secondary 68P10.

Key words and phrases. Median, coupling, Wasserstein metric, probability distances, Young functions.

0081–6906/97/\$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

that the expected number of comparisons, divided by n, converges to a finite limit as $n \to \infty$ and $k/n \to c$, $0 \leq c \leq 1$. In 1969 Singleton described a modification of the original algorithm, known as *median-of-three selection* [14]. Recently, Anderson and Brown have studied Hoare's and (Singleton's) algorithm from the combinatorist's point of view [1].

The first algorithm of linear cost even in the worst case was invented by Blum, Pratt, Tarjan, Floyd and Rivest [3]. Since then several papers have been devoted to pushing farther and farther down the upper bound for the worst case performance, see [11], [12], [13] or [15]. Bounds for the average case selections are derived in [3]. Generalizations of other types are also known, e.g., much effort was made to construct so-called parallel algorithms in various sorting and searching problems ([5] is only one example of those works).

The aim of the present note is to answer the following questions. Considering that the expected number of comparisons in Hoare's algorithm, divided by n, converges to a certain limit as $n \to \infty$ and $k/n \to c$, is it true that a limit distribution also exists? If yes, how can it be described? And what is the true rate of convergence?

For the thorough analysis let us repeat Hoare's algorithm in a more mathematical form leading to another interpretation which will then show what kind of limit distribution can be expected here.

The problem of finding the k-th smallest item from a list of n items $\{a_1, a_2, \dots, a_n\}$ will be addressed here as *the (n,k) selection problem.* Take the first item of the list and compare it with the remaining items. This makes $n-1$ comparisons and it turns out that a_1 is the j-th number in increasing order. Let $b_1, \dots, b_{j-1}$ denote the items less than a_1 and $c_1, \dots, c_{n-j}$ those greater than a_1. Now there are three possibilities.

If $j = k$, then a_1 is the number sought for.

If $j > k$, then we are reduced to the $(j-1, k)$ selection problem with the list $\{b_1, \dots, b_{j-1}\}$.

If $j < k$, then we are reduced to the $(n-j, k-j)$ selection problem with the list $\{c_1, \dots, c_{n-j}\}$.

If each permutation of $a_1, a_2, \dots, a_n$ is equally likely, then the events $\{j=1\}, \dots, \{j=n\}$ are also equiprobable. Let $X(n,k)$ denote the (random) number of comparisons needed to solve the (n,k) selection problem. Then one can easily check the following equations. The first one mirrors the symmetry between increasing and decreasing ordering. The rest of the lemma is a stochastic recursion for $X(n,i)$, $1 \leq i \leq n$.

LEMMA 1.1.

$$X(n,i) \stackrel{d}{=} X(n,n+1-i),$$
$$X(1,1) \equiv 0 \textit{ and}$$
$$X(n,i) \stackrel{d}{=} n-1+\sum_{j=1}^{i-1}\chi(j-1<W<j)X(n-j,i-j)$$
$$+\sum_{j=i}^{n-1}\chi(j<W<j+1)X(j,i)$$

where $\chi(\cdot)$ denotes the indicator of the event in brackets, W is uniformly distributed on $(0,n)$ and independent of the X's on the right-hand side. □

Making use of Lemma 1.1 one can write down a recursion for the expectations $\mathbf{E}X(n,i)$. Analysing the recurrence relation Knuth found the following interesting result. Let $H_n = 1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}$, then

$$\text{(1.1)}\quad \mathbf{E}X(n,k) = 2\Big((n+1)H_n-(n+3-k)H_{n+1-k}-(k+2)H_k+(n+3)\Big).$$

For $x\in[0,1]$ let $H(x,1-x)=x\log\frac{1}{x}+(1-x)\log\frac{1}{1-x}$, the entropy function (except that here log is meant to the base e). Then (1.1) easily implies that

$$\lim_{\substack{n\to\infty\\ k/n\to c}}\frac{1}{n}\mathbf{E}X(n,k)=2+2H(c,1-c).$$

In the second interpretation of $X(n,i)$ below we always choose at random from increasingly ordered lists instead of taking the first item of a randomly ordered list.

Let $U_1, U_2, \dots$ be independent random variables distributed uniformly on $(0,1)$, and let $1\le i\le n$. Let us start with the initial interval $I_0=(0,n)$ which we imagine sectioned into n pairwise disjoint unit intervals. These subintervals represent the items on the list in increasing order. We aim at selecting the item represented by the subinterval $(i-1,i)$. Let us choose a subinterval randomly: that can be done by observing the value $\lceil nU_1\rceil = j$, then subinterval $(j-1,j)$ represents the first item in the original list. Stop if $j=i$; and in the opposite case cut the interval $(j-1,j)$ out from I_0, thus obtaining two smaller intervals $(0,j-1)$ and (j,n) (one possibly empty). Let I_1 be the one containing the target unit interval. Then I_1 represents the remaining list after the first round of comparisons. In the second step we choose one from the unit intervals contained in I_1 at random, by taking the one with number $\lceil |I_1|U_2\rceil$. Then I_1 is bisected by cutting the chosen interval out from it, and I_2 is defined as that one of the two remaining subintervals which contains $(i-1,i)$, etc. In the k-th step I_{k-1} is bisected, and in doing that $|I_{k-1}|-1$ comparisons are made. Let τ be the number of steps needed till hitting the target interval, then

$$X(n,i)=\sum_{k=1}^{\tau}\Big(|I_{k-1}|-1\Big).$$

This model can be modified in a suitable way for obtaining a continuous analogue. That continuous model will be given in Section 2 together with its most important properties.

2. The continuous model

The second interpretation of Hoare's algorithm naturally leads to the following continuous analogue which can be considered as the limit of the discrete model as $n\to\infty$.

Let $U_1,U_2,\dots$ be independent random variables distributed uniformly on $(0,1)$, and let $0\leqq c\leqq a$. Let us start with the initial interval I_0 with endpoints $S_0=0$ and $T_0=a$. Let us divide I_0 into two subintervals by the random point $P_1=S_0+U_1(T_0-S_0)$. Let I_1 be the subinterval containing c. Thus the endpoints of I_1 are $S_1=P_1$, $T_1=T_0$ if $P_1<c$ and $S_1=S_0$, $T_1=P_1$ if $P_1\geqq c$. Similarly, for $j=2,3,\dots$ let the interval I_{j-1} be sectioned by the random point $P_j=S_{j-1}+U_j(T_{j-1}-S_{j-1})$, and I_j be the subinterval containing c, with endpoints $S_j=P_j$, $T_j=T_{j-1}$ if $P_j<c$ and $S_j=S_{j-1}$, $T_j=P_j$ if $P_j\geqq c$. Finally, let $Y(a,c)$ denote the sum of lengths of the random intervals I_j:

$$Y(a,c)=\sum_{j=0}^{\infty}|I_j|=\sum_{j=0}^{\infty}(T_j-S_j). \tag{2.1}$$

These quantities are meant to approximate the X's of the discrete model. The construction just described will be referred to as the *continuous model.*

It is easy to see that the infinite sum in (2.1) converges a.s. and in mean, too. Indeed, introducing the notations

$$x_{j-1}=(c-S_{j-1})/|I_{j-1}|,\qquad \eta_j=\max(U_j,\,1-U_j)$$

we have

$$|I_j|=\Big(\chi(U_j\leqq x_{j-1})(1-U_j)+\chi(U_j>x_{j-1})U_j\Big)|I_{j-1}|\leqq|I_{j-1}|\eta_j \tag{2.2}$$

hence

$$Y(a,c)\leqq a(1+\eta_1+\eta_1\eta_2+\eta_1\eta_2\eta_3+\cdots)=:a\xi \tag{2.3}$$

where η_j, $j=1,2,\dots$ are i.i.d. random variables, uniformly distributed on $(\frac{1}{2},1)$. Now it follows that

$$\mathbf{E}|I_j| \leqq \frac{3}{4}\mathbf{E}|I_{j-1}| \leqq \cdots \leqq \left(\frac{3}{4}\right)^j \mathbf{E}|I_0| = a\left(\frac{3}{4}\right)^j,$$

and so

$$\mathbf{E}Y(a,c) \leqq 4a.$$

From (2.3) it is also clear that all moments of $\mathbf{E}Y(a,c)$ are finite, because the same is true for ξ. More precisely, taking L_p-norm we obtain

$$\|\xi\|_p \leqq 1 + \|\eta_1\|_p + \|\eta_1\eta_2\|_p + \cdots = 1 + \|\eta_1\|_p + \|\eta_1\|_p^2 + \cdots = \frac{1}{1-\|\eta_1\|_p}.$$

Since

$$\|\eta_1\|_p = \left(\frac{2}{p+1}\left(1-\left(\frac{1}{2}\right)^{p+1}\right)\right)^{\frac{1}{p}} < \left(\frac{2}{p+1}\right)^{\frac{1}{p}},$$

for $p>1$ we have

$$\mathbf{E}(\xi^p) \leqq \frac{p+1}{2}\left(\left(\frac{p+1}{2}\right)^{\frac{1}{p}} - 1\right)^{-p} \leqq \left(\frac{p}{\log(\frac{p+1}{2})}\right)^p. \tag{2.4}$$

Sometimes we shall need a refinement of the continuous model. One striking difference between the constructions of X and Y is that the contribution of every single interval in Y is greater by 1 than that in the corresponding X. Therefore the distribution of $X(n,k)$ would be better approximated if we summed $(|I_k|-1)^+$ instead of $|I_k|$ in the continuous model. Thus, let $I_0, I_1, \dots$ be defined as above and

$$\mathcal{N} = \inf\{n:\ |I_n| < 1\},$$
$$Z(a,c) = \sum_{k=1}^{\mathcal{N}} \left(|I_{k-1}|-1\right).$$

For $a<1$ we have $\mathcal{N}=0$, hence $Z(a,c)=0$.

Some simple facts about the continuous model are collected in the following lemmas.

LEMMA 2.1.

$$Y(a,c) \overset{d}{=} Y(a,c-a), \qquad Z(a,c) \overset{d}{=} Z(a,c-a) \tag{2.5}$$

$$Y(a,c) \overset{d}{=} aY\left(1,\frac{c}{a}\right) \tag{2.6}$$

and

$$Y(a,c) \stackrel{d}{=} a + \chi(W<c)Y(a-W,c-W) + \chi(W \geqq c)Y(W,c), \tag{2.7}$$

$$Z(a,c) \stackrel{d}{=} (a-1)^+ + \chi(W<c)Z(a-W,c-W) + \chi(W \geqq c)Z(W,c) \tag{2.8}$$

where W is uniformly distributed on $(0,a)$ and independent of the Y's and Z's on the right-hand side. □

Therefore the expectation $f(c) = \mathbf{E}Y(1,c)$, $0 \leqq c \leqq 1$, satisfies the following equation

$$f(c) = 1 + \int_0^c (1-x) f\left(\frac{c-x}{1-x}\right) dx + \int_c^1 x f\left(\frac{c}{x}\right) dx. \tag{2.9}$$

It is not a great surprise that the solution of (2.9) is the limit of Knuth's formula

$$f(c) = 2 + 2H(c, 1-c) = \lim_{\substack{n\to\infty \\ k/n\to c}} \frac{1}{n}\mathbf{E}X(n,k),$$

thus we obtain the following result.

LEMMA 2.2. *For $0 \leqq c \leqq a$*

$$\mathbf{E}Y(a,c) = 2a(1+\log a) - 2c\log c - 2(a-c)\log(a-c),$$

where $0\log 0 = 0$. □

The assertion of our next lemma may be intuitively clear, but it needs a rigorous proof.

LEMMA 2.3. *Suppose $c' \leqq c$ and $a' - c' \leqq a - c$. Then $Y(a',c') \leqq_p Y(a,c)$, $Z(a',c') \leqq_p Z(a,c)$, where $\leqq_p$ denotes stochastic ordering.*

PROOF. The lemma can be proved via coupling technique. Let us define $Y(a',c')$ and $Y(a,c)$ by the help of the same sequence of i.i.d. random variables $U_1, U_2, \dots$. To this end let us shift I'_0 by $c-c'$ so that the target point, originally c', coincide with c. Thus let $S_0 = 0$, $T_0 = a$, and $S'_0 = c - c'$, $T'_0 = a' + c - c'$. Then still $I'_0 \subset I_0$. Firstly, construct P_j, I_j, $j = 1, 2, \dots$ as described in the beginning of this section. Then let the stopping times τ_j and intervals I'_j be defined successively as follows: $\tau_0 = 0$ and for $j \geqq 1$

$$\tau_j = \inf\{n > \tau_{j-1} :\ P_n \in I'_{j-1}\},$$
$$S'_j = P_{\tau_j} \ \text{ if } \ P_{\tau_j} < c, \ \text{ and } \ S'_j = S'_{j-1} \ \text{ if } \ P_{\tau_j} \geqq c,$$
$$T'_j = T_{j-1} \ \text{ if } \ P_{\tau_j} < c, \ \text{ and } \ T'_j = P_{\tau_j} \ \text{ if } \ P_{\tau_j} \geqq c.$$

In this way we clearly have $I'_j \subset I_{\tau_j}$. Since P_{τ_j} is uniformly distributed in I'_{j-1}, one can see that the joint distribution of the intervals I'_j, $j \geqq 0$, is identical with that described in the definition of $Y(a', c')$. Thus we obtain

$$Y(a', c') \stackrel{d}{=} \sum_{j=0}^{\infty} |I'_j| \leqq \sum_{j=0}^{\infty} |I_{\tau_j}| \leqq \sum_{j=0}^{\infty} |I_j| = Y(a, c),$$

and similarly

$$Z(a', c') \stackrel{d}{=} \sum_{j=0}^{\infty} (|I'_j| - 1)^+ \leqq \sum_{j=0}^{\infty} (|I_{\tau_j}| - 1)^+ \leqq \sum_{j=0}^{\infty} (|I_j| - 1)^+ = Z(a, c). \qquad \square$$

The next lemma shows how the continuous model relates in distribution to the discrete one.

LEMMA 2.4. *For every $n = 1, 2, \ldots,$ $i = 1, 2, \ldots, n$ and $i - 1 \leqq c \leqq i$ we have*

$$X(n, i) \leqq_p Z(n, c). \tag{2.10}$$

PROOF. This can be proved by induction in n. We have $X(1, 1) = Z(1, c) = 0$, and using the second interpretation of Hoare's algorithm we can write

$$X(n, i) \stackrel{d}{=} n - 1 + \sum_{j=1}^{i-1} \chi(j - 1 < W < j) X(n - j, i - j) + \sum_{j=1}^{n-1} \chi(j < W < j + 1) X(j, i),$$

where the random variable W is uniformly distributed on $(0, n)$ and independent of all X's appearing on the right-hand side. By the induction hypothesis

$$X(n, i) \leqq_p n - 1 + \sum_{j=1}^{i-1} \chi(j - 1 < W < j) Z(n - j, c - j) + \sum_{j=1}^{n-1} \chi(j < W < j + 1) Z(j, c).$$

From Lemma 2.3 and (2.8) it follows that

$$X(n,i) \leqq_p n-1+\sum_{j=1}^{i-1} \chi(j-1<W<j) Z(n-W, c-W)$$
$$+\sum_{j=1}^{n-1} \chi(j<W<j+1) Z(W, c)$$
$$\leqq_p n-1+\chi(W<c) Z(n-W, c-W)+\chi(W>c) Z(W, c) \stackrel{d}{=} Z(n, c). \qquad \square$$

From (2.10) it is clear that $Z(n,c)$ approximates $X(n,i)$ better than $Y(n,c)$ does, but the analogue of (2.6) does not remain true: $Z(\cdot,\cdot)$ is no more homogeneous.

3. The exact rate of convergence

Since all random variables mentioned so far are integrable, we can measure the rate of weak convergence in Wasserstein metric (also known as Gini's index of dissimilarity): for integrable random variables X and Y let

$$\kappa(X,Y)=\inf\{\mathbf{E}|X'-Y'|: X' \stackrel{d}{=} X,\ Y' \stackrel{d}{=} Y\},$$

where X' and Y' are defined on the same (suitably enlargened) probability space. This defines a metric on the space of (one dimensional) probability distributions with finite expectation. In terms of distribution functions one can write

$$\kappa(X,Y)=\int_{-\infty}^{+\infty} |F_X(t)-F_Y(t)|\,dt.$$

In case where one of the random variables stochastically majorizes the other one, simply $\kappa(X,Y)=|\mathbf{E}X-\mathbf{E}Y|$. Convergence in Wasserstein metric is equivalent to weak convergence plus convergence of expectations [10, Theorem 14.2.1].

At this point the first question of Section 1 can easily be answered.

THEOREM 3.1. *Let $0<c\leqq 1$ be fixed. Then the distribution of $\frac{1}{n}X(n,[cn])$ converges weakly to that of $Y(1,c)$ as $n\to\infty$.*

PROOF. From Lemma 2.4

$$\frac{1}{n}X(n,[cn]) \leqq_p \frac{1}{n}Y(n,cn) \stackrel{d}{=} Y(1,c),$$

hence

$$\kappa\Big(\frac{1}{n}X(n,[cn]), Y(1,c)\Big)=\mathbf{E}Y(1,c)-\frac{1}{n}\mathbf{E}X(n,[cn]) \to 0,$$

by Knuth's result. □

In this theorem $[cn]$ can be replaced by any sequence $(k_n,\ n \geqq 1)$ such that $1 \leqq k_n \leqq n$ and $\lim_{n\to\infty} k_n/n = c$, see the end of the section.

The main result of this section says that the rate of convergence in Theorem 3.1, measured by Wasserstein metric, is $\Theta(\frac{\log n}{n})$.

THEOREM 3.2. *Let* $\Delta(n,i) = \kappa(X(n,i), Y(n, i - \frac{1}{2}))$. *Then*

$$\Delta(n,i) \sim \log \frac{i^5(n-i+1)^5}{n^2}$$

holds uniformly in $1 \leqq i \leqq n$, *as* $n \to \infty$.

PROOF. Since $H_n = \log n + \gamma + \mathcal{O}(\frac{1}{n}) = \log(n - \frac{1}{2}) + \gamma + \mathcal{O}(\frac{1}{n})$, where γ is the Euler–Mascheroni constant, we obtain from (1.1) that

$$\begin{aligned}\mathbf{E}X(n,i) &= 2\Big((n+1)H_n - (n+3-i)H_{n+1-i} - (i+2)H_i + (n+3)\Big)\\ &= 2n\Big(1 + H\Big(\frac{i-0.5}{n}, 1 - \frac{i-0.5}{n}\Big)\Big)\\ &\quad - 5(\log i + \log(n-i+1)) + 2\log n + \mathcal{O}(1),\end{aligned}$$

thus

$$\Delta(n,i) = 5(\log i + \log(n-i+1)) - 2\log n + \mathcal{O}(1)$$

uniformly in i. □

REMARK 3.1. Comparing $X(n,i)$ with $Y(n+1,i)$ instead, one obtains

$$\kappa(X(n,i), Y(n+1,i)) = 4(\log i + \log(n+1-i)) + \mathcal{O}(1).$$

The following lemma estimates the modulus of continuity of the process $Y(1,c)$, $0 \leqq c \leqq 1$, in Wasserstein metric.

LEMMA 3.2. *Let* $0 \leqq c < c' \leqq 1/2$, $\delta = c' - c$. *Then*

$$\kappa(Y(1,c), Y(1,c')) < 4\delta\Big(1 + 2\log\frac{1}{\delta}\Big).$$

PROOF. Similarly to the proofs of Lemmas 2.3 and 2.4 coupling will be used again. We define $Y(a',c')$ and $Y(a,c)$ with one and the same sequence $U_1, U_2, \ldots$. Consider the stopping time

$$\tau = \inf\{n \geqq 1 : P_n \in (c, c')\} - \inf\{n : I_n \neq I'_n\}.$$

Then

$$|Y(1,c) - Y(1,c')| = \Big|\sum_{n=\tau}^{\infty}(|I_n| - |I'_n|)\Big| \leqq \sum_{n=1}^{\infty}\chi(\tau = n)\sum_{j=n}^{\infty}(|I_j| + |I'_j|).$$

Using (2.2) we obtain

$$|Y(1,c)-Y(1,c')| \leqq \sum_{n=1}^{\infty} \chi(\tau=n)(|I_n|+|I'_n|)\xi_n$$
$$=\sum_{n=1}^{\infty} \chi(\tau=n)|I_{n-1}|\xi_n=|I_{\tau-1}|\xi_n,$$

where the random variables $\xi_n=1+\eta_{n+1}+\eta_{n+1}\eta_{n+2}\dots$ are identically distributed. Since ξ_n and $|I_{n-1}|\chi(\tau=n)$ are independent for every n, so are ξ_τ and $|I_{\tau-1}|$. Let us determine the distribution of $|I_{\tau-1}|$.

Let us define $\tau_1=|\{n\leqq\tau:\ P_n>c\}|$, that is, if one only considers the (decreasing) subsequence of those P_j that are greater than c, τ_1 is the relative index of the first to fall in (c,c'). Similarly, we introduce $\tau_2=|\{n\leqq\tau:\ P_n<c'\}|$. Then $\tau=\tau_1+\tau_2-1$. Furthermore,

$$\tau_1 \stackrel{d}{=} \inf\Big\{n\geqq 1:\ V_1V_2\dots V_n<\frac{\delta}{1-c}\Big\},\quad T_{\tau-1}-c\stackrel{d}{=}(1-c)V_1V_2\dots V_{\tau_1-1},$$

and

$$\tau_2 \stackrel{d}{=} \inf\Big\{n\geqq 1:\ V_1V_2\dots V_n<\frac{\delta}{c'}\Big\},\quad c'-S_{\tau-1}\stackrel{d}{=}c'V_1V_2\dots V_{\tau_2-1},$$

where $V_1,V_2,\dots$ are i.i.d. random variables, uniformly distributed in $(0,1)$. Since $-\log(V_i)$ is of unit exponential distribution, we can take negative logarithm to bring our random variables into connection with a homogeneous Poisson process with unit intensity. For any $\lambda>0$ define

$$\nu=\max\{n\geqq 0:\ V_1V_2\dots V_n>e^{-\lambda}\}=\max\Big\{n\geqq 0:\ \sum_{i=1}^{n}(-\log V_i)\leqq\lambda\Big\},$$

then the distribution of $\lambda-\sum_{i=1}^{\nu}(-\log V_i)$ is exponential truncated at λ. With $\lambda=\log\frac{1-c}{\delta}$ we have

$$\begin{aligned} T_{\tau-1}-c &\stackrel{d}{=}(1-c)\exp\Big(-\sum_{i=1}^{\nu}(-\log V_i)\Big)= \\ &=(1-c)\exp(\mathcal{E}\wedge\lambda-\lambda)\stackrel{d}{=}\frac{\delta}{U}\wedge(1-c)\end{aligned}\tag{3.1}$$

where $\mathcal{E}$ is a mean 1 exponential random variable, $U=\exp(-\mathcal{E})$ is uniformly distributed on $(0,1)$ and $\wedge$ stands for minimum. Similarly,

$$c'-S_{\tau-1}\stackrel{d}{=}\frac{\delta}{U}\wedge c'.\tag{3.2}$$

Thus

$$\kappa(Y(1,c),Y(1,c')) \leqq \mathbf{E}|Y(1,c)-Y(1,c')| \leqq \mathbf{E}|I_{\tau-1}|\mathbf{E}\xi$$
$$=4\mathbf{E}\Big(\frac{\delta}{U}\wedge(1-c)+\frac{\delta}{U}\wedge c'-\delta\Big)<4\Big(\int_\delta^1\frac{\delta}{U}du+\delta\Big)=4\delta\Big(1+2\log\frac{1}{\delta}\Big). \quad \square$$

REMARK 3.2. Using (2.6) we immediately obtain for $0\leqq c<c'\leqq\frac{a}{2}$ that

$$\kappa(Y(a,c),Y(a,c'))<4\delta\Big(1+2\log\frac{a}{\delta}\Big). \tag{3.3}$$

On the other hand,

$$\kappa(Y(a,c),Y(a,c')) \geqq |\mathbf{E}Y(a,c)-\mathbf{E}Y(a,c')|$$
$$=2a\Big(H\Big(\frac{c'}{a},1-\frac{c'}{a}\Big)-H\Big(\frac{c}{a},1-\frac{c}{a}\Big)\Big).$$

$H(x,1-x)$ is convex, therefore the right-hand side attains its maximum for fixed a and δ when $c=0$ and $c'=\delta$. Then

$$\kappa(Y(a,0),Y(a,\delta)) \geqq 2aH\Big(\frac{\delta}{a},1-\frac{\delta}{a}\Big)>2\delta\log\frac{a}{\delta}.$$

Thus, the estimate of (3.3) for the modulus of continuity is sharp up to the order of magnitude.

COROLLARY 3.1. *Let the sequence of integers* $(k_n, n\geqq 1)$ *satisfy* $1\leqq k_n\leqq n$ *and* $\lim_{n\to\infty}\frac{k_n}{n}=c$. *Define* $\delta_n=\left|\frac{k_n}{n}-c\right|+\frac{1}{n}$, *then*

$$\kappa\Big(\frac{1}{n}X(n,k_n),Y(1,c)\Big)=\mathcal{O}\Big(\delta_n\log\frac{1}{\delta_n}\Big).$$

PROOF. This immediately follows from Theorem 3.2 and Lemma 3.2. $\square$

4. Convergence of higher moments

The surprising shortness of the proof of Theorem 3.2 was due to the fact that Knuth had been able to express $\mathbf{E}X(n,i)$ in a closed form. However, there exists another way of estimating the rate of convergence based on (2.7) and Lemma 1.1, which can be applied for other probability metrics as well. This method will require the following lemma.

LEMMA 4.1. *Let $\Delta(n,i)$, $1 \leqq i \leqq n$, $n=1,2,\dots$ be arbitrary nonnegative numbers satisfying*

$$\Delta(n,i) \leqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j)+\frac{1}{n}\sum_{j=i}^{n-1}\Delta(j,i)+n^p\varepsilon(n) \tag{4.1}$$

where $\{\varepsilon(n), n \geqq 1\}$ is a nondecreasing sequence of positive numbers and $p \geqq 0$. Then, for $p=0$

$$\Delta(n,i) \leqq \varepsilon(n)\Big(1+\log i+\log(n-i+1)\Big), \tag{4.2}$$

and for $p>0$

$$\Delta(n,i) \leqq \frac{p+1}{p-1+2^{-p}}n^p\varepsilon(n). \tag{4.3}$$

On the other hand, if

$$\Delta(n,i) \geqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j)+\frac{1}{n}\sum_{j=1}^{n-1}\Delta(j,i)+\varepsilon(n) \tag{4.4}$$

with a nonincreasing nonnegative sequence $\{\varepsilon(n), n \geqq 1\}$, then

$$\Delta(n,i) \geqq \frac{1}{2}\varepsilon(n)\Big(1+\log i+\log(n-i+1)\Big). \tag{4.5}$$

REMARK 4.1. If $p>0$ and

$$\Delta(n,i) \geqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j)+\frac{1}{n}\sum_{j=1}^{n-1}\Delta(j,i)+n^p\varepsilon(n)$$

with an arbitrary nonnegative sequence $\{\varepsilon(n), n \geqq 1\}$, then even the trivial lower estimation $\Delta(n,i)>n^p\varepsilon(n)$ attains the order of magnitude of the upper estimate (4.3).

PROOF. The proof will be performed by induction in n.

Let us first deal with the upper estimate. For $n=i=1$ (4.2) and (4.3) clearly hold. For $n \geqq 2$ let us apply the induction hypothesis and the monotonicity property of $\varepsilon(n)$ in (4.1). For $p=0$ we obtain

$$\Delta(n,i) \leqq \varepsilon(n)+\frac{1}{n}\sum_{j=1}^{i-1}\varepsilon(n-j)(1+\log(i-j)+\log(n-i+1))+$$
$$+\frac{1}{n}\sum_{j=i}^{n-1}\varepsilon(j)(1+\log i+\log(j-i+1))$$

$$\leqq \varepsilon(n)\Bigg(1+\frac{1}{n}\sum_{j=1}^{i-1}(1+\log(i-j)+\log(n-i+1))$$
$$+\frac{1}{n}\sum_{j=i}^{n-1}(1+\log i+\log(j-i+1))\Bigg)$$
$$\leqq \varepsilon(n)\Bigg(1+\frac{i-1}{n}\log(n-i+1)+\frac{n-i}{n}\log i+\frac{1}{n}\sum_{j=1}^{i-1}(1+\log j)$$
$$+\frac{1}{n}\sum_{j=1}^{n-i}(1+\log j)\Bigg)$$
$$\leqq \varepsilon(n)\Bigg(1+\frac{i-1}{n}\log(n-i+1)+\frac{n-i}{n}\log i+\frac{1}{n}\int_1^i(1+\log t)dt$$
$$+\frac{1}{n}\int_1^{n-i+1}(1+\log t)dt\Bigg)$$
$$=\varepsilon(n)\Big(1+\log i+\log(n-i+1)\Big).$$

For $p>0$ let C denote $\frac{p+1}{p-1+2^{-p}}$, then we similarly have

$$\Delta(n,i)\leqq n^p\varepsilon(n)+C\frac{1}{n}\Big(\sum_{j=1}^{i-1}(n-j)^p\varepsilon(n-j)+\sum_{j=i}^{n-1}j^p\varepsilon(j)\Big)$$
$$\leqq \varepsilon(n)\Big(n^p+C\frac{1}{n}\Big(\sum_{j=n-i+1}^{n-1}j^p+\sum_{j=i}^{n-1}j^p\Big)\Big)$$
$$\leqq \varepsilon(n)\Big(n^p+C\frac{1}{n(p+1)}\Big(\int_{n-i+1}^{n}t^p dt+\int_i^n t^p dt\Big)\Big)$$
$$=\varepsilon(n)\Big(n^p+C\frac{1}{n(p+1)}(2n^{p+1}-(n-i+1)^{p+1}-i^{p+1})\Big).$$

Let us make use of the convexity of the power function t^{p+1}.

$$\Delta(n,i)\leqq \varepsilon(n)\Big(n^p+C\frac{2}{n(p+1)}\Big(n^{p+1}-\Big(\frac{n+1}{2}\Big)^{p+1}\Big)\Big)$$
$$\leqq n^p\varepsilon(n)\Big(1+C\frac{2}{p+1}(1-2^{-p-1})\Big)=Cn^p\varepsilon(n).$$

The lower estimate is deduced in a similar way. $\Delta(1,1) > \varepsilon(1)/2$, and from (4.4) and the induction hypothesis we get

$$\Delta(n,i) \geqq \varepsilon(n) + \frac{1}{n}\sum_{j=1}^{i-1}\frac{1}{2}\varepsilon(n-j)(1+\log(i-j)+\log(n-i+1))+$$

$$+\frac{1}{n}\sum_{j=i}^{n-1}\frac{1}{2}\varepsilon(j)(1+\log i+\log(j-i+1))$$

$$\geqq \frac{1}{2}\varepsilon(n)\Bigg(2+\frac{1}{n}\sum_{j=1}^{i-1}(1+\log(i-j)+\log(n-i+1))$$

$$+\frac{1}{n}\sum_{j=i}^{n-1}(1+\log i+\log(j-i+1))\Bigg)$$

$$=\frac{1}{2}\varepsilon(n)\Bigg(2+\frac{i-2}{n}\log(n-i+1)+\frac{n-i+1}{n}\log i+\frac{1}{n}\sum_{j=2}^{i}(1+\log j)$$

$$+\frac{1}{n}\sum_{j=2}^{n-i+1}(1+\log j)\Bigg)$$

$$\geqq \frac{1}{2}\varepsilon(n)\Bigg(2+\frac{i-2}{n}\log(n-i+1)+\frac{n-i+1}{n}\log i+\frac{1}{n}\int_1^i(1+\log t)dt$$

$$+\frac{1}{n}\int_1^{n-i+1}(1+\log t)dt\Bigg)$$

$$=\frac{1}{2}\varepsilon(n)\Big(2+\frac{n-1}{n}(\log i+\log(n-i+1))\Big).$$

Since

$$\frac{1}{n}(\log i+\log(n-i+1))<1,$$

we have

$$\Delta(n,i)\geqq\frac{1}{2}\varepsilon(n)\Big(1+\log i+\log(n-i+1)\Big),$$

as asserted in (4.5). □

For the generalization of Theorem 3.2 and Corollary 3.1 let us replace the Wasserstein metric with the broader family of so called Φ-*average compound distances*, see [10, Example 3.3.1].

Let φ: $[0,\infty) \to [0,\infty)$ be a nondecreasing, left continuous function, $\varphi(0)=0$, and let ψ denote its generalized inverse: $\psi(s)=\sup\{t\geqq 0\colon \varphi(t)<s\}$. Define

$$\Phi(x)=\int_0^x \varphi(t)dt, \qquad \Psi(y)=\int_0^y \psi(s)ds.$$

Φ and Ψ are called a *pair of conjugate convex Young functions.* Several properties of convex Young functions and the corresponding Orlicz spaces are contained in [9]. Some important facts we state below.

A convex Young function Φ is said to satisfy *Orlicz's condition* if

$$p=p_\Phi=\sup_x \frac{x\varphi(x)}{\Phi(x)}<\infty.$$

This p is called *the characteristic exponent* of Φ. In that case the function is $\Phi(x)x^{-p}$ nonincreasing, hence $\Phi(x)\leqq\Phi(1)x^p$, $x\geqq 1$, and

$$K_\Phi=:\sup_x \frac{\Phi(2x)}{\Phi(x)}\leqq 2^p.$$

If the conjugate function Ψ also satisfies Orlicz's condition with characteristic exponent q, then $p^{-1}+q^{-1}\leqq 1$, and

$$\inf_x \frac{x\varphi(x)}{\Phi(x)}=\frac{q}{q-1}.$$

Furthermore, $\Phi(x)/x^{\frac{q}{q-1}}$ is nondecreasing, hence $\Phi(x)\geqq\Phi(1)x^{\frac{q}{q-1}}$, $x\geqq 1$, and

$$\inf_x \frac{\Phi(2x)}{\Phi(x)}\geqq 2^{\frac{q}{q-1}}.$$

For a convex Young function Φ satisfying Orlicz's condition let us define

$$\hat{\mu}_\Phi(X,Y)=\inf\{\mathbf{E}\Phi(|X'-Y'|)\colon X\overset{d}{=}X',\ Y\overset{d}{=}Y'\}$$

where X' and Y' are defined on the same probability space. This is a distance on the space of probability distributions with finite Φ-moments (the only difference from being a metric is the presence of a constant factor in the triangle inequality which now looks as $\hat{\mu}_\Phi(X,Y)\leqq K_\Phi(\hat{\mu}_\Phi(X,Z)+\hat{\mu}_\Phi(Z,Y))$. By the Cambanis–Simons–Stout formula [4]

$$\hat{\mu}_\Phi(X,Y)=\int_0^1 \Phi(|F_X^{-1}(t)-F_Y^{-1}(t)|)dt,$$

which shows that in the definition the minimum is attained when X' and Y' are related through the quantile transform. In the particular case $\Phi(t)=t^p$ $(p>1)$, $(\hat{\mu}_\Phi)^{\frac{1}{p}}$ is a metric, which induces weak convergence plus convergence of moments of order not greater than p. Formally, the definition of $\hat{\mu}_\Phi$ has sense for $\Phi(t)=t$, and it gives back the Wasserstein metric κ. In general, Jensen's inequality gives $\hat{\mu}_\Phi(X,Y)\geqq\Phi(\kappa(X,Y))$.

Since Φ is superadditive, $X\leqq_p Z\leqq_p Y$ implies $\hat{\mu}_\Phi(X,Z)+\hat{\mu}_\Phi(Z,Y)\leqq\hat{\mu}_\Phi(X,Y)$.

In the sequel we are going to estimate both distances $\hat{\mu}_\Phi(X(n,i), Y(n,i-\frac{1}{2}))$ and $\hat{\mu}_\Phi(\frac{1}{n}X(n,i),Y(1,\frac{i-0.5}{n}))$. They are no longer equivalent, for $\hat{\mu}_\Phi$ is not homogeneous unless $\Phi(x)=Cx^p$. That will be done in two steps: we first estimate the distance between $X(n,i)$ and $Z(n,i-\frac{1}{2})$, then between $Z(n,i-\frac{1}{2})$ and $Y(n,i-\frac{1}{2})$. One part of the proof deserves to be separated as a lemma.

LEMMA 4.2. *Let $\mathcal{N}$ be the stopping time in the definition of $Z(a,c)$, $a>1$. Then*

$$\mathbf{E}\Phi(\mathcal{N})\leqq\mathbf{E}\Phi(1+\pi_\lambda),$$

where π_λ is a Poisson random variable with expectation $\lambda=(\log a)/(1-\log 2)$.

PROOF. From (2.2) it follows that $|I_k|\leqq a\eta_1\eta_2\dots\eta_k$, where η_i's are independent random variables, each distributed uniformly on $(\frac{1}{2},1)$. Hence

$$\mathcal{N}\leqq\min\Big\{n:\sum_{k=1}^{n}(-\log\eta_k)>\log a\Big\}.$$

The distribution of the random variables $-\log\eta_i$ is equal to the conditional distribution of a standard exponential random variable given that it is less than $\log 2$. This distribution possesses the aging property NBUE (in fact, it even belongs to the smaller subclass IFR). Consider the renewal process defined by these variables, and let $N(t)$ denote the number of renewals before time t. According to Theorem 3.17 on p. 173 of [2], for arbitrary convex increasing function f

$$\mathbf{E}f(N(t))\leqq\mathbf{E}f(\pi_{t/\mu})$$

where $\mu=\mathbf{E}(-\log\eta_1)=1-\log 2$. Clearly, $\mathcal{N}=N(\log a)+1$, so let $f(x)=\Phi(x+1)$. □

THEOREM 4.1. *Suppose both Φ and its conjugate Ψ satisfy Orlicz's condition. Then*

$$\max_i\hat{\mu}_\Phi\Big(X(n,i),Y\Big(n,i-\frac{1}{2}\Big)\Big)=\Theta\Big(\frac{\Phi(n)}{n}\Big),\tag{4.6}$$

$$\max_i\hat{\mu}_\Phi\Big(\frac{1}{n}X(n,i),Y\Big(1,\frac{i-0.5}{n}\Big)\Big)=\Theta\Big(\frac{1}{n}\Big).\tag{4.7}$$

PROOF. The proof is based on the representation of $X(n,i)$ given by Lemma 1.1 and formulas (2.7)–(2.8) obtained for the continuous model. Then different ways of coupling will yield lower and upper estimations, respectively.

Let us start with the upper estimation in (4.6).

$$\begin{aligned}(4.8)\qquad &\hat\mu_\Phi\Big(Y\Big(n,i-\frac12\Big),X(n,i)\Big)\\ &\leqq K_\Phi\Big(\hat\mu_\Phi\Big(Z\Big(n,i-\frac12\Big),X(n,i)\Big)+\hat\mu_\Phi\Big(Y\Big(n,i-\frac12\Big),Z\Big(n,i-\frac12\Big)\Big)\Big).\end{aligned}$$

The first term on the right-hand side is the first to be treated. On the one hand,

$$\begin{aligned}X(n,i)\overset{d}{=}n-1+&\sum_{j=1}^{i-1}\chi(j-1<W<j)X(n-j,i-j)\\ &+\sum_{j=i}^{n-1}\chi(j<W<j+1)X(j,i)\end{aligned}$$

and on the other hand,

$$\begin{aligned}Z\Big(n,i-\frac12\Big)&\overset{d}{=}n-1+\chi\Big(W'<i-\frac12\Big)Z\Big(n-W',i-\frac12-W'\Big)\\ &\qquad+\chi\Big(i-\frac12<W'\Big)Z\Big(W',i-\frac12\Big)\\ &=n-1+\sum_{j=1}^{i-1}\chi(j-1<W'<j)Z\Big(n-W',i-\frac12-W'\Big)+\\ &\qquad+\sum_{j=i}^{n-1}\chi(j<W'<j+1)Z\Big(W',i-\frac12\Big)\\ &\qquad+\chi\Big(i-1<W'<i-\frac12\Big)Z\Big(n-W',i-\frac12-W'\Big)\\ &\qquad+\chi\Big(i-\frac12<W'<i\Big)Z\Big(W',i-\frac12\Big)\end{aligned}$$

with W and W' uniformly distributed over $(0,n)$.

By Lemma 2.3 we have

$$Z\Big(n,i-\frac12\Big)\leqq_p n-1+\sum_{j=1}^{i-1}\chi(j-1<W'<j)Z\Big(n-j+1,i-j+\frac12\Big)$$

$$+\sum_{j=i}^{n-1}\chi(j<W'<j+1)Z\Big(j+1,i-\frac{1}{2}\Big)$$
$$+\chi(i-1<W'<i-\frac{1}{2})Z(n-i+1,\frac{1}{2})+\chi(i-\frac{1}{2}<W'<i)Z(i,i-\frac{1}{2}).$$

(Here W' is independent of all the other random variables appearing on the same side.)

Consider the following one-to-one correspondence between W and W'. Let

$$\begin{aligned} W' = & W+1 && \text{on the event } \{0<W<i-\tfrac{3}{2}\},\\ & W-i+\tfrac{3}{2} && \text{on the event } \{i-\tfrac{3}{2}<W<i-\tfrac{1}{2}\},\\ & W+n-i-\tfrac{1}{2} && \text{on the event } \{i-\tfrac{1}{2}<W<i+\tfrac{1}{2}\},\\ & W-1 && \text{on the event } \{i+\tfrac{1}{2}<W<n\}.\end{aligned}$$

Then

$$Z(n,i-\frac{1}{2})\leqq_p n-1+\chi(i-\frac{3}{2}<W<i+\frac{1}{2})Z(n,i-\frac{1}{2})$$
$$+\sum_{j=1}^{i-2}\chi(j-1<W<j)Z(n-j,i-j-\frac{1}{2})$$
$$+\chi(i-2<W<i-\frac{3}{2})Z(n-i+1,\frac{1}{2})$$
$$+\sum_{j=i+1}^{n-1}\chi(j<W<j+1)Z(j,i-\frac{1}{2})+\chi(i+\frac{1}{2}<W<i+1)Z(i,i-\frac{1}{2}).$$

Hence

$$Z(n,i-\frac{1}{2})-X(n,i)$$
$$\leqq_p \sum_{j=1}^{i-2}\chi(j-1<W<j)\Big(Z(n-j,i-j-\frac{1}{2})-X(n-j,i-j)\Big)$$
$$+\chi(i-2<W<i-\frac{3}{2})\Big(Z(n-i+1,\frac{1}{2})-X(n-i+1,1)\Big) \tag{4.9}$$
$$+\sum_{j=i+1}^{n-1}\chi(j<W<j+1)\Big(Z(j,i-\frac{1}{2})-X(j,i)\Big)$$

$$+\chi\Big(i+\frac{1}{2}<W<i+1\Big)\Big(Z(i,i-\frac{1}{2})+X(i,i)\Big)$$
$$+\chi\Big(i-\frac{3}{2}<W<i+\frac{1}{2}\Big)Z\Big(n,i-\frac{1}{2}\Big).$$

Let each X in the right-hand side be the quantile transform of the corresponding Z. Then putting (4.9) into Φ and taking expectation we obtain

$$\Delta(n,i)=:\hat{\mu}_\Phi\Big(Z\Big(n,i-\frac{1}{2}\Big),X(n,i)\Big)$$
$$\leqq\frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j)+\frac{1}{n}\sum_{j=i}^{n-1}\Delta(j,i)+\frac{2}{n}\mathbf{E}\Phi\Big(Z\Big(n,i-\frac{1}{2}\Big)\Big).$$

Then, using Orlicz's condition on Φ, we have

$$\mathbf{E}\Phi\Big(Z\Big(n,i-\frac{1}{2}\Big)\Big)\leqq\mathbf{E}\Phi(n\xi)\leqq\mathbf{E}(\xi^p)\Phi(n).$$

Orlicz's condition on Ψ implies that

$$\frac{2}{n}\mathbf{E}\Phi\Big(Z\Big(n,i-\frac{1}{2}\Big)\Big)\leqq 2\mathbf{E}(\xi^p)\frac{\Phi(n)}{n}=n^{\frac{1}{q-1}}\varepsilon(n)$$

where $\varepsilon(n)=2\mathbf{E}(\xi^p)\Phi(n)/n^{\frac{q}{q-1}}$ is nondecreasing. Thus, (4.1) is satisfied. Hence (4.3) gives

$$\hat{\mu}_\Phi(Z(n,i-\frac{1}{2}),X(n,i))=\mathcal{O}\Big(\frac{\Phi(n)}{n}\Big). \tag{4.10}$$

Let us continue with the second term on the right-hand side of (4.8). By definition

$$Y(a,c)=Z(a,c)+\mathcal{N}+Y(T_\mathcal{N}-S_\mathcal{N},c-S_\mathcal{N}),$$

hence

$$Y(a,c)-Z(a,c)\leqq_p\mathcal{N}+\xi,$$

where ξ is independent of $\mathcal{N}$. From this it follows that

$$\hat{\mu}_\Phi(Y(n,i-\frac{1}{2}),Z(n,i-\frac{1}{2}))<\mathbf{E}\Phi(\mathcal{N}+\xi)<K_\Phi(\mathbf{E}\Phi(\mathcal{N})+\mathbf{E}\Phi(\xi)).$$

The latter does not change with n. As to $\mathbf{E}\Phi(\mathcal{N})$, it can be estimated by Lemma 4.2:

$$\mathbf{E}\Phi(\mathcal{N})\leqq\mathbf{E}\Phi(1+\pi_\lambda)\leqq\Phi(1)\mathbf{E}((1+\pi_\lambda)^p),$$

where π_λ is a Poisson random variable with expectation $\lambda=(\log n)/(1-\log 2)$. It is easy to see that $\mathbf{E}((1+\pi_\lambda)^p)\sim\lambda^p$ for fixed $p>0$ and increasing λ. Hence $\mathbf{E}\Phi(\mathcal{N})=\mathcal{O}((\log n)^p)$ as $n\to\infty$. This is asymptotically negligible compared to $\Phi(n)/n \geqq \Phi(1)n^{\frac{1}{q-1}}$. Combining this with (4.10) we obtain

$$\hat{\mu}_\Phi\Big(Y\Big(n,i-\frac{1}{2}\Big),X(n,i)\Big)=\mathcal{O}\Big(\frac{\Phi(n)}{n}\Big).$$

The lower estimation in (4.6) is quite obvious. Since $\mathbf{P}(X(n,i)=n-1)\geqq\frac{1}{n}$ and $Y(n,i-\frac{1}{2})\geqq n+\min(i-\frac{1}{2},n-i+\frac{1}{2})$, clearly

$$\hat{\mu}_\Phi\Big(Y\Big(n,i-\frac{1}{2}\Big),X(n,i)\Big)\geqq\frac{1}{n}\Phi(\min(i,n-i)).$$

Thus, with $0<\varepsilon<\frac{1}{2}$, for every i between $n\varepsilon$ and $n(1-\varepsilon)$ we have

$$\hat{\mu}_\Phi(Y(n,i-\frac{1}{2}),X(n,i))\geqq\frac{1}{n}\Phi(n\varepsilon)\geqq\varepsilon^p\frac{\Phi(n)}{n}. \tag{4.11}$$

The proof of (4.6) is completed.

Estimation (4.7) can be proved either in a similar way, or by applying (4.6) with the Young function $\overline{\Phi}(x)=\Phi\Big(\frac{x}{n}\Big)$. Its conjugate is $\overline{\Psi}(y)=\Psi(ny)$ and $\overline{p}=p$, $\overline{q}=q$. Since all estimations in the proof of (4.6) depended on Φ only through p and q, (4.7) immediately follows. □

REMARK 4.2. Now we outline how to estimate $\Delta(n,i)=\kappa(X(n,i), Y(n,i-\frac{1}{2}))$ without Knuth's explicite formula, only by using Lemma 4.1. Namely, we shall prove that

$$1+\log i+\log(n-i+1)\leqq\Delta(n,i)\leqq C\Big(1+\log i+\log(n-i+1)\Big) \tag{4.12}$$

holds for every $1\leqq i\leqq n$, $n=1,2,\dots$, with $C=4(1+\log 2)$.

On the one hand, in a similar way that led to (4.9) we obtain

$$\begin{aligned} &Y\Big(n,i-\frac{1}{2}\Big)-X(n,i)\\ &\leqq_p\sum_{j=1}^{i-1}\chi(j-1<W<j)\Big(Y(n-j,i-j-\frac{1}{2})-X(n-j,i-j)\Big)\\ &\quad+\sum_{j=i}^{n-1}\chi(j<W<j+1)\Big(Y\Big(j,i-\frac{1}{2}\Big)-X(j,i)\Big)+1\\ &\quad+\chi(i-\frac{3}{2}<W<i+\frac{1}{2})Y(n,i-\frac{1}{2})\\ &-\chi(i-\frac{3}{2}<W<i-1)Y(n-i+1,\frac{1}{2})-\chi(i<W<i+\frac{1}{2})Y(i,i-\frac{1}{2}). \end{aligned}$$

Hence

$$\Delta(n,i) \leqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j) + \frac{1}{n}\sum_{j=i}^{n-1}\Delta(j,i)$$
$$+\frac{2}{n}\mathbf{E}Y\Big(n,i-\frac{1}{2}\Big)+1-\frac{1}{2n}\Big(\mathbf{E}Y\Big(n-i+1,\frac{1}{2}\Big)+\mathbf{E}Y\Big(i,i-\frac{1}{2}\Big)\Big)$$
$$\leqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j) + \frac{1}{n}\sum_{j=i}^{n-1}\Delta(j,i) + C,$$

that is, (4.1) holds with $\varepsilon(n) = C$.

For getting a lower estimation we can proceed analogously. Let us build $Y(n, i-\frac{1}{2})$ with the same uniformly distributed random variable W that we used for constructing $X(n,i)$.

$$Y\Big(n,i-\frac{1}{2}\Big) \overset{d}{=} n + \sum_{j=1}^{i-1}\chi(j-1<W<j)Y\Big(n-W,i-\frac{1}{2}-W\Big)$$
$$+\sum_{j=i}^{n-1}\chi(j<W<j+1)Y\Big(W,i-\frac{1}{2}\Big)$$
$$+\chi\Big(i-1<W<i-\frac{1}{2}\Big)Y\Big(n-W,i-\frac{1}{2}-W\Big)+\chi\Big(i-\frac{1}{2}<W<i\Big)Y\Big(W,i-\frac{1}{2}\Big)$$
$$\geqq_p n + \sum_{j=1}^{i-1}\chi(j-1<W<j)Y\Big(n-j,i-j-\frac{1}{2}\Big)$$
$$+\sum_{j=i}^{n-1}\chi(j<W<j+1)Y\Big(j,i-\frac{1}{2}\Big)$$
$$+\chi\Big(i-1<W<i-\frac{1}{2}\Big)Y\Big(n-i+\frac{1}{2},0\Big)+\chi\Big(i-\frac{1}{2}<W<i\Big)Y\Big(i-\frac{1}{2},i-\frac{1}{2}\Big).$$

Comparing this with the representation of $X(n,i)$ we can see that

$$Y(n,i-\frac{1}{2}) - X(n,i)$$
$$\leqq_p \sum_{j=1}^{i-1}\chi(j-1<W<j)\Big(Y(n-j,i-j-\frac{1}{2})-X(n-j,i-j)\Big)$$
$$+\sum_{j=i}^{n-1}\chi(j<W<j+1)\Big(Y(j,i-\frac{1}{2})-X(j,i)\Big)$$
$$+\chi\Big(i-1<W<i-\frac{1}{2}\Big)Y\Big(n-i+\frac{1}{2},0\Big)+\chi\Big(i-\frac{1}{2}<W<i\Big)Y\Big(i-\frac{1}{2},0\Big)+1.$$

This, after integration, leads us to

$$\Delta(n,i) \geqq \frac{1}{n}\sum_{j=1}^{i-1}\Delta(n-j,i-j)+\frac{1}{n}\sum_{j=i}^{n-1}\Delta(j,i)+2,$$

that is, (4.3) is satisfied with $\varepsilon(n)=2$.

Now Lemma 4.1 completes the proof of (4.12). □

Finally, we prove the analogue of Corollary 3.1. First, let us extend Lemma 3.2 to $\hat{\mu}_\Phi(Y(a,c),Y(a,c'))$.

LEMMA 4.3. *Suppose both* Φ *and* Ψ *comply with Orlicz's condition. Let* $0 \leqq c < c' \leqq a/2$, $\delta = c' - c$. *Then*

$$\hat{\mu}_\Phi(Y(a,c),Y(a,c'))=\mathcal{O}\Big(\frac{\Phi(a)}{a}\delta\Big),$$

where the constant involved in the $\mathcal{O}$ *notation only depends on* Φ.

PROOF. From (3.1) and (3.2) it follows that

$$\begin{aligned}\Delta =: \hat{\mu}_\Phi(Y(a,c),Y(a,c')) &\leqq \mathbf{E}\Phi(|Y(a,c)-Y(a,c')|)\\ &= \mathbf{E}\Phi(a|Y(1,c)-Y(1,c')|) \leqq \mathbf{E}\Phi(a|I_{\tau-1}|\xi_\tau)\\ \leqq \mathbf{E}(\xi_\tau^p(a|I_{\tau-1}|)) &\leqq K_\Phi^2\mathbf{E}(\xi^p)\mathbf{E}\Big(\Phi\Big(\frac{\delta}{U}\wedge(a-c)\Big)+\Phi\Big(\frac{\delta}{U}\wedge c'\Big)+\Phi(\delta)\Big)\end{aligned}$$

by the triangle inequality for Φ. Here U is uniformly distributed on $(0,1)$. Thus, using the convexity of Φ and the growth condition on Ψ we can write

$$\Delta \leqq K_\Phi^2\mathbf{E}(\xi^p)\Big(2\mathbf{E}\Phi\Big(\frac{\delta}{U}\wedge a\Big)+\Phi(\delta)\Big)=K_\Phi^2\mathbf{E}(\xi^p)\Big(2\int_{\delta/a}^{1}\Phi\Big(\frac{\delta}{U}\Big)du+\Phi(\delta)\Big)$$

$$\leqq K_\Phi^2\mathbf{E}(\xi^p)\Big(2\int_{\delta/a}^{1}\Phi(a)(\delta/au)^{\frac{q}{q-1}}du+\delta\Phi(a)/a\Big)$$

$$\leqq K_\Phi^2\mathbf{E}(\xi^p)\Phi(a)\frac{\delta}{a}(2q-1). \qquad \square$$

COROLLARY 4.2. *Let the sequence of integers* $(k_n, n \geqq 1)$ *satisfy* $1 \leqq k_n \leqq n$ *and* $\lim_{n\to\infty}\frac{k_n}{n}=c$. *Define again* $\delta_n=\left|\frac{k_n}{n}-c\right|+\frac{1}{n}$, *then*

$$\hat{\mu}_\Phi\Big(\frac{1}{n}X(n,k_n)Y(1,c)\Big)=\mathcal{O}(\delta_n). \qquad \square$$

ACKNOWLEDGEMENT. This work was supported by the Hungarian National Foundation for Scientific Research Grant Nos. T-016384, T-016237.

NOTE added in proof. After the final version of the present paper had been accepted, similar results were published by R. Grübel and U. Rösler (*Adv. Appl. Probab.* **28** (1996), 252–269).

REFERENCES

[1] ANDERSON, D. H. and BROWN, R., Combinatorial aspects of C. A. R. Hoare's FIND algorithm, *Australas. J. Combin.* **5** (1992), 109–119. *MR* **94**e:05005

[2] BARLOW, R. E. and PROSCHAN, F., *Statistical theory of reliability and life testing. Probability models*, Holt, Rinehart and Winston, New York, 1975. *MR* **55** #11534

[3] BLUM, M., PRATT, V., TARJAN, R. E., FLOYD, R. W. and RIVEST, R. L., Time bounds for selection, *J. Comput. System Sci.* **7** (1973), 448–461. *MR* **48** #8256

[4] CAMBANIS, S., SIMONS, G. and STOUT, W., Inequalities for $\mathcal{E}k(X,Y)$ when the marginals are fixed, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **36** (1976), 285–294. *MR* **54** #8790

[5] COLE, R. and YAP, CHEE K., A parallel median algorithm, *Inform. Process. Lett.* **20** (1985), 137–139. *MR* **86k**:68004

[6] CUNTO, W. and MUNRO, J. I., Average case selection, *J. Assoc. Comput. Mach.* **36** (1989), 270–279. *MR* **91f**:68024

[7] HOARE, C. A., Algorithm 65 FIND, *Comm. ACM* **4** (1961), 321.

[8] KNUTH, D. E., Mathematical analysis of algorithms, *Information processing '71* (Proc. IFIP Congress, Ljubljana, 1971), Vol. 1: Foundations and systems, North-Holland, Amsterdam, 1972, 19–27. *MR* **53** #7122

[9] KRASNOSEL'SKIĬ, M. A. and RUTICKIĬ, YA. B., *Convex functions and Orlicz spaces*, Noordhoff, Leiden, 1961. *MR* **23** #A4016

[10] RACHEV, S. T., *Probability metrics and the stability of stochastic models*, Wiley, New York, 1991. *MR* **93b**:60012

[11] RIVEST, R. and FLOYD, R. W., Bounds for the expected time for median computations (extended abstract), *Combinatorical algorithms (Courant Comput. Sci. Sympos.* **9**, New York Univ., New York, 1972), Algorithmics Press, New York, 1973, 69–76. *MR* **50** #11849

[12] RIVEST, R. and FLOYD, R. W., Expected time bounds for selections, *Comm. ACM* **18** (1975), 165–172. *Zbl* **296**.68049

[13] SCHÖNHAGE, A., PATERSON, M. and PIPPENGER, N., Finding the median, *J. Comput. System Sci.* **13** (1976), 184–199. *MR* **55** #1814

[14] SINGLETON, R. C., Algorithm 347: An efficent algorithm for sorting with minimal storage, *Commun. ACM* **12** (1969), 185–186. (See *MR* **94**e:05005.)

[15] YAP, CHEE K., New upper bounds for selection, *Comm. ACM* **19** (1976), 501–508. *MR* **54** #6559

(Received June 23, 1995)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS STATISZTIKAI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

moritamas@ludens.elte.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 209–238

THE CENTRAL LIMIT THEOREM FOR L-STATISTICS

R. NORVAIŠA

Dedicated to E. Csáki for his sixtieth birthday

Summary

We consider a linear combination $L_n = n^{-1} \sum_{1 \leq i \leq n} c_{ni} h(X_{n:i})$ of a function of order statistics $X_{n:1} \leqq \ldots \leqq X_{n:n}$ corresponding to a sample of independent random variables with a common distribution function. Two improved variants of known sufficient conditions for the central limit theorem for L_n to hold are given. The first one concerns the case when the weight constants $c_{n1}, \ldots, c_{nn}$ are given by $c_{ni} = n \int_{(i-1)/n}^{i/n} J\, d\lambda$, where J is a Lebesgue integrable function. The second one allows the weight constants to be arbitrary. To obtain these results, we invoke differentiability of superposition (or Nemytskiĭ) operators induced by an integral representation of L_n and a central limit theorem for the empirical process with sample paths in a Banach function space.

1. Introduction and results

Let $X_1, \ldots, X_n$ be independent random variables with a common distribution function (df) F, and let $X_{n:1} \leqq \cdots \leqq X_{n:n}$ be their corresponding order statistics. Consider a linear combination of a function of these order statistics, or an L-statistic, of the following form:

$$L_n := \frac{1}{n} \sum_{i=1}^{n} c_{ni} h(X_{n:i}),$$

where the function h is assumed to be an indefinite integral with the corresponding Lebesgue–Stieltjes signed measure dh and the total variation measure μ_h.

The first main result is an asymptotic normality of the L-statistic L_n when the weight constants $c_{n1}, \ldots, c_{nn}$ are given by a function J, i.e. when

$$(1.1) \qquad c_{ni} = n \int_{(i-1)/n}^{i/n} J\, d\lambda, \qquad i = 1, \ldots, n.$$

1991 *Mathematics Subject Classification.* Primary 62G30; Secondary 60B12, 47H30.
Key words and phrases. L-statistic, empirical process, central limit theorem, superposition operator, differentiability, Banach function space.

The function J defined on $[0,1]$ is assumed to be integrable with respect to Lebesgue measure λ and is often called a score function. Recall that a point s is called a *Lebesgue point* of J if

$$\lim_{x\to 0}\frac{1}{x}\int_s^{s+x}|J(t)-J(s)|\,dt=0.$$

THEOREM 1.1. *Let* $0<p,q<\infty$. *Consider a df* F *and an indefinite integral* h *such that*

$$\int F^{p/2}(1-F)^{q/2}d\mu_h<\infty. \tag{1.2}$$

Let a score function J *be such that*

$$\mu_h(\{t: F(t) \text{ is not a Lebesgue point of } J\})=0 \tag{1.3}$$

and assume that there exists a finite constant C *such that*

$$|J(u)|\leqq Cu^{p/2-1/2}(1-u)^{q/2-1/2}, \quad \text{for } \lambda\text{-a.a. } u\in[0,1]. \tag{1.4}$$

Then the L-statistic L_n *given by the weight constants* (1.1) *satisfies the central limit theorem, i.e.,*

$$\sqrt{n}(L_n-L(J,F)) \xrightarrow{d} N(0,\sigma^2(J,F)), \quad as \quad n\to\infty,$$

where

$$L(J,F)=\int Jh\circ F^{-1}d\lambda \tag{1.5}$$

and

$$\sigma^2(J,F)=\iint[F(t\wedge s)-F(t)F(s)]J(F(t))J(F(s))\,h(dt)h(ds). \tag{1.6}$$

Example 3 in Shorack [28] and example 5.6 in Stigler [33] show that the central limit theorem ceases to hold if a score function J and a "probabilistic" inverse function F^{-1} have common discontinuities. To prevent such pathologies it is customary to require for J to be continuous at $F(t)$ for μ_h-a.a. t, even when the less stringent condition (1.3) is sufficient (see e.g. the proof of Theorem 1 in Boos [2]). Also, the discontinuity points of J in the above mentioned examples are not Lebesgue points. Using more general integrals in (1.1) than that of Lebesgue, one may hope to weaken condition (1.3) (see Remark 3.3 below). Mason and Shorack [18] proved that the sufficient condition (1.2) can be replaced by the necessary condition $\sigma^2(F,J)<\infty$, whenever the L-statistic L_n is slightly trimmed and a score function J behaves regularly. We do not know whether (1.2) may also be weakened in the present context.

The second main result extends the first one to arbitrary weight constants and incorporates a more general integral condition on a score function than that of (1.4). It costs us the restriction of the parameter p (and q) to the range $[1, \infty)$ and an additional assumption on the image measure $F(\mu_h)$ when $p > 1$. To be more precise we need some more notation. As usual, for a given measure space (T, μ) denote by $\mathbb{L}_r(T, \mu) = \mathbb{L}_r(\mu)$ a Lebesgue space of μ-measurable functions f on T for which the norm

$$\|f\|_{L_r(\mu)} := \begin{cases} (\int |f|^r d\mu)^{1/r}, & \text{if } 1 \leqq r < \infty, \\ \text{ess sup}|f|, & \text{if } r = \infty, \end{cases}$$

is finite. Let J be a Lebesgue integrable function on an interval $I \subset \mathbb{R}$. The *Hardy–Littlewood maximal function* MJ of J is defined by

$$(MJ)(x) := \sup_{x_1 < x < x_2} \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} |J| \, d\lambda, \quad x \in I. \tag{1.7}$$

The operator $M : J \to MJ$ is called the *Hardy–Littlewood maximal operator.* Let $1 < r < \infty$. The condition on the measure $F(\mu_h)$ we are going to assume is that there exist another measure ν on $[0, 1]$ and a finite constant C_r such that

$$\int |MJ|^r dF(\mu_h) \leqq C_r \int |J|^r d\nu. \tag{1.8}$$

If $F(\mu_h)$ is a Lebesgue measure on $[0, 1]$, then one may take for ν the same Lebesgue measure and (1.8) becomes a classical Hardy–Littlewood maximal theorem (see e.g. Stein [30]). A more general result is used in Example 3.9 below. If $r = \infty$, then a sup variant of (1.8) always holds for $\nu = F(\mu_h)$. For given arbitrary weight constants $c_{n1}, \ldots, c_{nn}$, define a sequence of functions $\{J_n;\, n \geqq 1\}$ on $[0, 1]$ by

$$J_n(u) = c_{ni}, \quad \text{if} \quad u \in ((i-1)/n, i/n] \quad \text{and} \quad i = 1, \ldots, n, \tag{1.9}$$

and $J_n(0) = c_{n1}$.

Now we can state our second main result:

THEOREM 1.2. *Let $1 \leqq p < \infty$. Consider a df F and an indefinite integral h such that*

$$\int [F(1-F)]^{p/2} d\mu_h < \infty. \tag{1.10}$$

Assume that there exists a measure ν on $[0, 1]$ such that (1.8) holds for $r = p' := p/(p-1)$. Suppose also that $\{J_n;\, n \geqq 1\} \subset \mathbb{L}_{p'}(\nu)$ and that there exists a Lebesgue integrable function $J \in \mathbb{L}_{p'}(\nu)$ such that (1.3) holds and

$$\lim_{n \to \infty} \|J_n - J\|_{L_{p'}(\nu)} = 0.$$

Then the L-statistic L_n satisfies the central limit theorem, i.e.,

$$\sqrt{n}(L_n - L(J_n, F)) \xrightarrow{d} N(0, \sigma^2(J, F)), \quad as \quad n \to \infty,$$

where $L(J,F)$ and $\sigma^2(J,F)$ are given by (1.5) *and* (1.6), *respectively.*

The proofs of both theorems are carried out through sections 3, 4 and 5 where we deal with differentiability of superposition operators, the central limit theorem for empirical processes and L-statistics, respectively. Section 2 contains some notations and results related to a theory of function spaces equipped with a structure of Banach lattice (or Riesz space).

Now we briefly comment on known results and relations of our paper with some of them. One can find a more complete survey and historical comments in Stigler [31], [32], Shorack [27] and Serfling [25]. Most general results on the asymptotic normality of L-statistics have been given by Shorack [28], Stigler [33], Boos [2], Serfling [25] and Mason [17]. The conditions of Theorem 1 in Shorack [28] as it applies to L_n requires the existence of a finite constant C such that

$$|h(x)| \leqq Cx^{-p/2+\epsilon}(1-x)^{-q/2+\epsilon},$$

for some $\epsilon > 0$, in addition to assumptions of Theorem 1.1 above. The main ingredient in his proofs was a "special construction" of an empirical process and a Brownian bridge (see for example Shorack and Wellner [29] page 93). The method of projection was used in Stigler [33] for L-statistics L_n with the weight constants given by

$$c_{ni} = J(i/(n+1)), \quad i = 1, \dots, n.$$

His results have been extended by Mason [17]. Theorem 1 in Mason [17] asserts the asymptotic normality of slightly trimmed L-statistics L_n centered by EL_n in fact under the same conditions as in Theorem 1.1 above. Centering by $L(J,F)$ then requires Hölder type smoothness of a score function J (see Theorem 2 in Mason [17] and Theorem 4 in Stigler [33]). The differentiable statistical function approach has been used to L-statistics by Reeds [22], Boos [2], Serfling [25]. For example, if a score function J is continuous then our Theorem 1.1 in the case of $p=q=1$ coincides with Theorem C on p. 284 in Serfling [25]. Later on a theory of differentiable statistical functionals received considerable attention (see e.g. Fernholz [10], Esty et al. [8], Shao [26]). Unfortunately, these results as applied to L-statistics require trimming and/or appropriate smoothness of a score function.

In this paper we follow an approach proposed by Shorack [28] for the first step. Namely, by performing integration by parts one may represent the L-statistic L_n as (or approximately in probability by) a composition of a nonlinear superposition operator (induced by a df F and a score function J) acting from some Banach function space into $\mathbb{L}_1(\mu_h)$ and a linear functional $f \to \int f\, dh$ acting from $\mathbb{L}_1(\mu_h)$ into $\mathbb{R}$ and use a general idea of the above mentioned theory of differentiable statistical functionals. New ingredients we invoke in this way are a technique developed in the theory of integral equations for handling superposition operators and a central limit theorem in an arbitrary Banach space. In particular, for a class of weakly compact sets $\mathcal{C}$, we prove $\mathcal{C}$-differentiability of a composition operator (a special case of a superposition operator) previously known to be Hadamard differentiable, i.e. $\mathcal{C}$-differentiable with $\mathcal{C}$ being a class of norm compact sets. This and other statements from Section 3 are closely related to corresponding results due to Reeds [22], Fernholz [10] and Dudley [5]. It is worthwhile to mention also that continuity results of superposition operators and laws of large numbers in Banach function spaces for the empirical df have been used in an above mentiond way in Norvaiša [21] to obtain laws of large numbers for L-statistics.

2. Banach function spaces

In this section we consider Banach spaces of measurable functions compatible with an order structure. This class of spaces includes classical Lebesgue spaces $\mathbb{L}_p$, $1 \leqq p \leqq \infty$, Orlicz spaces and their generalizations (such as Musielak–Orlicz spaces), Lorentz, Marcinkiewicz and symmetric spaces.

Let $(T, \mu) = (T, \mathcal{T}, \mu)$ be a complete σ-finite measure space with a σ-algebra $\mathcal{T}$ of subsets of T. Denote by $\mathbb{M} = \mathbb{M}(T, \mu)$ the linear space of all equivalence classes of μ-measurable real-valued functions defined and finite μ-a.e. on T. A map $\|\cdot\|: \mathbb{M} \to [0, \infty]$ is called a *function norm* if

(1) $\|\cdot\|$ is a norm;
(2) $|f| \leqq |g|$ (μ-a.e.) implies $\|f\| \leqq \|g\|$;
(3) if $E \subset T$ is of finite μ-measure and χ_E is its indicator function then $\|\chi_E\| < \infty$.

Given a function norm on $\mathbb{M}$, define the set

$$\mathbb{B}(T, \mu) := \{ f \in \mathbb{M}(T, \mu) : \|f\| < \infty \}.$$

Then $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$ is a normed linear space. If $\mathbb{B}$ is complete, it is called a *Banach function space* (B.f.s.). We will also assume further that B.f.s.'s are order complete (or Dedekind complete). We refer to Zaanen [34] for notation not explained here.

Let $\{f_n : n \geqq 1\}$ be a sequence in $\mathbb{M} = \mathbb{M}(T, \mu)$. As usual, f_n converges in $\mathbb{M}$, if it converges in μ-measure on every finite measure subset of T. The convergence to zero μ-a.e. of a sequence $\{f_n : n \geqq 1\}$, say $f_n \to 0$ μ-a.e., means its *convergence in order in* $\mathbb{M}$. Let $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$ be a Banach function space. A sequence $\{f_n : n \geqq 1\}$ *converges in order in* $\mathbb{B}$ to zero if $f_n \to 0$ μ-a.e. and there exists $f \in \mathbb{B}$ such that $|f_n| \leqq |f|$ for all $n \geqq 1$. A subset $A \subset \mathbb{B}$ of a B.f.s. $\mathbb{B}$ is said to be of *uniformly absolutely continuous norm* (u.a.c. norm) whenever, given $\epsilon > 0$ and a sequence $\{E_n : n \geqq 1\}$ of μ-measurable sets with $E_n \downarrow \emptyset$, there exists an index N such that $\|f \chi_{E_n}\| < \epsilon$ holds for all $n \geqq N$ and all $f \in A$ simultaneously. An element $f \in \mathbb{B}$ is said to have an *absolutely* *continuous norm* if the set $\{f\}$ is of u.a.c. norm. A B.f.s. $\mathbb{B}$ is said to be *order continuous*, (or to have an absolutely continuous norm) whenever every element of $\mathbb{B}$ has an absolutely continuous norm. It is worthwile to recall that, by Theorem 1.3.7 in Luxemburg [15], a B.f.s. $\mathbb{B}=(\mathbb{B}(T,\mu),\|\cdot\|)$ is separable if and only if $\mathbb{B}$ is order continuous, and the measure μ is separable. Note that $\mathbb{L}_p$, $1\leqq p<\infty$, are order continuous B.f.s.'s and that an Orlicz space $\mathbb{L}_\phi$ is order continuous if and only if ϕ satisfies Δ_2-condition at infinity.

The following lemma is essentially due to Luxemburg [15].

LEMMA 2.1. *Let $\{f_n\colon n\geqq 1\}$ be a sequence of elements of a B.f.s. $\mathbb{B}=(\mathbb{B}(T,\mu),\ \|\cdot\|)$. If $f_n\to 0$ in μ-measure and the set $\{f_n\colon n\geqq 1\}$ is of u.a.c. norm, then $\|f_n\|\to 0$. The converse implication holds true if, in addition, $\mathbb{B}$ is order continuous.*

REMARK. For the B.f.s. L_p, $1\leqq p<\infty$ this statement is known also as Vitali's theorem (see Theorem III.3.6 in Dunford and Schwartz [6]). One can find in van Eldik and Grobler [7] (see Theorem 2.5) even a more general theorem than stated here.

PROOF. Due to Lemma 1.2.2 in Luxemburg [15], one only needs to notice that if $A\subset\mathbb{B}$ has u.a.c. norm and if E is a set of finite μ-measure, then $\|f\chi_{E_n}\|\to 0$ uniformly over A for every sequence of subsets $\{E_n\colon n\geqq 1\}$ of E such that $\mu(E_n)\to 0$.

A subset $A\subset\mathbb{B}$ of a B.f.s. $\mathbb{B}$ is said to be *L-weakly compact* if it is norm bounded and if $\|f_n\|\to 0$ for every disjoint sequence $\{f_n\colon n\geqq 1\}$ in the positive part of the solid hull of A. By Satz II.2. in Meyer-Nieberg [19], a norm bounded subset A is L-weakly compact if and only if given $\epsilon>0$ there exists a positive element $g\in\mathbb{B}$ with order continuous norm such that $A\subset[-g,g]+V(\epsilon)$, with $V(\epsilon)=\{f\in\mathbb{B}:\|f\|<\epsilon\}$. The following statement has been proved in van Eldik and Grobler [7] (see Proposition 2.8).

LEMMA 2.2. *Let $\mathbb{B}$ be an order complete B.f.s. A norm bounded subset $A\subset\mathbb{B}$ is L-weakly compact iff A is of u.a.c. norm.*

One can find in Dodds and Fremlin [4] additional characterization results of L-weakly compact sets. For example, by their Theorem 4.2, relative (norm) compactness implies L-weak compactnes whenever $\mathbb{B}$ is order continuous.

3. Superposition operators

Let (T,μ) be a complete σ-finite measure space, and let $\phi=\phi(t,x)$ be a real-valued function defined on $T\times\mathbb{R}$. Given a function $f=f(t)$ on T, one can associate another function

$$\Phi f(t) := \phi(t,f(t)) \tag{3.1}$$

that is also defined on T. In this way the function ϕ induces a map Φ called *superposition operator* (or Nemytskiĭ operator). A function ϕ is said to be *sup-measurable* if the operator Φ maps every measurable function into measurable function, i.e., the superposition $\phi(t, f(t))$ is measurable for every measurable function f. This property allows to consider Φ as a nonlinear operator acting between B.f.s.'s induced by function norms on $\mathbb{M}(T,\mu)$. We refer to Appell and Zabrejko [1] for more information on the subject. Here we are interested in differentiability properties of superposition operators defined by (3.1).

Let $(\mathbb{B}_i, \|\cdot\|_i)$, $i=1,2$, be two normed spaces, and let $\mathcal{C} = \mathcal{C}(\mathbb{B}_1)$ be a collection of bounded subsets of $\mathbb{B}_1$, containing all singletons $\{f\}$, $f \in \mathbb{B}_1$. An operator Φ between $\mathbb{B}_1$ and $\mathbb{B}_2$ is said to be *$\mathcal{C}$-differentiable at* $f_0 \in \mathbb{B}_1$ if there is a bounded linear operator $\Phi'(f_0)$ from $\mathbb{B}_1$ into $\mathbb{B}_2$ such that

$$\Phi(f_0+f) - \Phi(f_0) = \Phi'(f_0)\,f + \Delta(f)$$

and for every $C \in \mathcal{C}$

$$\lim_{x\to 0} \frac{1}{x} \|\Delta(xf)\|_2 = 0$$

uniformly for $f \in C$. The linear operator $\Phi'(f_0)$ is called the derivative of Φ at f_0. $\mathcal{C}$-differentiability was defined by Sebastião e Silva [24]. In statistics usually there are three particular types of differentiation that are of interest:

(1) $\mathcal{C}(\mathbb{B}_1) = \{$bounded subsets of $\mathbb{B}_1\}$; this corresponds to Fréchet differentiation.

(2) $\mathcal{C}(\mathbb{B}_1) = \{$(norm) compact subsets of $\mathbb{B}_1\}$; this corresponds to Hadamard (or compact) differentiation.

(3) $\mathcal{C}(\mathbb{B}_1) = \{$single point subsets of $\mathbb{B}_1\}$; this corresponds to Gâteaux differentiation.

Other collections $\mathcal{C}$ of subsets of $\mathbb{B}_1$ have been considered recently by Dudley [5]. It is shown below that Hadamard differentiability of a superposition operator Φ may be extended in some cases to $\mathcal{C}_L(\mathbb{B}_1)$-differentiability, where

$$\mathcal{C}_L(\mathbb{B}_1) = \{\text{L-weakly compact subsets of } \mathbb{B}_1\}. \tag{3.2}$$

Now we pass to differentiability conditions for the superposition operator Φ induced by a sup-measurable function ϕ and acting between two B.f.s.'s. The arguments from the paragraphs just before equation (2.55) in Appell and Zabrejko [1] assures that the Gâteaux derivative $\Phi'(f_0)$ of Φ has necessarily the form

$$\Phi'(f_0)\,f = a^{f_0} f,$$

i.e., $\Phi'(f_0)$ is always a multiplication operator by some measurable function. Moreover, Gâteaux differentiability yields that

$$a^{f_0}(t) = \mu\text{-}\lim_{x\to 0} \frac{1}{x}[\phi(t, f_0(t)+x) - \phi(t, f_0(t))], \quad t \in T, \tag{3.3}$$

where μ- lim denotes the convergence in μ-measure on each set of finite measure.

For a given B.f.s. $\mathbb{B} = (\mathbb{B}(T,\mu), \|\cdot\|)$ and for a non-negative function $w \in \mathbb{M}(T,\mu)$, define a weighted B.f.s. $\mathbb{B}_w = (\mathbb{B}_w(T,\mu), \|\cdot\|_w)$ by

$$\mathbb{B}_w(T,\mu) := \{f \in \mathbb{M}(T,\mu) : \|fw\| < +\infty, \ \forall f \in \mathbb{B}\},$$

with the norm $\|\cdot\|_w := \|\cdot w\|$. Now our first statement reads as follows:

PROPOSITION 3.1. *Let $\mathbb{B} = (\mathbb{B}(T,\mu), \|\cdot\|)$ be an order continuous B.f.s., and let ϕ be a sup-measurable function. For a given function $f_0 \in \mathbb{M}(T,\mu)$ such that $\Phi(f_0) \in \mathbb{B}$, assume that there exists a function $a^{f_0} \in \mathbb{M}(T,\mu)$ such that (3.3) holds. Suppose also that there exist a non-negative function $w \in \mathbb{M}(T,\mu)$ and finite constants C_1, C_2 such that*

$$|a^{f_0}| \leqq C_1 w, \tag{3.4}$$

$$|\Phi(f_0 + x) - \Phi(f_0)| \leqq C_2|x|w, \quad \forall x \in \mathbb{R}, \tag{3.5}$$

and $f_0 w \in \mathbb{B}$. Then the superposition operator Φ maps the weighted B.f.s. $\mathbb{B}_w$ into $\mathbb{B}$ and Φ is $C_L(\mathbb{B}_w)$-differentiable at f_0 with the derivative

$$\Phi'(f_0)\,f = a^{f_0} f,$$

where the class $C_L(\mathbb{B})$ is defined by (3.2).

PROOF. By (3.5) it follows that

$$\|\Phi(f)\| \leqq C_2\|(f - f_0)w\| + \|\Phi(f_0)\| < \infty,$$

for all $f \in \mathbb{B}_w$. Hence the superposition operator Φ acts between $\mathbb{B}_w$ and $\mathbb{B}$. By (3.4), the derivative $\Phi'(f_0)$ is a linear bounded operator from $\mathbb{B}_w$ into $\mathbb{B}$. The remainder in the claimed differentiation is

$$\Delta(f) = \Phi(f_0 + f) - \Phi(f_0) - a^{f_0} f.$$

All single point subsets of $\mathbb{B}_w$ belong to $C_L(\mathbb{B}_w)$ since $\mathbb{B}$, and hence $\mathbb{B}_w$ too, are order continuous B.f.s.'s. Let C be an L-weakly compact subset of $\mathbb{B}_w$. It is enough to show that

$$\lim_{x_n \to 0} \frac{1}{x_n} \|\Delta(x_n f_n)\| = 0 \tag{3.6}$$

for an arbitrary sequence $\{f_n : n \geqq 1\} \subset C$. By (3.4) and (3.5), it follows that

$$|\Delta(x_n f_n)|/x_n \leqq (C_1 + C_2)|f_n|w$$

for all $n \geq 1$. This in conjunction with Lemma 2.2, yields that the sequence $\{\Delta(x_n f_n)/x_n \colon n \geq 1\}$ in $\mathbb{B}$ has u.a.c. norm. Since every ball in a B.f.s. is a bounded subset in $\mathbb{M}$, it can be shown that (3.3) yields

$$\mu\text{-}\lim_{x_n \to 0} \frac{1}{x_n} |\Delta(x_n f_n)| = 0.$$

Now, the desired relation (3.6) follows from Lemma 2.1 and the proof is complete.

Now consider a special case of a superposition operator corresponding to a composition of two functions, i.e., the case when the function $\phi(t,\cdot) \equiv \phi(\cdot)$ for all $t \in T$, and

$$\Phi f(t) = \phi \circ f(t), \quad \forall t \in T. \tag{3.7}$$

A function ϕ from $\mathbb{R}$ into $\mathbb{R}$ is called *Lipschitz* if for some $K < \infty$, $|\phi(x) - \phi(y)| \leq K|x - y|$ for all x, y. Then for Lebesgue almost all x, the derivative $\phi'(x)$ exists, with $|\phi'(x)| \leq K$. The following statement improves and extends Proposition 6.1.2. in Fernholz [10], where ϕ was assumed to be continuous and piecewise differentiable with bounded derivative, and C was taken to be a class of (norm) compact sets in the B.f.s. $\mathbb{B} = \mathbb{L}_p[0,1]$, $1 \leq p < \infty$.

COROLLARY 3.2. *Let $\mathbb{B} = (\mathbb{B}(T,\mu), \|\cdot\|)$ be an order continuous B.f.s., and let ϕ be a Borel measurable Lipschitz function from $\mathbb{R}$ into $\mathbb{R}$. For a given function $f_0 \in \mathbb{B}$ such that $\phi \circ f_0 \in \mathbb{B}$ assume that*

$$\mu(\{\, t \in T : \phi \text{ is not differentiable at } f_0(t)\}) = 0. \tag{3.8}$$

Then the superposition (composition) operator Φ given by (3.7) maps $\mathbb{B}$ into $\mathbb{B}$ and Φ is $C_L(\mathbb{B})$-differentiable at f_0 with the derivative

$$\Phi'(f_0)f = (\phi' \circ f_0)f.$$

PROOF. Take the function w in Proposition 3.1 to be equal to the Lipschitz constant K.

REMARK 3.3. Lipschitz condition and (3.8) imply that (3.3) holds μ-a.e. Thus one may hope to weaken condition (3.8) if ϕ would not be require to be a Lipschitz function. For this one may invoke a variant of a generalized derivative considered among others by Khintchine [13]. Namely, *la dérivée générale* of ϕ is a function ϕ'_g, defined a.e., such that for any $\epsilon > 0$

$$\lim_{h \to 0} \lambda(\{\, x : |\frac{\phi(x+h) - \phi(x)}{h} - \phi'_g(x)| > \epsilon\,\}) = 0.$$

The existence of an approximate derivative ϕ'_{ap} a.e. on an interval implies the existence of ϕ'_g on the interval, and $\phi'_{ap} = \phi'_g$ a.e., while Khintchine constructed a function ϕ such that ϕ'_g exists on $[0,1]$, and ϕ'_{ap} only on a null set. In connection with L-statistics one may consider a Denjoy–Khintchine integrable score function J. Recall that a function ϕ on $[a,b]$ is the Denjoy–Khintchine integral of a function J provided ϕ is ACG on $[a,b]$, and $\phi'_{ap} = J$ a.e. (see Saks [23] for details).

Let Θ be a normed space of parameters θ. We wish to consider a linear family of sup-measurable functions $\phi = \phi(\theta; t, x) = \phi_\theta(t, x)$ on $\Theta \times T \times \mathbb{R}$ such that

(1) $\phi_\theta(\cdot,\cdot)$ is a sup-measurable function on $T \times \mathbb{R}$ for all $\theta \in \Theta$;

(2) $\phi(\cdot; t, x)$ is a linear function on Θ for all $(t,x) \in T \times \mathbb{R}$.

Then one can define an operator Φ from $\Theta \times \mathbb{M}(T,\mu)$ into $\mathbb{M}(T,\mu)$ by

$$\Phi(\theta, f)(t) := \phi(\theta; t, f(t)), \quad \forall t \in T. \tag{3.9}$$

Note that $\Phi(\cdot, f)$ is a linear operator on Θ whenever a function f is fixed. A simple example of a linear family of sup-measurable functions may be given by any normed space Θ of sup-measurable functions $\psi = \psi(t,x)$ on $T \times \mathbb{R}$ or Borel measurable functions $\psi = \psi(x)$ on $\mathbb{R}$; in these cases $\phi(\psi; t, x) := \psi(t,x)$ or $\phi(\psi; x) := \psi(x)$, respectively. The last example induces the composition operator $\Phi: (\psi, f) \to \psi \circ f$. Still another example will be considered below (see (3.19)). Define also the maximal operator $M\Phi$ on $\Theta \times \mathbb{M}$ by

$$M\Phi(\theta, f)(t) := \sup_{x \neq 0} \left| \frac{1}{x} [\phi_\theta(t, f(t) + x) - \phi_\theta(t, f(t))] \right|,$$

where sup means a lattice supremum in $\mathbb{M}$. Let $\mathbb{B}_i = (\mathbb{B}_i(T,\mu); \|\cdot\|_i)$, $i = 1, 2$, be a pair of B.f.s.'s. The *generalized dual space* (or multiplicator space) $\mathbb{B}_2/\mathbb{B}_1$ is defined to be the set

$$(\mathbb{B}_2/\mathbb{B}_1)(T,\mu) := \{ f \in \mathbb{M}(T,\mu) : fg \in \mathbb{B}_2 \text{ for each } g \in \mathbb{B}_1 \}.$$

Equipped with the natural norm

$$\|f\|_{2/1} := \sup\{ \|fg\|_2 : \|g\|_1 \leqq 1 \quad g \in \mathbb{B}_1 \},$$

the set $\mathbb{B}_2/\mathbb{B}_1$ becomes a B.f.s. $((\mathbb{B}_2/\mathbb{B}_1)(T,\mu), \|\cdot\|_{2/1})$ (see Maligranda and Person [16], or p. 62 in Appell and Zabrejko [1]).

Now we are ready to state and prove the following differentiability result for the operator Φ given by (3.9).

PROPOSITION 3.4. *Let $\mathbb{B}_i = (\mathbb{B}_i(T,\mu), \|\cdot\|_i)$, $i = 1, 2$, be a pair of B.f.s.'s with a generalized dual $\mathbb{B}_2/\mathbb{B}_1 \neq \{0\}$, and let $(\Theta, \|\cdot\|)$ be a normed space. Suppose a linear family of sup-measurable functions $\phi = \phi(\theta; t, x)$ on $\Theta \times T \times \mathbb{R}$ and elements $f_0 \in \mathbb{B}_1$, $\theta_0 \in \Theta$ are such that for some finite constant C the following hold:*

$$\|M\Phi(\theta, f_0)\|_{2/1} \leqq C\|\theta\|, \tag{3.10}$$

$$\Phi(\theta, f_0) \in \mathbb{B}_2, \tag{3.11}$$

for all $\theta \in \Theta$. Assume also that there exists a μ-measurable function $a = a(\theta_0, f_0)(\cdot)$ on T such that

$$a(t) = \lim_{x \to 0} \frac{1}{x} [\phi_{\theta_0}(t, f_0(t) + x) - \phi_{\theta_0}(t, f_0(t))] \tag{3.12}$$

in μ-measure. Then the operator Φ defined by (3.9) *maps a product space $(\Theta \times \mathbb{B}_1, \|\cdot\| + \|\cdot\|_1)$ into $\mathbb{B}_2$ and Φ is C-differentiable at (θ_0, f_0) with the derivative*

$$\Phi'(\theta_0, f_0)(\theta, f) = \Phi(\theta, f_0) + a(\theta_0, f_0) f, \tag{3.13}$$

where C is a collection of all sets $B \times K$ such that B is a ball in Θ and K is an L-weakly compact set in $\mathbb{B}_1$.

PROOF. For any real-valued function f on T, and for any element θ, just by the definition of the maximal operator $M\Phi$ we have

$$|\Phi(\theta, f)(t)| \leqq |f(t) - f_0(t)|\, M\Phi(\theta, f_0)(t) + |\Phi(\theta, f_0)(t)|,$$

for all $t \in T$. Due to the definition of the generalized dual space $\mathbb{B}_2/\mathbb{B}_1$ and by the assumptions (3.10), (3.11), it follows that

$$\|\Phi(\theta, f)\|_2 \leqq C\|f - f_0\|_1 \|\theta\| + \|\Phi(\theta, f_0)\|_2 < \infty,$$

for all $(\theta, f) \in \Theta \times \mathbb{B}_1$. Therefore the operator Φ maps $\Theta \times \mathbb{B}_1$ into $\mathbb{B}_2$. Moreover, the linear operator $\Phi'(\theta_0, f_0)$ is bounded. Let K be an L-weakly compact set in $\mathbb{B}_1$, and let B be a ball in Θ. It is sufficient to show for arbitrary sequences $\{f_n \colon n \geqq 1\} \subset K$ and $\{\theta_n \colon n \geqq 1\} \subset B$ that

$$\lim_{x_n \to 0} \frac{1}{x_n} \|\Delta(x_n \theta_n, x_n f_n)\|_2 = 0, \tag{3.14}$$

where the remainder in the claimed differentiation is

$$\begin{aligned} \Delta(\theta, f) &= \Phi(\theta_0 + \theta, f_0 + f) - \Phi(\theta_0, f_0) - \Phi(\theta, f_0) - af \\ &= [\Phi(\theta_0, f_0 + f) - \Phi(\theta_0, f_0) - af] + [\Phi(\theta, f_0 + f) - \Phi(\theta, f_0)] \\ &=: \Delta_1(\theta_0, f) + \Delta_2(\theta, f). \end{aligned} \tag{3.15}$$

To estimate the first term $\Delta_1(\theta_0, f)$, we will use Lemma 2.1 for the sequence

$$\{\Delta_1(x_n f_n, \theta_0)/x_n;\ n \geqq 1\} \subset \mathbb{B}_2. \tag{3.16}$$

Let E be a μ-measurable set. By the definition of the generalized dual space $\mathbb{B}_2/\mathbb{B}_1$ we have

$$\frac{1}{x_n} \|\chi_E \Delta_1(\theta_0, x_n f_n)\|_2 \leqq \|M\Phi(\theta_0, f_0) + a\|_{2/1} \|\chi_E f_n\|_1,$$

for all $n \geqq 1$. Then, for any sequence of μ-measurable sets E_k with $E_k \downarrow \emptyset$, by Lemma 2.2 it follows that

$$\lim_{k\to\infty} \sup_n \|\chi_{E_k} \Delta_1(\theta_0, x_n f_n)/x_n\|_2$$

$$\leqq \|M\Phi(\theta_0, f_0) + a\|_{2/1} \lim_{k\to\infty} \sup_{f\in K} \|\chi_{E_k} f\|_1 = 0.$$

Therefore, the sequence (3.16) is of u.a.c. norm. Since the sequence $\{f_n; n \geqq 1\}$ is norm bounded in $\mathbb{B}_1$, it is bounded in $\mathbb{M}$, too. This fact, in conjunction with the assumption (3.12), yields that the sequence (3.16) tends to zero in μ-measure. Thus, by Lemma 2.1, it follows that

$$\lim_{x_n\to 0} \frac{1}{x_n} \|\Delta_1(\theta_0, x_n f_n)\|_2 = 0. \tag{3.17}$$

As to the second term in (3.15), due to the definition of the generalized dual space $\mathbb{B}_2/\mathbb{B}_1$ and by the assumption (3.10) we have

$$\frac{1}{x_n} \|\Delta_2(x_n\theta_n, x_n f_n)\|_2 \leqq C x_n \|\theta_n\| \, \|f_n\|_1,$$

for all $n \geqq 1$. This inequality combined with (3.17) yields (3.14) and the proof of Proposition 3.4 is now complete.

In general, $\mathcal{C}$-differentiability of the operator Φ cannot be extended to Fréchet differentiability, even for a fixed parameter $\theta \in \Theta$, without additional assumptions. By Theorem 2.15 in Appell and Zabrejko [1], there is a class of pairs of B.f.s.'s $\mathbb{B}_1$ and $\mathbb{B}_2$ such that the generalized dual are not order continuous and Fréchet differentiability of a superposition operator Φ given by (3.1) yields that the function ϕ is equivalent to an affine function in x. This class includes a pair $\mathbb{B}_1 = \mathbb{B}_2 = \mathbb{L}_p(T, \mu)$, for any $1 \leqq p < \infty$ and for the atomic free measure μ on T.

A stronger assumption than sup-measurability, often made on the superposition operator given by (3.1), is that ϕ is a *Carathéodory function*, that is to say $\phi(\cdot, x)$ is μ-measurable for each $x \in \mathbb{R}$ and $\phi(t, \cdot)$ is continuous for μ-a.a. $t \in T$. The following lemma is a crucial ingredient in the proof of Fréchet differentiability of the operator Φ.

LEMMA 3.5. *Let $\mathbb{B}_i$, $i = 1, 2$, be a pair of B.f.s.'s, and let ψ be a Carathéodory function on $T \times \mathbb{R}$ such that the induced superposition operator Ψ maps $\mathbb{B}_1$ into $\mathbb{B}_2$. Then the operator Ψ is continuous whenever the B.f.s. $\mathbb{B}_2$ is order continuous.*

PROOF. It is a part of Theorem 2.6 proved in Appell and Zabrejko [1].

We are now ready to formulate conditions for the Fréchet differentiability of the operator Φ given by (3.9).

PROPOSITION 3.6. *Under the hypothesis of Proposition 3.4, assume in addition that the generalized dual space $\mathbb{B}_2/\mathbb{B}_1$ is order continuous, and ϕ is a linear family of Carathéodory functions, i.e., $\phi(\theta;\cdot,\cdot)$ is a Carathéodory function for every $\theta \in \Theta$ and $\phi(\cdot;t,x)$ is a linear function on Θ for every $(t,x) \in T \times \mathbb{R}$. Moreover, assume that* (3.12) *holds μ-a.e. Then the operator Φ defined by* (3.9) *maps a product space $(\Theta \times \mathbb{B}_1, \|\cdot\| + \|\cdot\|_1)$ into $\mathbb{B}_2$ and Φ is Fréchet differentiable at (θ, f_0) with the derivative* (3.13).

PROOF. The proof goes along the same lines as that of Proposition 3.4. Here, in addition, we have to establish (3.17) for an arbitrary bounded sequence $\{f_n : n \geqq 1\} \subset \mathbb{B}_1$. To this end, we use Lemma 3.5 for the function ψ defined by

$$\psi(t,x) := \begin{cases} \dfrac{1}{x}[\phi_{\theta_0}(t, f_0(t)+x) - \phi_{\theta_0}(t, f_0(t))], & \text{if } x \neq 0, \\ a(t), & \text{if } x = 0. \end{cases}$$

Since (3.12) holds μ-a.e., ψ is a Carathéodory function, and the induced superposition operator Ψ acts between B.f.s.'s $\mathbb{B}_1$ and $\mathbb{B}_2/\mathbb{B}_1$. Now, by Lemma 3.4, it follows that

$$\limsup_{x_n \to 0} \frac{1}{x_n} \|\Delta_1(\theta_0, x_n f_n)\|_2$$
$$\leqq \sup_n \|f_n\|_1 \limsup_{x_n \to 0} \|\Psi(x_n f_n) - \Psi(0)\|_{2/1} = 0,$$

since $\|x_n f_n\|_1 \to 0$, as $n \to \infty$. This, we noted earlier, is what had to be proved and the proof of Proposition 3.6 is now complete.

Reeds [22], in the proof of Theorem 6.4.3, and Fernholz [10], by Proposition 6.1.6, show that the composition operator

$$\Phi(\phi, f) = \phi \circ f \tag{3.18}$$

is Hadamard differentiable from $D[0,1] \times \mathbb{L}_p([0,1],\lambda)$ into $\mathbb{L}_p([0,1],\lambda)$ at (ϕ_0, f_0) where ϕ_0 is the identity, f_0 is a diffeomorphism and $D[0,1]$ is the space of right-continuous functions with left limits in the supremum norm. Dudley [5] extended this result with respect to the directions of differentiability ϕ from sup (norm) compact sets to much larger sets. Namely, his Theorem 5.1 says in particular that (3.18) is jointly differentiable at (ϕ_0, f_0) from $R_{p/q} \times \mathbb{L}_p([0,1],\lambda)$ into $\mathbb{L}_q([0,1],\lambda)$ whenever $1 \leqq q < p < +\infty$, for Fréchet differentiability in f and $\mathcal{C}$-differentiability in ϕ, where ϕ_0 is Lipschitz function, f_0 is increasing with $f_0' \geqq \beta > 0$ a.e. and $\mathcal{C}$ is the class of uniformly (p/q)-Riemann sets C, i.e., the restrictions of functions in C to any bounded interval are uniformly Riemann and

$$\sup\{\, |\phi(x) - \phi(0)|/(1 + |x|^{p/q}) : x \in \mathbb{R},\ \phi \in C\} < +\infty.$$

To conclude this section, we give an example illustrating the above statements. Let μ be a σ-finite Lebesgue–Stieltjes measure on $T = \mathbb{R}$. Consider a pair of B.f.s.'s $\mathbb{B}_2 = \mathbb{L}_q(\mathbb{R}, \mu)$, $1 \leqq q < +\infty$, and $\mathbb{B}_1 = \mathbb{L}_p(\mathbb{R}, \mu)$, $q \leqq p < +\infty$. Then $\mathbb{B}_2/\mathbb{B}_1 = \mathbb{L}_r(\mathbb{R}, \mu)$ with $r = pq/(p-q)$ is a generalized dual space. Let J be a Lebesgue integrable function on $\mathbb{R}$ from a parameter set Θ to be specified below, and let $f_0 \in \mathbb{L}_p(\mathbb{R}, \mu)$. Define a linear family of Carathéodory functions ϕ by

$$\phi(J; t, x) := \int_{f_0(t)}^{x} J \, d\lambda, \qquad t \in T,\ x \in \mathbb{R}. \tag{3.19}$$

Consider the operator Φ defined by (3.9). Note that $\Phi(J, f_0) = 0$ (cf. (3.11)) and

$$M\Phi(J, f_0)(t) \leqq (MJ)(f_0(t)), \quad \forall t \in T,$$

where MJ is the Hardy–Littlewood maximal function of J defined by (1.7). Let $f_0(\mu)$ denote the image measure on $\mathbb{R}$ given by $f_0(\mu)(A) = \mu(f_0^{-1}(A))$ for all measurable sets A. Thus, it follows by the image measure theorem that

$$\int_{\mathbb{R}} |M\Phi(J, f_0)|^r \, d\mu \leqq \int_{\mathbb{R}} |MJ|^r \, df_0(\mu).$$

To estimate the right side one may use the Hardy–Littlewood maximal theorem whenever $f_0(\mu)$ is a Lebesgue measure. This classical result has been extended to inequality of the type

$$\int_I |MJ|^r v \, d\lambda \leqq C_r \int_I |J|^r v \, d\lambda, \tag{3.20}$$

where $1 < r < \infty$, I is a fixed interval, C_r is a finite constant independent of J and MJ, and v is a non-negative function. Stein [30] showed that (3.20) is true for $I = \mathbb{R}$ and $v(x) = |x|^a$ for $-1/r < a < 1 - 1/r$. Fefferman and Stein [9] showed that (3.20) is true for $I = \mathbb{R}$ if $Mu \leqq Cu$ a.e.. It was Muckenhoupt [20] who gave a characterization of a weight function v for (3.20) to hold. Namely, he proved that (3.20) is true if and only if there is a finite constant K such that

$$\Big[\int_E v \, d\lambda\Big] \Big[\int_E v^{-1/(r-1)} d\lambda\Big]^{r-1} \leqq K \lambda^r(E),$$

where E is any subinterval of I. Moreover, it is said that v *satisfies condition* A_r *on* I whenever the later property holds true. The 70's and 80's witnessed a real flood of papers on weighted inequalities triggered by this result and extensions were obtained to many different directions (see e.g. Garcia-Cuerva and Rubio de Francia [11] for a survey). Due to obvious reasons we do not attempt here to use the most general results.

CONDITION 3.7. Assume that the measure $f_0(\mu)$ on $\mathbb{R}$ is absolutely continuous with respect to the non-negative density function $df_0(\mu)/d\lambda$ satisfying condition A_r on $\mathbb{R}$.

Then, by (3.20), it follows that

$$\int_{\mathbb{R}} |M\Phi(J, f_0)|^r d\mu \leqq C_r \int_{\mathbb{R}} |J|^r df_0(\mu),$$

for some finite constant C_r depending on r only. This gives us condition (3.8) if one takes a parameter set Θ to be $\mathbb{L}_r(\mathbb{R}, f_0(\mu))$.

CONDITION 3.8. Assume $f_0 \in \mathbb{L}_p(\mu)$ and let $J_0 \in \mathbb{L}_r(f_0(\mu))$ to be such that

$$\mu(\{\, x \in \mathbb{R}\colon\ f_0(x) \text{ is not a Lebesgue point of } J_0 \,\}) = 0.$$

Then, by Lebesgue's theorem on derivation of the indefinite integral, (3.12) holds μ-a.e. for $\theta_0 = J_0$ and for $a = J_0 \circ f_0$. Thus , by Propositions 3.4 and 3.6 we have:

EXAMPLE 3.9. *Under the previous notation, assume that conditions 3.7 and 3.8 hold. Then the operator* Φ *defined by* (3.9) *with* ϕ *given by* (3.19) *maps the product space* $\mathbb{L}_r(f_0(\mu)) \times \mathbb{L}_p(\mu)$ *into* $\mathbb{L}_q(\mu)$, *where* $r = pq/(p-q)$ *and* $p \geqq q$. *Moreover,* Φ *is Fréchet (C-)differentiable at* (J_0, f_0) *with the derivative*

$$\Phi'(J_0, f_0)(J, f) \;=\; J_0 \circ f_0 \cdot f,$$

whenever $p > q$ $(p = q$, *respectively).*

4. Empirical processes

To use $\mathcal{C}$-differentiability of a superposition operator defined on a B.f.s. $\mathbb{B}$ for L-statistics we will need paths of the empirical process to be concentrated in probability on sets from the class $\mathcal{C}$. Here we show for a large class of B.f.s.'s $\mathbb{B}$ that this property (say $\mathcal{C}$-tightness) of the empirical process is equivalent to the central limit theorem whenever $\mathcal{C}$ contains all (norm) compact sets of $\mathbb{B}$. A more precise statement follows in Corollary 4.5 below which may be considered as a main result of this section whose proof will be shown to be a consequence of some results from the theory of Probability in Banach spaces.

DEFINITION F. Let F be a non-degenerate df, i.e., assume that F is not a df of a constant rv, and put

$$a := \inf\{t \in \mathbb{R}\colon\ F(t) > 0\}, \qquad b := \sup\{t \in \mathbb{R}\colon\ F(t) < 1\}.$$

Then we can and will consider F to be defined on a non-empty set T which is assumed to be equal to the interval (a, b) with endpoints $\{a\}$ and/or $\{b\}$

included or not included depending on $F(a) > 0$ and/or $F(b-) < 1$, respectively.

Let F_n be the empirical df based on a sample of independent identically distributed real rv's $X_1, \ldots, X_n$ with non-degenerate df F, and let $\alpha_n = \{\alpha_n(t);\ t \in T\}$ be the corresponding empirical process given by

$$\alpha_n(t) := \sqrt{n}(F_n(t) - F(t)), \quad t \in T. \tag{4.1}$$

We give another representation of α_n (see (4.4) below) which is better suited to characterize the central limit theorem and to handle L-statistics in the next section. First define a symmetrized empirical df S_n by

$$S_n(t) := \begin{cases} -F_n(t), & \text{if } t < c, \\ 1 - F_n(t), & \text{if } t \geqq c, \end{cases} \tag{4.2}$$

for some point $c \in T$ to be specified by Condition H in the next section and put

$$m(t) := \mathbb{E}S_n(t) = \begin{cases} -F(t), & \text{if } t < c, \\ 1 - F(t), & \text{if } t \geqq c. \end{cases} \tag{4.3}$$

Then

$$S_n(t) = \frac{1}{n} \sum_{i=1}^{n} \psi(X_i, t), \quad t \in T,$$

where the function ψ on $T \times T$ is given by

$$\psi(x, \cdot) = \begin{cases} -\chi_{[x,c)}, & \text{if } x < c, \\ \chi_{[c,x)}, & \text{if } x \geqq c, \end{cases}$$

for all $x \in T$, and

$$\alpha_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y(X_i, t), \quad t \in T, \tag{4.4}$$

where the function Y on $T \times T$ is given by

$$Y(x,t) = \psi(x,t) - m(t), \quad x, t \in T.$$

Note also that $F_n(t) - F(t) = m - S_n(t)$.

Let $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$ be a B.f.s., where μ is a σ-finite Lebesgue–Stieltjes measure on T such that $\mu([c,d]) < \infty$ for all compact subsets $[c,d] \subset T$. We will assume throughout this section that $\mathbb{B}$ is separable. Since μ is a separable measure, the B.f.s. $\mathbb{B}$ is separable if and only if it is order continuous. Let $\xi = \{\xi(t);\ t \in T\}$ be a $\Pr \times \mu$-measurable stochastic process with a.a. sample paths in $\mathbb{B}$. Then, in a standard way (see e.g. Cremers and Kadelka [3]),

ξ induces a Borel measurable map from Ω into $\mathbb{B}$, say ξ again, called a $\mathbb{B}$-rv, and a probability distribution $\mathcal{L}(\xi)$ on $\mathbb{B}$. Concerning the empirical process α_n, it follows from the representation (4.4) that α_n induces a $\mathbb{B}$-rv if and only if $m \in \mathbb{B}$.

Let $\mathbb{B}$ be a separable Banach space. A $\mathbb{B}$-rv η is Gaussian if $g(\eta)$ is a real-valued Gaussian rv for any continuous linear functional g on $\mathbb{B}$. Recall that one may describe any Gaussian $\mathbb{B}$-rv η by an operator R from the Banach space of continuous linear functionals $\mathbb{B}^*$ into $\mathbb{B}$, called *Gaussian covariance*, which appears in the expression of the characteristic function of η. The description of Gaussian covariances on B.f.s.'s is given by Gorgadze et al. [12]. In particular, they proved:

PROPOSITION 4.1. *Let $\mathbb{B}$ be an order continuous B.f.s.. Then the following is true:*

(1) *If a symmetric positive operator $R\colon \mathbb{B}^* \to \mathbb{B}$ is a Gaussian covariance, then there exists a measurable function $r : T \times T \to \mathbb{R}$ such that*

$$(Rg)(t) = \int_T g(s)r(s,t)\mu(ds), \quad t \in T,\ g \in \mathbb{B}^*, \tag{4.5}$$

and the function $t \to \sqrt{r(t,t)} \in \mathbb{B}$.

(2) *The operator R given by (4.5) with the symmetric positive definite function r such that $t \to \sqrt{r(t,t)} \in \mathbb{B}$, is Gaussian covariance if and only if $\mathbb{B}$ does not contain l_∞^n uniformly.*

According to this statement, a measurable version of a Brownian bridge $\eta = \{\eta(t)\colon t \in T\}$ with the covariance structure $F(t \wedge s) - F(t)F(s)$ has a.a. sample paths in $\mathbb{B}$, and hence induces a $\mathbb{B}$-rv, if and only if

$$\sqrt{F(1-F)} \in \mathbb{B}. \tag{4.6}$$

Let $\{\xi_n\colon n \geqq 1\}$ be a sequence of measurable stochastic processes with a.a. sample paths in a B.f.s. $\mathbb{B}$. We say that ξ_n satisfies the *central limit theorem (CLT) in* $\mathbb{B}$ if there exists a Gaussian $\mathbb{B}$-rv η such that $\mathcal{L}(\xi_n)$ converges weakly to $\mathcal{L}(\eta)$. Note that if $\zeta_1, \ldots, \zeta_n, \ldots$ are iid $\mathbb{B}$-rv's and if

$$\xi_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i, \tag{4.7}$$

then we say also that ζ_1 satisfies the CLT in $\mathbb{B}$. A $\mathbb{B}$-rv ξ is said to be *pregaussian* if there exists a Gaussian $\mathbb{B}$-rv $G(\xi)$ such that $Eg^2(\xi) = Eg^2(G(\xi))$ for all $g \in \mathbb{B}^*$. Note that ζ_1, as well as ξ_n itself, are pregaussian $\mathbb{B}$-rv's whenever the sequence ξ_n given by (4.7) satisfies the CLT in $\mathbb{B}$. We say that a Banach space $\mathbb{B}$ satisfies the inequality Ros(p), $1 \leqq p < +\infty$, if there is a constant C

such that for any finite sequence of independent pregaussian $\mathbb{B}$-rv's $\zeta_1, \ldots, \zeta_n$ with associated Gaussian $\mathbb{B}$-rv's $G(\zeta_1), \ldots, G(\zeta_n)$ (which may be assumed to be independent) we have

$$\mathbb{E}\Big\|\sum_i \zeta_i\Big\|^p \leqq C\Big(\sum_i \mathbb{E}\|\zeta_i\|^p + \mathbb{E}\Big\|\sum_i G(\zeta_i)\Big\|^p\Big).$$

In particular, $\mathbb{L}_p$-spaces with $1 \leqq p \leqq 2$, satisfy inequalities $\mathrm{Ros}(q)$ for every q, $1 \leqq q < \infty$, and $\mathbb{L}_p$-spaces with $2 < p < \infty$, satisfy the inequality $\mathrm{Ros}(p)$ for the corresponding p. For the proof of the following statement we refer to Theorem 10.10 in Ledoux and Talagrand [14].

PROPOSITION 4.2. *Let $\mathbb{B}$ be a separable Banach space satisfying the inequality* $\mathrm{Ros}(p)$ *for some $p > 2$, and let $\{\xi_n\colon n \geqq 1\}$ be a sequence of $\mathbb{B}$-rv's given by* (4.7), *where $\zeta_1, \ldots, \zeta_n, \ldots$ are iid $\mathbb{B}$-rv's. Then ξ_n satisfies the CLT in $\mathbb{B}$ if and only if ζ_1 is pregaussian and*

$$\lim_{x\to\infty} x^2 \Pr(\{\|\zeta_1\| > x\}) = 0. \tag{4.8}$$

According to this statement, the empirical process α_n given by (4.4) satisfies the CLT in a B.f.s. $\mathbb{B}$ if and only if (4.6) holds and

$$\lim_{x\to\infty} x^2 \Pr(\{\|\psi(X,\cdot)\| > x\}) = 0. \tag{4.9}$$

Next we show that the last condition is superfluous for the empirical processes.

LEMMA 4.3. *Let F be a df defined on T, and let $\mathbb{B} = (\mathbb{B}(T,\mu), \|\cdot\|)$ be an order continuous B.f.s. Then* (4.6) *implies* (4.9).

PROOF. Suppose first that

$$F(a) = 0 \quad \text{and} \quad F(b-) = 1. \tag{4.10}$$

Define a family $G = \{g_s\colon s \in T\}$ of functions on T by

$$g_s(t) := \begin{cases} \sqrt{F(s)}\,\chi_{[s,c)}(t), & \text{if } s < c, \\ \sqrt{1-F(s-)}\,\chi_{[c,s)}(t), & \text{if } s \geqq c, \end{cases}$$

for all $t \in T$. It is easy to check that

$$g_s \leqq 2\sqrt{F(1-F)}, \quad \forall s \in T.$$

Thus, by (4.6), the family G has u.a.c. norms. Moreover, by the assumption (4.10), it follows that

$$\lim_{s\downarrow a,\, s\uparrow b} g_s(t) = 0, \quad \forall t \in T.$$

Hence, by Lemma 2.1, one may conclude that

$$\lim_{s\downarrow a, s\uparrow b} \|g_s\| = 0.$$

Choose an arbitrary number $\epsilon > 0$, and take $s_1, s_2 \in T$ such that

$$\|g_s\| \leqq \sqrt{\epsilon/2}, \quad \forall s \in T \setminus (s_1, s_2).$$

Then for all $x > \|\chi_{(s_1,s_2)}\|^2 \vee \epsilon/2$, we have

$$\begin{aligned}
&x \Pr(\{ \|\psi(X,\cdot)\| > \sqrt{x} \}) = \\
&= x\Big[\Pr(\{ \|g_X\| > \sqrt{xF(X)},\ X \leqq s_1 \}) + \\
&\qquad + \Pr(\{ \|g_X\| > \sqrt{x(1-F(X-))},\ X \geqq s_2 \})\Big] \leqq \\
&\leqq x[\Pr(\{ F(X) \leqq \epsilon/(2x) \}) + 1 - \Pr(\{ F(X-) \leqq 1 - \epsilon/(2x) \})] \leqq \epsilon,
\end{aligned}$$

where in the last step we have used the inequalities

$$\Pr(\{ F(X) \leqq u \}) \leqq u$$

and

$$\Pr(\{ F(X-) \leqq u \}) \geqq u,$$

for all $u \in [0,1]$. See, e.g., p. 5 in Shorack and Wellner [29] for the proof of the first inequality. The second one follows from

$$u \leqq F(F^{-1}(u)), \quad \forall u \in [0,1],$$

and from the fact that $x \leqq F^{-1}(u)$ if and only if $F(x-) \leqq u$, where $F^{-1}(u) = \inf\{ x : F(x) \geqq u \}$ (see also p. 5 in Shorack and Wellner [29]). Since ϵ is an arbitrary positive number, Lemma 4.3 is proved under the assumption (4.10). Suppose now that $F(a) > 0$ and $F(b-) < 1$. Then (4.6) implies $\|\chi_T\| < \infty$, and hence (4.9) is true in this case as well. If $F(a) > 0$ and $F(b-) = 1$ or, conversely, $F(a) = 0$ and $F(b-) < 1$, then one may reduce the task to $\mathbb{B}$-rv's $\psi(X,\cdot)\chi_{[c,b)}$ and $\psi(X,\cdot)\chi_{(a,c)}$, respectively. Now (4.9) follows from (4.6) making use of analogous arguments to those in the case of (4.10). This completes the proof of Lemma 4.3.

COROLLARY 4.4. *Let α_n be the empirical process based on a df F, and let $\mathbb{B} = (\mathbb{B}(T,\mu), \|\cdot\|)$ be an order continuous B.f.s. satisfying the inequality* Ros(p) *for some $p > 2$. Then α_n satisfies the CLT in $\mathbb{B}$ if and only if* (4.6) *holds true.*

Assume for a moment that the empirical process α_n is bounded in probability in $\mathbb{B}$, i.e., for each $\epsilon > 0$ one can find a finite number M such that

$$\sup_n \Pr(\{ \|\alpha_n\| > M \}) \leqq \epsilon.$$

Due to the representation (4.4), one may invoke a terminology from Ledoux and Talagrand [14] and say that α_n satisfies the *bounded CLT in* $\mathbb{B}$. Then, by Theorem 10.3 in Ledoux and Talagrand [14], the $\mathbb{B}$-rv $Y(X_1,\cdot)$ is pregaussian whenever $\mathbb{B}$ does not contain an isomorphic copy of c_0. Thus, for the class of B.f.s.'s which satisfies the conditions of Corollary 4.4, the bounded CLT for the empirical process is equivalent to the (usual) CLT.

5. L-statistics

This section contains the proofs of main results.

Let F be a non-degenerate df defined on the interval T (see Definition F in the previous section), and let $X_{n:1} \leqq \ldots \leqq X_{n:n}$ be the order statistics corresponding to a sample from the df F. We consider a linear combination of a function of these order statistics, an L-statistic, given by

$$L_n = \frac{1}{n}\sum_{i=1}^{n} c_{ni} h(X_{n:i}),$$

for some weights constant $c_{n1}, \ldots, c_{nn}$ and for a function h described by

Condition H. Assume a function h to be left-continuous and of bounded variation on every compact subset of T. Suppose also that there is a point $c \in T$ such that $h(c) = 0$.

For any function h satisfying Condition H, there exists a signed Lebesgue–Stieltjes measure dh on T such that

$$h(x) = \int\limits_{[c,x)} dh = \begin{cases} \int\limits_{[c,x)} dh, & \text{if } x > c, \\ 0, & \text{if } x = c, \\ -\int\limits_{[x,c)} dh, & \text{if } x < c. \end{cases}$$

Associated with any such h, denote the induced total variation measure by μ_h. We note in passing that by the definition of the interval T and due to Condition H, μ_h is a finite measure on $[a,b]$ whenever $F(a) > 0$ and $F(-b) < 1$.

The L-statistic L_n may be expressed in the form

$$L_n = \int\limits_0^1 h(F_n^{-1}) J_n d\lambda = \int\limits_T h\, d\Big(\int\limits_{1/2}^{F_n} J_n d\lambda\Big), \tag{5.1}$$

where F_n, F_n^{-1} are the empirical df, the empirical quantile function, respectively, corresponding to a sample from the df F and J_n is defined by (1.9).

The integral representation (5.1) may be considered as a functional on a class of step functions. To extend it to larger classes of functions, the following formalities seem to be useful. Define a finite measure ν on $[0,1]$ by

$$\nu((c,d]) := \int_c^d J\,d\lambda, \qquad \forall\, 0 \leqq c < d \leqq 1,$$

and for any measurable map $H : [0,1] \to T$ define $H(\nu)$ to be the image measure on T. Note that for any df F

$$F^{-1}(\nu)((x_1,x_2]) = \int_{F(x_1)}^{F(x_2)} J\,d\lambda, \qquad (x_1,x_2] \subset T, \tag{5.2}$$

where $F^{-1}(x) = \inf\{t \in T : F(x) \geqq t\}$. Thus, by the image measure theorem and due to (5.2), for any function $h \in \mathbb{L}_1(T, F^{-1}(\nu))$ we have $h \circ F^{-1} \in \mathbb{L}_1([0,1],\nu)$ and

$$\begin{aligned} \int_0^1 h \circ F^{-1} J\,d\lambda &= \int_0^1 h \circ F^{-1} d\nu = \int_T h\,dF^{-1}(\nu) \\ &= \int_T h\,d\Big(\int_{1/2}^F J\,d\lambda\Big) =: L_h(J,F) = L(J,F). \end{aligned} \tag{5.3}$$

Note that, by (5.1), $L_n = L(J_n, F_n)$.

The following representation of an L-statistic goes back to Shorack [28] and its various forms have been used later on in many papers.

LEMMA 5.1. *Consider a df F, a function h satisfying Condition H, and a Lebesgue integrable function J over $[0,1]$. Assume that*

$$\int_{(a,c)} \Big|\int_0^F J\,d\lambda\Big|\,d\mu_h + \int_{[c,b)} \Big|\int_F^1 J\,d\lambda\Big|\,d\mu_h < \infty. \tag{5.4}$$

Then the functional $L_h(J,F)$ given by (5.3) exists and

$$L(J,F_n) - L(J,F) = -\int_T \Big[\int_F^{F_n} J\,d\lambda\Big]\,dh \qquad a.s.. \tag{5.5}$$

PROOF. First assume that

$$F(a)=0 \quad \text{and} \quad F(b-)=1. \tag{5.6}$$

One may rewrite (5.3) into the following form

$$L(J,F) = \int_{(a,c]} h\, d\Big(\int_0^F J\, d\lambda\Big) - \int_{(c,b)} h\, d\Big(\int_F^1 J\, d\lambda\Big). \tag{5.7}$$

Take any point $d\in(a,c)$. Integration by parts for Lebesgue–Stieltjes integrals yields

$$\int_{(d,c]} h\, d\Big(\int_0^F J\, d\lambda\Big) = -h(d)\int_0^{F(d)} J\, d\lambda - \int_{[d,c)} \Big[\int_0^F J\, d\lambda\Big]\, dh.$$

By the assumption (5.7), it follows that

$$\lim_{d\downarrow a} \chi_{[d,c)}(t) \int_0^{F(d)} |J|\, d\lambda = 0 \qquad \forall\, t\in T.$$

Hence, due to (5.4), by the dominated convergence theorem we get

$$\lim_{d\downarrow a} |h(d)| \int_0^{F(d)} |J|\, d\lambda = 0,$$

since $|h(d)|\leq \mu_h([d,c))$. Thus, letting d go to a, one may conclude the existence of the first integral in the right side of (5.7). The existence of the second one follows in the same way. Now, if the assumption (5.6) does not hold, i.e., if the df F has a jump at one or both endpoints, then one may use integration by parts for the corresponding integrals over the intervals $[a,c]$, $(c,b]$, or $[a,b]$, respectively. The representation (5.5) follows by performing integration by parts in the same way for all samples such that $a<X_{n:1}\leqq X_{n:n}<b$ whenever $F(a)=0$ and/or $F(b-)=1$. Now the proof of Lemma 5.1 is complete.

For the following statement, recall the definition of the symmetrized empirical df S_n and its expectation m, respectively given by (4.2) and (4.3) above.

PROPOSITION 5.2. *Consider the L-statistic L_n corresponding to a sample from a df F and with a function h satisfying condition H.*

I. *Assume the weights constant to be given by a score function J (see (1.1)) such that the representation (5.5) holds. Suppose also that there exists a B.f.s. $\mathbb{B}$ and a family of operators $\{\Phi_\epsilon : \epsilon > 0\}$ such that:*

(i) *the empirical process α_n given by (4.1) has a.a. sample paths in $\mathbb{B}$ and the sequence $\{\alpha_n : n \geqq 1\}$ is uniformly $C(\mathbb{B})$-tight for some class of bounded subsets $C(\mathbb{B})$ of $\mathbb{B}$, i.e., for every $\epsilon > 0$ there exists $K \in C(\mathbb{B})$ such that*

$$\sup_n \Pr(\{\alpha_n \in K^c\}) \leqq \epsilon; \tag{5.8}$$

(ii) *for every $\epsilon > 0$ and for all sufficiently large $n \geqq 1$ there exists a subset $A_{n\epsilon}$ of Ω having $\Pr(A_{n\epsilon}) > 1 - \epsilon$ on which*

$$\Phi_\epsilon(S_n) - \Phi_\epsilon(m) = \int_F^{F_n} J \, d\lambda;$$

(iii) *for every $\epsilon > 0$, Φ_ϵ maps $\mathbb{B}$ into $\mathbb{L}_1(T, \mu_h)$ and Φ_ϵ is $C(\mathbb{B})$-differentiable at m with the derivative*

$$\Phi'_\epsilon(m)(f) = a(F)(\cdot)\, f = a\, f,$$

for all $f \in \mathbb{B}$.

Then the L-statistic L_n satisfies the central limit theorem, i.e.,

$$\sqrt{n}(L_n - L(J, F)) \xrightarrow{d} N(0, \sigma^2(a, F)), \tag{5.9}$$

where

$$\sigma^2(a, F) = \int_T \int_T [F(t \wedge s) - F(t)F(s)] a(t) a(s)\, h(dt)\, h(ds).$$

II. *Assume the weights constant to be arbitrary and let the representation (5.5) hold for all $J = J_n$, $n \geqq 1$ given by (1.9). Suppose also that there exists a B.f.s. $\mathbb{B}$, a normed space of functions $(\Theta, \|\cdot\|)$ and an operator Φ such that (i) of I holds;*

(ii) *$\{J_n : n \geqq 1\} \subset \Theta$ and there exists $J \in \Theta$ such that $\|J_n - J\| \to 0$;*

(iii)

$$\Phi(J_n, S_n) - \Phi(J, m) = \int_F^{F_n} J_n d\lambda, \quad \forall n \geqq 1;$$

(iv) *Φ maps $\Theta \times \mathbb{B}$ into $\mathbb{L}_1(T, \mu_h)$ and Φ is C-differentiable at (J, m) with the derivative*

$$\Phi'(J, m)(\theta, f) = a(J, F)(\cdot) f = a\, f,$$

for all $f \in \mathbb{B}$ *and* $\theta \in \Theta$, *where* $\mathcal{C} = \{B \times K : B$ *is a ball in* Θ, $K \in \mathcal{C}(\mathbb{B})\}$.

Then the L-statistic L_n *satisfies the central limit theorem, i.e.,* (5.9) *holds with* $L(J_n, F)$ *instead of* $L(J, F)$.

PROOF. We prove the first part only because the second one is analogous. Due to the representation (5.5), we have

$$\sqrt{n}(L_n - L(J,F)) = -\int_T \left\{ \sqrt{n} \int_F^{F_n} J\, d\lambda + a\, \alpha_n \right\} dh + \int_T \alpha_n a\, dh$$
$$=: r_n + \langle \alpha_n, a \rangle .$$

Note that

$$< \alpha_n, a > = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i,$$

where $Y_1, \ldots, Y_n, \ldots$ are independent identically distributed zero mean rv's with $\mathbb{E}Y_1^2 = \sigma^2(a, F)$. Due to the classical central limit theorem our task is to prove that

$$\lim_{n\to\infty} r_n = 0 \quad \text{in probability} . \tag{5.10}$$

The remainder in the differentiation of the operator Φ_ϵ is

$$\Delta^\epsilon(f) = \Phi_\epsilon(m+f) - \Phi_\epsilon(m) - a\, f,$$

for all $f \in \mathbb{B}$. Then, by the assumption (ii) and since $\alpha_n = -\sqrt{n}(S_n - m)$, we have on the set $A_{n\epsilon}$

$$r_n = -\int_T \sqrt{n}\{\Phi_\epsilon(S_n) - \Phi_\epsilon(m) + a\, \alpha_n/\sqrt{n}\} dh$$

$$= -\int_T \sqrt{n}\{\Delta^\epsilon(-\alpha_n/\sqrt{n})\} dh.$$

Choose an arbitrary number $\epsilon > 0$. In virtue of the assumption (i), there exists a set $K \in \mathcal{C}(\mathbb{B})$ such that (5.8) holds. By $\mathcal{C}(\mathbb{B})$-differentiability of Φ_ϵ, there exists a number $N \geqq 1$ such that

$$\sqrt{n} \| \Delta^\epsilon(f/\sqrt{n}) \|_{L_1(T,\mu_h)} < \epsilon,$$

uniformly for $f \in K$ and all $n \geqq N$. Therefore, we have the inequality

$$\Pr(\{ |r_n| \geqq \epsilon \}) \leqq \Pr(A_{n\epsilon}^c) + \Pr(\{ \alpha_n \in K^c \})$$

$$+ \Pr\bigg(\bigg\{\int_T \sqrt{n}|\Delta^\epsilon(-\alpha_n/\sqrt{n})|d\mu_h \geqq \epsilon\bigg\}\bigg) \leqq 2\epsilon$$

for all $n \geqq N$. Since ϵ is an arbitrary number, the desired relation (5.10) holds true and the proof of Proposition 5.2 is now complete.

Now, we are ready to give the proof of the first main result.

Proof of Theorem 1.1. It is based on the part I of Proposition 5.2. First note that the representation (5.5) holds by Lemma 5.1, since (5.4) follows from (1.2) and (1.4). To verify assumptions (i)–(iii), we are going to use Corollary 4.4 and Proposition 3.1. For the B.f.s. $\mathbb{B}$ take a weighted B.f.s. $(\mathbb{L}_1(T,\mu_h))_{w_{p,q}} = \mathbb{L}_1(T, w_{p,q}\mu_h)$, where

$$w_{p,q} = F^{p/2-1/2}\chi_{(a,c)} + (1-F)^{q/2-1/2}\chi_{[c,b)}. \tag{5.11}$$

Then, by Corollary 4.4 and (1.2), the assumption (i) holds with the class $\mathcal{C}(\mathbb{B})$ being all norm compact sets of $\mathbb{L}_1(T, w_{p,q}\mu_h)$. To define the family of operators $\{\Phi_\epsilon\colon \epsilon > 0\}$, recall the linear bounds in probability for the empirical df F_n (see van Zuijlen [35] for the case when F is an arbitrary df); namely, for given an $\epsilon > 0$ there exist $M_\epsilon \in (0,1)$ and a subset $A_{n\epsilon}$ of Ω such that $\Pr(A_{n\epsilon}) > 1-\epsilon$ and on $A_{n\epsilon}$

$$|S_n(t)| \leqq |m(t)|/M_\epsilon, \qquad \forall\, t \in T.$$

Moreover, by the Glivenko–Cantelli theorem, for any $\gamma \in (1/2 \vee \sup_t |m(t)|, 1)$ there exists a finite number N such that a.s.

$$|S_n(t)| \leqq |F_n(t) - F(t)| + |m(t)| \leqq \gamma, \qquad \forall\, t \in T,$$

and all $n \geqq N$. Define a family of functions $\{\phi_\epsilon\colon \epsilon > 0\}$ on $T \times \mathbb{R}$ by

$$\phi_\epsilon(t,x) := \int_{|m(t)|}^{|x|\wedge|m(t)|/M_\epsilon\wedge\gamma} J^c(s,t)\,ds, \qquad t\in T,\ x \in \mathbb{R}, \tag{5.12}$$

where

$$J^c(s,t) = J(s)\chi_{(a,c)}(t) - J(1-s)\chi_{[c,b)}(t), \qquad s \in [0,1],\ t \in T. \tag{5.13}$$

Define a family of superposition operators $\{\Phi_\epsilon\colon \epsilon > 0\}$ by (3.1) and note that the assumption (ii) holds. We will show that the assumption (iii) holds, too, using Proposition 3.1 for the B.f.s. $\mathbb{B} = \mathbb{L}_1(T,\mu_h)$, for the weight function $w = w_{p,q}$ given by (5.11), for the sup-measurable function $\phi = \phi_\epsilon$ given by (5.12), for $f_0 = m$ and $a^{f_0} = J \circ F$. Note that $\Phi_\epsilon(m) = 0$ and (3.3) is true due to Lebesgue's theorem on derivation of the indefinite integral and by (1.3).

It is easy to see that (3.4) is nothing else than (1.4) and $mw_{p,q} \in \mathbb{L}_1(T, \mu_h)$ because

$$|mw_{p,q}| \leqq F^{p/2}(1-F)^{q/2} \in \mathbb{L}_1(T, \mu_h).$$

All what is left is the verification of the most tedious condition (3.5). Using the properties of the number γ, by (1.4) it follows that we have

$$|\Phi_\epsilon(m+x) - \Phi_\epsilon(m)|$$

$$\leqq C_1\chi_{(a,c)} \int_{|m|}^{|m+x|\wedge|m|/M_\epsilon} s^{p/2-1/2}ds + C_2\chi_{[c,b)} \int_{|m|}^{|m+x|\wedge|m|/M_\epsilon} s^{q/2-1/2}ds$$

$$=: C_1 I_p \chi_{(a,c)} + C_2 I_q \chi_{[c,b)}. \tag{5.13}$$

We estimate the integral I_p only, since the estimation of I_q is identical. Invoking the inequality ($p > 0$)

$$||a|^{p+1} - |b|^{p+1}| \leqq (p+1)2^p[|a-b|^{p+1} + |b|^p|a-b|],$$

we arrive at

$$\int_{|m|}^{|m+x|} s^{p/2-1/2}ds = \frac{2}{p+1}\frac{|m+x|^{p+1} - |m|^{p+1}}{|m+x|^{p/2+1/2} + |m|^{p/2+1/2}}$$

$$\leqq 2^{p+1}|x|[|x|^p|m|^{-p/2-1/2} + |m|^{p/2-1/2}].$$

For all x such that $|m+x| \leqq |m|/M_\epsilon$, we have $|x| \leqq (1+1/M_\epsilon)|m|$ and

$$I_p \leqq 2^{p+1}[1 + (1+1/M_\epsilon)]|m|^{p/2-1/2}|x|.$$

Otherwise, $|x| \geqq (1/M_\epsilon - 1)|m|$ and

$$I_p = \int_{|m|}^{|m|/M_\epsilon} s^{p/2-1/2}ds \leqq \frac{2}{p+1}(1/M_\epsilon - 1)(M_\epsilon^{-p/2-1/2} - 1)|m|^{p/2-1/2}|x|.$$

From these we get the bound

$$I_p \leqq C_{p,\epsilon}|x||m|^{p/2-1/2},$$

for some finite constant $C_{p,\epsilon}$. Returning to (5.13), one may conclude that the desired condition (3.5) holds true. Thus, by Proposition 3.1, the superposition operator Φ_ϵ maps $\mathbb{L}_1(T, w_{p,q}\mu_h)$ into $\mathbb{L}_1(T, \mu_h)$ and Φ_ϵ is $\mathcal{C}_L$-differentiable with the derivative

$$\Phi_\epsilon'(m)f = J \circ F \cdot f,$$

for every $\epsilon > 0$. Since $\mathbb{L}_1(T, w_{p,q}\mu_h)$ is order continuous B.f.s., by Theorem 4.2 in Dodds and Fremlin [4] every norm compact set is L-weakly compact set too, i.e., $\mathcal{C}(\mathbb{B}) \subset \mathcal{C}_L$. Thus, the assumption (iii) of Proposition 5.2 is satisfied. Now, the statement of Theorem 1.1 is a consequence of Proposition 5.2 and this also completes the proof.

We conclude with the proof of the second main result.

PROOF OF THEOREM 1.2. We will deduce it from part II of Proposition 5.2. Assume first that $p > 1$. To check the representation (5.5) for $J = J_n$, one can verify the condition (5.4) of Lemma 5.1. Using Hölder's inequality, the image measure theorem and Muckenhoupt's weighted version of Hardy–Littlewood maximal theorem (see (3.20) for the case $r = p'$, $J = J_n$ and $v = dF(\mu_h)/d\lambda$), we have for the first integral in (5.4)

$$\int\limits_{(a,c)} \left| \int\limits_0^F J_n d\lambda \right| d\mu_h \leqq \|F\chi_{(a,c)}\|_{L_p(\mu_h)} \|MJ_n\|_{L_{p'}(F(\mu_h))}$$

$$\leqq C_{p'} \|\sqrt{F}\chi_{(a,c)}\|_{L_p(\mu_h)} \|J_n\|_{L_{p'}(F(\mu_h))} < +\infty.$$

Since the estimation of the second integral in (5.4) is analogous, by Lemma 5.1 we conclude that the representation (5.5) holds for $J = J_n$. To verify the other conditions of Proposition 5.2, take a B.f.s. $\mathbb{B} = \mathbb{L}_p(T, \mu_h)$, a normed space $\Theta = \mathbb{L}_{p'}([0,1], F(\mu_h))$ and a superposition operator Φ defined by (3.9) and by a linear family of Carathéodory functions ϕ on $\Theta \times \mathbb{R} \times T$ defined by

$$\phi(J; t, x) = \int\limits_{|m(t)|}^{|x|} J^c(s,t) ds,$$

where J^c is given by (5.13). The assumption (i) follows from Corollary 4.4 and (1.10) with the class $\mathcal{C}(\mathbb{B})$ of all norm compact sets in $\mathbb{B}$. Since the assumption (ii) is obviously satisfied and it is easy to see that

$$\Phi(J_n, S_n)(t) - \Phi(J, m)(t) = \int\limits_{|m(t)|}^{|S_n(t)|} J_n^c(s,t)\, ds = \int\limits_{F(t)}^{F_n(t)} J_n d\lambda, \quad \forall t \in T,$$

we have to check only the assumption (iv). For this purpose we use Proposition 3.6 where the pair of B.f.s.'s is taken to be $\mathbb{B}_1 = \mathbb{L}_p(T, \mu_h)$ and $\mathbb{B}_2 = \mathbb{L}_1(T, \mu_h)$, Θ as above, $f_0 = m$, $\theta_0 = J$ and $a = J \circ F$. Then the generalized dual space $\mathbb{B}_2/\mathbb{B}_1 = \mathbb{L}_{p'}(T, \mu_h)$ is order continuous because we have assumed $p > 1$. It is plain that

$$M\Phi(J, m)(t) = \sup_{x \neq 0} \left| \frac{1}{x} \int\limits_{|m(t)|}^{|m(t)+x|} J^c(s,t)\, ds \right| \leqq (MJ)(t),$$

where MJ is a Hardy–Littlewood maximal function (1.7). Thus, (3.10) is a consequence of Muckenhoupt's weighted version of the Hardy–Littlewood maximal theorem (see (3.20) for $r=p'$ and $v=dF(\mu_h)/d\lambda$). Since $\Phi(J,m)=0$ and (3.12) holds by Lebesgue's theorem on derivation of the indefinite integral, by Proposition 3.6 one may conclude that Φ maps $\mathbb{L}_{p'}([0,1],F(\mu_h)) \times \mathbb{L}_p(T,\mu_h)$ into $\mathbb{L}_1(T,\mu_h)$ and Φ is Fréchet differentiable at (J,m) with the derivative

$$\Phi'(J,m)(\theta,f) = J \circ F \cdot f.$$

This yields the assumption (iv) with $a = J \circ F$, and by Proposition 5.2 we may infer that the statement of Theorem 1.2 holds in the case $p > 1$. For the case $p = 1$ one may follow the pattern of the previous case only using Proposition 3.4 instead of Proposition 3.6. We omit obvious details and Theorem 1.2 is thus established.

REFERENCES

[1] APPELL, J. and ZABREJKO, P. P., *Nonlinear superposition operators*, Cambridge tracts in mathematics, **95**, Cambridge University Press, Cambridge, 1990. *MR* **91k**: 47168

[2] BOOS, D. D., A differential for L-statistics, *Ann. Statist.* **7** (1979), 955–959. *MR* **80i**:62014

[3] CREMERS, H. and KADELKA, D., On weak convergence of stochastic processes with Lusin path spaces, *Manuscripta Math.* **45** (1984), 115–125. *MR* **85i**:60005

[4] DODDS, P. G. and FREMLIN, D. H., Compact operators in Banach lattices, *Israel J. Math.* **34** (1979), 287–320. *MR* **81g**:47037

[5] DUDLEY, R. M., Differentiability of the composition and inverse operators for regulated and a.e. continuous functions, 1991 (preprint). See *MR* **94f**:60011

[6] DUNFORD, N. and SCHWARTZ, J. T., *Linear operators. Part I. General theory*, Interscience Publishers, Inc., New York, 1958, 1967. *MR* **22** #8302

[7] ELDIK, P. VAN and GROBLER, J. J., Lebesgue-type convergence theorems in Banach lattices with applications to compact operators, *Nederl. Akad. Wetensch. Indag. Math.* **41** (1979), 425–437. *MR* **81b**:46031

[8] ESTY, W., GILLETTE, R., HAMILTON, M. and TAYLOR, D., Asymptotic distribution theory of statistical functionals: the compact derivative approach for robust estimators, *Ann. Inst. Statist. Math.* **37** (1985), 109–129. *MR* **86j**:62042

[9] FEFFERMAN, C. and STEIN, E. M., Some maximal inequalities, *Amer. J. Math.* **93** (1971), 107–115. *MR* **44** #2026

[10] FERNHOLZ, L. T., *von Mises calculus for statistical functionals*, Lecture Notes in Statist., **19**, Springer, New York, 1983. *MR* **85j**:62005

[11] GARCIA-CUERVA, J. and RUBIO DE FRANCIA, J. L., *Weighted norm inequalities and related topics*, Notas de matematica, **104**, North-Holland Mathematics Studies, **116**, North-Holland, Amsterdam, 1985. *MR* **87d**:42023

[12] GORGADZE, Z. G., TARIELADZE, V. I. and ČOBANJAN, S. A., Gaussian covariances in Banach sublattices of the space $L_0(T,\Sigma,\nu)$, *Dokl. Akad. Nauk SSSR* **241** (1978), 528–531. *MR* **80a**:60042; *Soviet Math. Dokl.* **19** (1978), 885–888.

[13] KHINTCHINE, A., Recherches sur la structure des fonctions mesurables, II, *Fund. Math.* **9** (1927), 212–279. *Jb. Fortschritte Math.* **53**, 229

[14] LEDOUX, M. and TALAGRAND, M., *Probability in Banach spaces. Isoperimetry and processes*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **23**, Springer-Verlag, Berlin, 1991. *MR* **93c**:60001

[15] LUXEMBURG, W. A. J., *Banach function spaces*, Thesis, Technische Hogeschool te Delft, Delft, 1955. *MR* **17**, 285a

[16] MALIGRANDA, L. and PERSSON, L.-E., Generalized duality of some Banach function spaces, *Nederl. Akad. Wetensch. Indag. Math.* **51** (1989), 323–338. *MR* **91b**: 46028

[17] MASON, D. M., Asymptotic normality of linear combinations of order statistics with a smooth score function, *Ann. Statist.* **9** (1981), 899–904. *MR* **83a**:62104

[18] MASON, D. M. and SHORACK, G. R., Necessary and sufficient conditions for asymptotic normality of trimmed L-statistics, *J. Statist. Plann. Inference* **25** (1990), 111–139. *MR* **91h**:62047

[19] MEYER-NIEBERG, P., Zur schwachen Kompaktheit in Banachverbänden, *Math. Z.* **134** (1973), 303–315. *MR* **48** #9341

[20] MUCKENHOUPT, B., Weighted norm inequalities for the Hardy maximal function, *Trans. Amer. Math. Soc.* **165** (1972), 207–226. *MR* **45** #2461

[21] NORVAIŠA, R., Laws of large numbers for L-statistics, *J. Applied Math. Stochastic Anal.* **7** (1994), 125–143. *MR* **95h**:60052

[22] REEDS, J. A., *On the definition of von Mises functionals*, Ph.D. Thesis, Research Rep. No. S–44, Dept. Statist., Harvard Univ., Cambridge, MA, 1976. See *MR* **91d**:62042

[23] SAKS, S., *Theory of the integral*, 2., revised ed., Monografie Matematyczne, 7, G. E. Stechert and Co., New York, NY, 1937. *Zbl* **7**, 105 and **17**, 300–301

[24] SEBASTIÃO E SILVA, J., Le calcul différentiel et intégral dans les espaces localement convexes, réels ou complexes I, II, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat.* (8) **20** (1956), 743–750; **21** (1956), 40–46. *MR* **19**, 561e; *MR* **19**, 561f

[25] SERFLING, R. J., *Approximation theorems of mathematical statistics*, Wiley series in probability and mathematical statistics, John Wiley and Sons, New York, 1980. *MR* **82a**:62003

[26] SHAO, J., Some results for differentiable statistical functionals, *Nonparametric statistics and related topics* (Ottawa, ON, 1991), North-Holland, Amsterdam, 1992, 179–188. *MR* **94f**:62073

[27] SHORACK, G. R., Asymptotic normality of linear combinations of functions of order statistics, *Ann. Math. Statist.* **40** (1969), 2041–2050. *MR* **40** #6672

[28] SHORACK, G. R., Functions of order statistics, *Ann. Math. Statist.* **43** (1972), 412–427. *MR* **46** #1001

[29] SHORACK, G. R. and WELLNER, J. A., *Empirical processes with applications to statistics*, Wiley series in probability and mathematical statistics: Probability and mathematical statistics, John Wiley and Sons, New York, 1986. *MR* **88e**:60002

[30] STEIN, E. M., *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, NJ, 1970. *MR* **44** #7280

[31] STIGLER, S. M., Linear functions of order statistics, *Ann. Math. Statist.* **40** (1969), 770–788. *MR* **41** #9413

[32] STIGLER, S. M., Simon Newcomb, Percy Daniell, and the history of robust estimation 1885–1920, *J. Amer. Statist. Assoc.* **68** (1973), 872–879. *MR* **50** #15015

[33] STIGLER, S. M., Linear functions of order statistics with smooth weight functions, *Ann. Statist.* **2** (1974), 676–693. *MR* **51** #9354

[34] ZAANEN, A. C., *Riesz spaces, II*, North-Holland mathematical library, Vol. 30, North-Holland, Amsterdam, 1983. *MR* **86b**:46001

[35] ZUIJLEN, M. C. A. VAN, *Empirical distributions and rank statistics*, Mathematical Centre Tracts, No. **79**, Mathematisch Centrum, Amsterdam, 1977. *MR* **55** #6684

(Received April 8, 1996)

INSTITUTE OF MATHEMATICS AND INFORMATICS
LITHUANIAN ACADEMY OF SCIENCES
AKADEMIJOS 4
VILNIUS 2600
LITHUANIA

norvaisa@ktl.mii.lt

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 239–250

MODERATE DEVIATION OF A BRANCHING WIENER PROCESS

P. RÉVÉSZ

Dedicated to Professor E. Csáki on the occasion of his 60th birthday

1. Introduction

Consider the following model:

(i) a particle starts from the position $0 \in \mathbb{R}^d$ and executes a Wiener process $W(t) \in \mathbb{R}^d$,

(ii) arriving at time $t=1$ to the new location $W(1)$ it dies,

(iii) at death it is replaced by Y offspring where

$$\mathbf{P}\{Y=l\}=p_l \qquad (l=0,1,2,\dots)$$

and

$$p_l \geqq 0, \qquad \sum_{l=0}^{\infty} p_l = 1,$$

(iv) each offspring, starting from where its ancestor dies, executes a Wiener process (from its starting point) and repeats the above given steps and so on. All Wiener processes and offspring-numbers are assumed independent of one another.

A more formal definition is given in Chapter 6 of [3].

Let $A \subset \mathbb{R}^d$ be a Borel set and let $\lambda(A,t)$ $(t=0,1,2,\dots)$ be the number of particles located in A at time t. Then

$$B(t)=\lambda(\mathbb{R}^d,t)$$

is the number of particles living at t and $\{B(t), t=0,1,2,\dots\}$ is a branching process. From now on we assume that

$$1<m=\sum_{k=0}^{\infty} kp_k < \infty$$

1991 *Mathematics Subject Classification.* Primary 60J80; Secondary 60J65.
Key words and phrases. Branching process, Wiener process, central limit theorem.

and

$$0<\sigma^2=\sum_{k=0}^{\infty}(k-m)^2p_k<\infty.$$

It is well known (cf. [2]) that the limit

$$\lim_{t\to\infty}\frac{B(t)}{m^t}=B$$

exists a.s. and

$$\mathbf{E}\left|\frac{B(t)}{m^t}-B\right|=O(m^{-t/2})\qquad (t=1,2,\dots) \tag{1.1}$$

where B is a non-negative r.v. with

$$\mathbf{E}B=1, \tag{1.2}$$

$$\mathbf{P}\{B=0\}=q \tag{1.3}$$

and $q<1$ depends on the distribution $\{p_k\}$.

It is easy to see that

$$\mathbf{E}(\lambda(A,T)\mid B(T))=B(T)\int_A\varphi(x,T)dx \tag{1.4}$$

where

$$\varphi(x,T)=\varphi_d(x,T)=(2\pi T)^{-d/2}\exp\Big(-\frac{x^2}{2T}\Big).$$

(1.4) suggests that $\lambda(A,T)$ as $T\to\infty$ should behave like the right-hand side of (1.4). A result, saying that it is indeed so, is the following

THEOREM A ([3]). *Let*

$$\mathcal{C}(x)=\mathcal{C}(x_1,x_2,\dots,x_d)=$$
$$=\Big\{y=(y_1,y_2,\dots,y_d):|x_i-y_i|\leqq\frac{1}{2}\ (i=1,2,\dots,d)\Big\},$$
$$\lambda(x,T)=\lambda(\mathcal{C}(x),T)$$

and $x=x(T)\in\mathbb{Z}^d$ be a sequence with $\|x\|\leqq T^\gamma$ $(0\leqq\gamma\leqq1)$. Then for any $\varepsilon>0$

$$\mathbf{P}\Big\{T^{(d+2-2\gamma-2\varepsilon)/2}\Big|\frac{\lambda(x,T)}{m^T}-H(\mathcal{C}(x),0,T)B\Big|\geqq1\Big\}\leqq\exp(-CT^\delta)$$

where $C>0$,

$$H(A,y,t)=\begin{cases}\int_A\varphi(x-y,t)dx & \text{if } t>0,\\ 1 & \text{if } t=0 \text{ and } y\in A,\\ 0 & \text{if } t=0 \text{ and } y\notin A\end{cases}$$

and δ is a small enough positive number.

In case $x=0$ we get the following consequence

THEOREM B. *For any $\varepsilon>0$ there exist a $C=C(\varepsilon)>0$ and a $\delta=\delta(\varepsilon)>0$ such that*

$$\mathbf{P}\left\{T^{1-\varepsilon}\left|\frac{\lambda(0,T)}{m^T}(2\pi T)^{d/2}-B\right|\geqq 1\right\}\leqq \exp(-CT^{\delta})$$

for any $T=1,2,\ldots$.

It is worthwhile to mention that this theorem is the best possible in the following sense:

THEOREM C ([3]). *For any $C>0$ there exists a $\delta=\delta(C)>0$ such that*

$$\mathbf{P}\left\{T\left|\frac{\lambda(0,T)}{m^T}(2\pi T)^{d/2}-B\right|\geqq C\right\}\geqq \delta.$$

Note that in case $\|x\|=T^{\gamma}$ $(\gamma>1/2)$ Theorem A does not say too much on the limit behaviour of $m^{-T}\lambda(x,T)$. The case $\|x\|\gg T^{1/2}$ is studied by Biggins [1]. In fact it is proved that the limit behaviour of $\lambda(x,T)$ is similar to that of the right-hand side of (1.4) even if x is large but no rate of convergence is given. In the present paper we intend to study the rate of convergence when $\|x\|\gg T^{1/2}$ but $\|x\|\ll T$. The expression "moderate deviation" refers to this fact.

Our main result is the following:

THEOREM. *Let $h=h(T)$ $(T=1,2,\ldots)$ be a function with*

$$(\log T)^{1+\varepsilon}\leqq h\leqq T(\log T)^{-\varepsilon-1}\qquad(\varepsilon>0).$$

Then

$$\mathbf{P}\left\{\frac{T}{h(T)(\log T)^{1+\varepsilon/2}}\sup_{\substack{x\in Z^d\\ \|x\|\leqq h}}\left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right|\geqq(\log T)^{-\varepsilon/4}\right\}\leqq T^{-2}.$$

Consequently,

$$\lim_{T\to\infty}\frac{T}{h(T)(\log T)^{1+\varepsilon/2}}\sup_{\substack{x\in Z^d\\ \|x\|\leqq h}}\left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right|=0\qquad a.s..$$

In order to enlighten the meaning of the above Theorem we give two examples.

EXAMPLE 1. Let

$$h=T^{\gamma}(\log T)^{1+\varepsilon}\qquad(0\leqq\gamma<1,\ \varepsilon>0).$$

Then

$$\mathbf{P}\left\{\frac{T^{1-\gamma}}{(\log T)^{2+2\varepsilon}} \sup_{\substack{x\in\mathbb{Z}^d \\ ||x||\leqq T^{\gamma}(\log T)^{1+\varepsilon}}} \left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right| \geqq (\log T)^{-\varepsilon/4}\right\} \leqq T^{-2}$$

and

$$\lim_{T\to\infty} \frac{T^{1-\gamma}}{(\log T)^{2+2\varepsilon}} \sup_{\substack{x\in\mathbb{Z}^d \\ ||x||\leqq T^{\gamma}(\log T)^{1+\varepsilon}}} \left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right| = 0 \quad \text{a.s.}.$$

Note that this result is clearly stronger than Theorem A and slightly stronger than Theorem B.

EXAMPLE 2. Let

$$h = T(\log T)^{-\alpha} \qquad (\alpha > 1).$$

Then for any $\varepsilon > 0$

$$\mathbf{P}\left\{(\log T)^{\alpha-1-\varepsilon} \sup_{\substack{x\in\mathbb{Z}^d \\ ||x||\leqq T(\log T)^{-\alpha}}} \left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right| \geqq (\log T)^{-\varepsilon/4}\right\} \leqq T^{-2}$$

and

$$\lim_{T\to\infty} (\log T)^{\alpha-1-\varepsilon} \sup_{\substack{x\in\mathbb{Z}^d \\ ||x||\leqq T(\log T)^{-\alpha}}} \left|\frac{\lambda(x,T)}{m^T}(\varphi(x,T))^{-1}-B\right| = 0 \quad \text{a.s.}.$$

The proof of the Theorem in case $d > 1$ is the same as in case $d = 1$. Hence the proof will be presented only in case $d = 1$.

2. Two lemmas

Introduce the following notations:

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-u^2/2} du,$$

$$\mathcal{J}_1 = \mathcal{J}_1(x,T) = \int_{x-1/2}^{x+1/2} \varphi(u,T) du,$$

$$\mathcal{J}_2 = \mathcal{J}_2(x,T,t) =$$

$$= \int_{-\infty}^{+\infty} \left(\int_{x-1/2}^{x+1/2} \varphi(\xi - y, T - t) d\xi \right)^2 \varphi(y, t) dy =$$

$$= \int_{-\infty}^{+\infty} (\mathcal{J}_1(x - y, T - t))^2 \varphi(y, t) dy.$$

LEMMA 1. *Assume that* $|x| \leqq 2T$. *Then*

$$\exp\Big(-\frac{1}{8T}\Big)\varphi(x,T) \leqq \mathcal{J}_1 \leqq \exp\Big(\frac{|x|}{4T}\Big)\varphi(x,T). \tag{2.1}$$

PROOF. Clearly

$$\mathcal{J}_1 = \int_{-1/2}^{1/2} \varphi(x+v,T)dv = \varphi(x,T) \int_{-1/2}^{1/2} \exp\Big(-\frac{xv}{T}\Big) \exp\Big(-\frac{v^2}{2T}\Big) dv,$$

$$\exp\Big(-\frac{1}{8T}\Big) \leqq \exp\Big(-\frac{v^2}{2T}\Big) \leqq 1 \qquad \Big(-\frac{1}{2} \leqq v \leqq \frac{1}{2}\Big)$$

and

$$\int_{-1/2}^{1/2} \exp\Big(-\frac{xv}{T}\Big) dv = \frac{T\Big(\exp\Big(\frac{x}{2T}\Big) - \exp\Big(-\frac{x}{2T}\Big)\Big)}{x}.$$

Since

$$1 \leqq \frac{T\Big(\exp\Big(\frac{x}{2T}\Big) - \exp\Big(-\frac{x}{2T}\Big)\Big)}{x} \leqq \exp\Big(\frac{|x|}{4T}\Big)$$

we have (2.1).

LEMMA 2. *Assume that*

$$1 \leqq |x| \leqq \frac{T}{\log T} \qquad \textit{and} \qquad 0 < t \leqq \frac{T}{\log T}.$$

Then

$$\begin{aligned} \mathcal{J}_2 \leqq{}& \exp\Big(\frac{2|x|}{T} + \frac{tx^2}{T^2}\Big)\Big(1 + \frac{t^2}{T^2}\Big)(\varphi(x,T))^2 + \\ &+ \Big(\frac{2t}{\pi}\Big)^{1/2} \frac{1}{|x|} \exp\Big(-\frac{|x|^2}{2t}\Big). \end{aligned} \tag{2.2}$$

PROOF. By (2.1)

$$\int_{-|x|}^{|x|} (\mathcal{J}_1(x-y,T-t))^2\varphi(y,t)dy \leqq$$

$$\leqq \int_{-|x|}^{|x|} \exp\Big(\frac{|x-y|}{2(T-t)}\Big)(\varphi(x-y,T-t))^2\varphi(y,t)dy \leqq$$

$$\leqq \exp\Big(\frac{2|x|}{T}\Big)\int_{-\infty}^{+\infty}(\varphi(x-y,T-t)^2\varphi(y,t)dy$$

where

$$\int_{-\infty}^{+\infty}(\varphi(x-y,T-t))^2\varphi(y,t)dy=$$

$$=\frac{(2\pi)^{-3/2}t^{-1/2}}{T-t}\int_{-\infty}^{+\infty}\exp\Big(-\frac{1}{2}\Big(\frac{2(x-y)^2}{T-t}+\frac{y^2}{t}\Big)\Big)dy=$$

$$=\frac{(2\pi)^{-3/2}t^{-1/2}}{T-t}\exp\Big(-\frac{x^2}{T+t}\Big)\int_{-\infty}^{+\infty}\exp\Big(-\frac{T+t}{2t(T-t)}\Big(y-2x\frac{t}{T+t}\Big)^2\Big)dy=$$

$$=(2\pi)^{-1}(T^2-t^2)^{-1/2}\exp\Big(-\frac{x^2}{T+t}\Big)=$$

$$=(\varphi(x,T))^2T(T^2-t^2)^{-1/2}\exp\Big(\frac{x^2}{T}-\frac{x^2}{T+t}\Big)\leqq$$

$$\leqq(\varphi(x,T))^2\Big(1+\frac{t^2}{T^2}\Big)\exp\Big(\frac{tx^2}{T^2}\Big).$$

Now we have

$$\mathcal{J}_2=\int_{-|x|}^{|x|}(\mathcal{J}_1(x-y,T-t))^2\varphi(y,t)dy+$$

$$+\int_{|y|\geqq|x|}(\mathcal{J}_1(x-y,T-t))^2\varphi(y,t)dy.$$

We also get

$$\int\limits_{|y|\geqq|x|} (\mathcal{J}_1(x-y,T-t))^2\varphi(y,t)dy \leqq$$

$$\leqq \int\limits_{|y|\geqq|x|} \varphi(y,t)dy = 2\Big(1-\Phi\Big(\frac{|x|}{t^{1/2}}\Big)\Big) \leqq$$

$$\leqq \Big(\frac{2t}{\pi}\Big)^{1/2}\frac{1}{|x|}\exp\Big(-\frac{|x|^2}{2t}\Big).$$

Hence we have (2.2).

3. The moments of $\lambda(x,T)$

Let $P_1=P_1(T)$, $P_2=P_2(T),\dots,P_{B(T)}=P_{B(T)}(T)$ be the locations of the $B(T)$ particles living at time T in an arbitrary but fixed order. Consider the paths (Wiener processes) $\{W_1(t),W_2(t),\dots,W_{B(T)}(t),\ 0\leqq t\leqq T\}$ of these particles, i.e. $W_i(T)=P_i(T)$ $(i=1,2,\dots,B(T)$. Define a partition $C_1,C_2,\dots,C_T$ of the set $\{P_1,P_2,\dots,P_{B(T)}\}$ as follows: $P_j\in C_i$ if

$$W_j(t)=W_1(t) \qquad \text{for} \quad t\leqq i$$

and for any $\varepsilon>0$ there exists a $0<\varepsilon_1=\varepsilon_1(\varepsilon)<\varepsilon$ such that

$$W_j(i+\varepsilon_1)\neq W_1(i+\varepsilon_1).$$

Finally let

$$I_i=I_i(T,x)=\begin{cases}1 & \text{if } P_i\in C(x),\\ 0 & \text{if } P_i\notin C(x)\end{cases}$$

and

$$|C_i|=c_i.$$

Note that

$$\mathbf{E}c_i=(m-1)m^{T-i} \qquad (i=1,2,\dots,T-1).$$

Since

$$\lambda(x,T)=\sum_{i=1}^{B(T)} I_i,$$

by Lemma 1 we have

LEMMA 3. *Assume that* $|x| \leqq 2T$. *Then*

$$\exp\Big(-\frac{1}{8T}\Big)\varphi(x,T)B(T) \leqq \mathcal{J}_1 B(T) =$$
$$= \mathbf{E}(\lambda(x,T) \mid B(T)) \leqq \exp\Big(\frac{|x|}{4T}\Big)\varphi(x,T)B(T).$$

Consequently,

$$\exp\Big(-\frac{1}{8T}\Big)\varphi(x,T)m^T \leqq \mathbf{E}\lambda(x,T) \leqq \exp\Big(\frac{|x|}{4T}\Big)\varphi(x,T)m^T.$$

LEMMA 4. *Assume that*

$$1 \leqq |x| \leqq \frac{T}{\log T}.$$

Then

$$\mathbf{E}\lambda^2(x,T) \leqq (2+\varepsilon)(m-1)m^{2T}\varphi^2(x,T) \tag{3.1}$$

for any $\varepsilon > 0$ *if* T *is big enough.*

PROOF. By Lemma 2

$$\mathbf{E}I_1 I_j = \mathcal{J}_2(x,T,k) \leqq$$
$$\leqq \exp\Big(\frac{2|x|}{T} + \frac{kx^2}{T^2}\Big)\Big(1+\frac{k^2}{T^2}\Big)(\varphi(x,T))^2 +$$
$$+ \Big(\frac{2k}{\pi}\Big)^{1/2}\frac{1}{|x|}\exp\Big(-\frac{x^2}{2k}\Big)$$

if

$$P_j \in C_k, \qquad k \leqq \frac{T}{\log T} \quad \text{and} \quad 1 \leqq |x| \leqq \frac{T}{\log T}.$$

Hence

$$\mathbf{E}\Big(I_1 \sum_{k=1}^{T/\log T} \sum_{P_j \in C_k} I_j\Big) =$$
$$= (m-1)m^T \sum_{k=1}^{T/\log T} m^{-k}\mathcal{J}_2(x,T,k) \leqq$$
$$\leqq (m-1)m^T(\varphi(x,T))^2 \sum_{k=1}^{T/\log T} \exp\Big(\frac{2|x|}{T} - k\Big(\log m - \frac{x^2}{T^2}\Big)\Big)\Big(1+\frac{k^2}{T^2}\Big) +$$

$$(m-1)m^T \frac{1}{|x|}\Big(\frac{2}{\pi}\Big)^{1/2} \sum_{k=1}^{T/\log T} k^{1/2} \exp\Big(-\frac{x^2}{2k}\Big) m^{-k} \leqq$$
$$\leqq (1+\varepsilon)(m-1)m^T(\varphi(x,T))^2$$

for any $\varepsilon > 0$ if T is big enough.

Since

$$\mathbf{E}\Big(I_1 \sum_{k=T/\log T}^{T} \sum_{P_j \in C_k} I_j\Big) \leqq$$
$$\leqq \mathbf{E}\Big(I_1 \sum_{k=T/\log T}^{T} c_k\Big) \leqq (m-1)m^T \exp\Big(\frac{|x|}{4T}\Big)\varphi(x,T) \sum_{k=T/\log T}^{T} m^{-k} \leqq$$
$$\leqq (1+\varepsilon)(m-1)m^T(\varphi(x,T))^2,$$

we have (3.1).

4. Proof of the Theorem

Let $y_1, y_2, \ldots, y_{B(t)}$ be the locations of the $B(t)$ particles living at time t.

Let $\lambda_i(x,T,t)$ be the number of those offsprings of the i-th particle which are located in x at time T. Clearly

$$\lambda(x,T) = \sum_{i=1}^{B(t)} \lambda_i(x,T,t).$$

LEMMA 5. *Let* $K > 2(\log m + 2)$ *and*

$$\Omega_1 = \Omega_1(t,K) = \{|y_i| \leqq Ks,\ \ i = 1,2,\ldots,B(s),\ \ s \geqq t\}.$$

Then

$$\mathbf{P}\{\Omega_1\} \geqq 1 - e^{-t} \tag{4.1}$$

for any $t = 1, 2, \ldots$.

PROOF.

$$\mathbf{P}\{\Omega - \Omega_1\} \leqq \mathbf{E}\Big(\sum_{s=t}^{\infty} B(s) \exp\Big(-\frac{Ks}{2}\Big)\Big) \leqq e^{-t}.$$

Hence we have (4.1).

Let

$$\mathcal{F}(t)=\mathcal{F}\{\lambda(x,s),\ x\in\mathbb{R}^1,\ s=0,1,2,\dots,t\}$$

be the smallest σ-algebra with respect to which the array

$$\{\lambda(x,s),\ x\in\mathbb{R}^1,\ s=0,1,2,\dots,t\}$$

is measurable.

LEMMA 6. *Let*

$$|x|\leqq\frac{T}{(\log T)^{1+\varepsilon}},$$
$$t\leqq A\log T \qquad (A>0),$$
$$K>2(\log m+2).$$

Then on the set $\Omega_1(t,K)$ *we have*

$$(1-\Delta)m^{T-t}\varphi(x,T)\leqq\mathbf{E}(\lambda_i(x,T,t)\mid\mathcal{F}(t))\leqq(1+\Delta)m^{T-t}\varphi(x,T), \tag{4.2}$$

where

$$\Delta=\Delta(A,K,x,T)=\frac{2KA|x|\log T+2K^2A^2(\log T)^2}{T}.$$

PROOF. Observe that on the set $\Omega_1(t,K)$

$$|x-y_i|\leqq\frac{T}{(\log T)^{1+\varepsilon}}+Kt<2(T-t).$$

Hence by Lemma 3

$$\mathbf{E}(\lambda_i(x,T,t)\mid\mathcal{F}(t))\leqq$$
$$\leqq\exp\Big(\frac{|x-y_i|}{4(T-t)}\Big)\varphi(x-y_i,T-t)m^{T-t}\leqq$$
$$\leqq\exp\Big(\frac{|x|+KA\log T}{T}\Big)\varphi(x,T)m^{T-t}\frac{\varphi(x-y_i,T-t)}{\varphi(x,T)}.$$

Since

$$\frac{\varphi(x-y_i,T-t)}{\varphi(x,T)}=\Big(\frac{T}{T-t}\Big)^{1/2}\exp\Big(-\frac{1}{2}\Big(\frac{(x-y_i)^2}{T-t}-\frac{x^2}{T}\Big)\Big)\leqq$$
$$\leqq\Big(\frac{T}{T-t}\Big)^{1/2}\exp\Big(\frac{x|y_i|}{T-t}+\frac{x^2t}{2T(T-t)}+\frac{y_i^2}{2(T-t)}\Big)\leqq$$
$$\leqq\Big(\frac{T}{T-t}\Big)^{1/2}\exp\Big(\frac{|x|KA\log T}{T-t}+\frac{x^2A\log T}{2T(T-t)}+\frac{K^2A^2(\log T)^2}{2(T-t)}\Big)\leqq$$
$$\leqq\Big(1+\frac{A\log T}{T}\Big)\exp\Big(\frac{2|x|KA\log T+K^2A^2(\log T)^2}{T}\Big)\leqq 1+\Delta,$$

we have the upper part of (4.2). The lower part can be seen in the same way.

LEMMA 7. *Under the conditions of Lemma 6 on the set* $\Omega_1(t,K)$ *we have*

$$\mathbf{E}(\lambda_i^2(x,T,t)\mid\mathcal{F}(t))\leqq(2+\varepsilon)(m-1)(1+2\Delta)m^{2(T-t)}\varphi^2(x,T)\leqq$$
$$\leqq 3(m-1)m^{2(T-t)}(\varphi(x,T))^2.$$

PROOF is the same as that of Lemma 6.

LEMMA 8. *Under the conditions of Lemma 6 on the set* $\Omega_1(t,K)$ *we have*

$$(1-\Delta)m^{T-t}B(t)\varphi(x,T)\leqq\mathbf{E}(\lambda(x,T)\mid\mathcal{F}(t))\leqq$$
$$\leqq(1+\Delta)m^{T-t}B(t)\varphi(x,T)$$

and

$$\operatorname{Var}(\lambda(x,T)\mid\mathcal{F}(t))\leqq 3(m-1)m^{2(T-t)}B(t)\varphi^2(x,T).$$

PROOF OF THE THEOREM. Let

$$\mathcal{K}=\{|\lambda(x,T)-\mathbf{E}(\lambda(x,T)\mid\mathcal{F}(t))|\geqq m^{t/4}(\operatorname{Var}(\lambda(x,T)\mid\mathcal{F}(t)))^{1/2}\}.$$

By Lemma 8

$$\mathcal{K}\supset\mathcal{L}\supset\overline{\mathcal{M}}$$

where

$$\mathcal{L}=\left\{|\lambda(x,T)-\mathbf{E}(\lambda(x,T)\mid\mathcal{F}(t))|\geqq m^T\varphi(x,T)\Big(3(m-1)\frac{B(t)}{m^{3t/2}}\Big)^{1/2}\right\},$$

$$\mathcal{M}=\left\{(1-\Delta)\frac{B(t)}{m^t}-\Big(3(m-1)\frac{B(t)}{m^t}\Big)^{1/2}m^{-t/4}\leqq\right.$$
$$\left.\leqq\frac{\lambda(x,T)}{m^T\varphi(x,T)}\leqq(1+\Delta)\frac{B(t)}{m^t}+\Big(3(m-1)\frac{B(t)}{m^t}\Big)^{1/2}m^{-t/4}\right\}$$

and $\overline{\mathcal{M}}$ is the complement of $\mathcal{M}$. By Chebyshev inequality on the set $\Omega_1(t,K)$ we have

$$\mathbf{P}\{\overline{\mathcal{M}}\mid\mathcal{F}(t)\}\leqq\mathbf{P}\{\mathcal{K}\mid\mathcal{F}(t)\}\leqq m^{-t/2}.$$

Hence

$$\mathbf{P}\left\{\Big|\frac{\lambda(x,T)}{m^T\varphi(x,T)}-\frac{B(t)}{m^t}\Big|\geqq\Delta\frac{B(t)}{m^t}+\Big(3(m-1)\frac{B(t)}{m^t}\Big)^{1/2}m^{-t/4}\right\}\leqq m^{-t/2}$$

and

$$\mathbf{P}\left\{\Big|\frac{T}{h(T)(\log T)^{1+\varepsilon/2}}\sup_{\substack{x\in\mathbb{Z}^d\\|x|\leqq h(T)}}\Big|\frac{\lambda(x,T)}{m^T\varphi(x,T)}-B\Big|\geqq(\log T)^{-\varepsilon/4}\right\}\leqq m^{-t/2}h(T)$$

if T is big enough. Choosing A of Lemma 6 big enough we have the Theorem.

REFERENCES

[1] BIGGINS, J. D., Uniform convergence of martingales in the branching random walk, *Ann. Probab.* **20** (1992), 137–151. *MR* **93b**:60094

[2] HARRIS, T. E., *The theory of branching processes*, Springer-Verlag, Berlin, 1963. *MR* **29** #664

[3] RÉVÉSZ, P., *Random walks of infinitely many particles*, World Scientific Publ. Co., Singapore, 1994.

(Received May 15, 1995)

TECHNISCHE UNIVERSITÄT WIEN
INSTITUT FÜR STATISTIK
WIEDNER HAUPTSTRASSE 8–10
A–1040 WIEN
AUSTRIA

revesz@ci.tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 251–262

ON LAST EXIT DECOMPOSITIONS OF LINEAR DIFFUSIONS

P. SALMINEN

Dedicated to Professor E. Csáki for his sixtieth birthday

Abstract

Let X be a regular one-dimensional diffusion living on an interval $I \subseteq \Re$. In this note we study the last exit decompositions of X at a fixed time t and at the life time. It is seen that these decompositions can be simply proved by using symmetry properties of X. Some further implications are also presented.

1. Introduction and notation

Let $X = \{X_t;\ t \geqq 0\}$ be a one-dimensional diffusion (in the sense of Itô and McKean [9]) living on an interval $I \subseteq \Re$. We assume that X is regular, i.e., $\mathbf{P}_x(H_y < \infty) > 0$ for every $x, y \in I$ where $H_y := \inf\{t\colon X_t = y\}$ and $\mathbf{P}_x$ is the probability measure associated to X when started from x. It is proved in Itô and McKean [9] p. 149 ff. (see also McKean [11]) using the theory of eigen-differential expansions that X has a transition density, denoted $p(t; x, y)$, $t > 0$, $x, y \in I$, with respect to its speed measure m, i.e., for every $t > 0$, $x \in I$ and $f \in \mathcal{B}_b(I)$ (:= the set of all bounded real-valued Borel-measurable functions on I)

$$\mathbf{E}_x(f(X_t)) = \int_I p(t; x, y) f(y) m(dy).$$

The function p is jointly continuous in all variables, non-negative, and, which is important here, symmetric in x and y, that is

$$p(t; x, y) = p(t; y, x) \quad \text{for all } x, y \in I.$$

The Green function is given by

$$g_\alpha(x, y) := \int_0^\infty e^{-\alpha t} p(t; x, y)\, dt.$$

1991 *Mathematics Subject Classification.* Primary 60J60; Secondary 60J25.

Key words and phrases. Last exit time, bridges, excessive transforms, scale function, speed measure, Green function.

0081–6906/97/$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

Using the fundamental increasing and decreasing solutions ψ and φ, respectively, the Green function can be expressed for $x \leqq y$ as

$$g_\alpha(x,y) = w_\alpha^{-1}\psi_\alpha(x)\varphi_\alpha(y)$$

where

$$w_\alpha := \psi_\alpha^+(x)\,\varphi_\alpha(x) - \psi_\alpha(x)\,\varphi_\alpha^+(x) = \psi_\alpha^-(x)\,\varphi_\alpha(x) - \psi_\alpha(x)\,\varphi_\alpha^-(x)$$

is a constant, so called Wronskian. Here, letting s denote the scale function, e.g.,

$$\varphi_\alpha^-(x) = \lim_{y\uparrow x} \frac{\varphi_\alpha^-(x) - \varphi_\alpha^-(y)}{s(x)-s(y)}, \quad \varphi_\alpha^+(x) = \lim_{y\downarrow x} \frac{\varphi_\alpha^-(y) - \varphi_\alpha^-(x)}{s(y)-s(x)}.$$

Recall also the usual normalization $w_0 = 1$. Further, p satisfies the Chapman–Kolmogorov equation for all $s < t$

$$p(t;x,y) = \int_I p(u;x,z)p(t-u;z,y)\,m(dz).$$

Existence and some regularity properties of p have been proved also by Rogers [14] for diffusions which can be obtained as solutions of stochastic differential equations with smooth coefficients.

Recall further (Itô and McKean [9] p. 154) that the $\mathbf{P}_y$-distribution of H_x has a density $n_y(t,x)$, $t>0$. From the eigen-differential expansion it is seen that $n_y(t,x)$ is continuous (at least) in t and y, non-negative, and satisfies for all $s<t$

$$n_y(t,x) = \int_I \hat{p}(u;y,z)n_z(t-u,x)\,m(dz), \tag{1}$$

where $\hat{p}$ is the transition density of the diffusion obtained from X by killing it at the time H_x.

For given x and $t>0$ introduce

$$G_t^x = \sup\{u<t\colon X_u = x\} \quad \text{and} \quad D_t^x = \inf\{u>t\colon X_u = x\}.$$

Then the celebrated *last exit decomposition* says that for $u<t<v$

$$\begin{aligned} &\mathbf{P}_x(G_t^x \in du,\ X_t \in dy,\ D_t^x \in dv) = \\ &= p(u;x,x)n_y(t-u,x)n_y(v-t,x)du\,dv\,m(dy). \end{aligned} \tag{2}$$

This is usually stated without D_t^x which enters into the formula by a standard application of the Markov property. Last exit decompositions are valid, of course, for very general Markov processes. In Getoor and Sharpe [5], [6] the result is proved for Hunt and standard processes. For continuous time

Markov chains, see Chung [3], and Williams [18]. In Maisonneuve [10] last exit decompositions are put into a general framework, called *exit system*, which shows connections to excursion theory. This is further developed in Getoor [4] and Getoor and Sharpe [8]. For Brownian motion see Chung [2].

In this note we prove (2) using time reversal. Although the approach is natural and intuitive we are not aware of any work where this is exploited. The idea is not, of course, new. In fact, in Williams [18] p. 222 this point of view is taken up for continuous time Markov chains. In spite of this we believe that it is worthwhile to present the following study mainly due to its simplicity.

Our approach leads also to complete characterizations of the laws of the processes $\{X_u:\ 0\leqq u\leqq G_t^x\}$, $\{X_u:\ G_t^x\leqq u\leqq t\}$, and $\{X_u:\ t\leqq u\leqq D_t^x\}$. Moreover, we prove, in the transient case, last exit decompositions at the life time. These provide a new derivation for the distribution of the last exit time; a result originally due to Pitman and Yor [12] in a special case, see also Getoor and Sharpe [7]. For other proofs see Salminen [15, 16].

2. Time reversal of diffusion bridges

We consider X in space-time, that is, we study the process $\bar{X}=\{(X_t,t):\ t\geqq 0\}$. Assume that $X_0=\alpha$ and let $\beta\in I$. Introduce

$$h(x,u):=h(x,u;\beta,t):=\begin{cases} p(t-u;x,\beta), & 0\leqq u<t,\ x\in I,\\ 0, & \text{otherwise.}\end{cases}$$

Using the Chapman–Kolmogorov equation it is seen that h is excessive for $\bar{X}$:

$$\mathbf{E}_{x,u}(h(X_v,v;\beta,t))=\begin{cases} h(x,u;\beta,t), & 0\leqq u<v<t,\\ 0, & \text{otherwise.}\end{cases}$$

For $f\in\mathcal{B}_b(I)$ and $u<v\wedge t$ let $P_{v,u}^h$ be the semigroup defined by

$$P_{v,u}^h f(x):=\mathbf{E}_x\Big(f(\omega(v-u))\frac{h(\omega(v-u),v)}{h(x,u)}\Big).$$

Here and in many cases below we consider X in its canonical framework and let $\omega:[0,\infty)\to I$ denote a generic element in the space of continuous functions. Let $(X^{\alpha,t,\beta},\mathbf{P}^{\alpha,t,\beta})$ be the strong non-time-homogeneous Markov process induced by $P_{v,u}^h$. We refer to Sharpe [17] p. 298 and (62.19) Theorem p. 296 where existence is proved in the general framework using multiplicative functionals. Because $0<h<\infty$ on $I\times[0,t)$ and equals zero elsewhere $X^{\alpha,t,\beta}$ lives on $I\times[0,t)$. Due to the symmetry of the transition density p the process $X^{\alpha,t,\beta}$ has a very clean time reversal property stated in the next proposition. The sign "$\sim$" means that the processes on the left- and the right-hand side are identical in law.

PROPOSITION 1.

$$\{X_{t-u}^{\alpha,t,\beta} : 0 < u < t\} \sim \{X_u^{\beta,t,\alpha} : 0 < u < t\}.$$

PROOF. For $0 < u_1 < \cdots < u_n < t$ we have

$$\begin{aligned}\mathbf{P}^{\alpha,t,\beta}&(\omega(t-u_1)\in dy_1,\dots,\omega(t-u_n)\in dy_n)\\ &= \mathbf{E}_\alpha(h(y_1,t-u_1;\beta,t);\omega(t-u_n)\in dy_n,\dots,\omega(t-u_1)\in dy_1)\\ &= p(t-u_n;\alpha,y_n)m(dy_n)p(u_n-u_{n-1};y_n,y_{n-1})m(dy_{n-1})\\ &\qquad \cdot\ldots\cdot p(u_2-u_1;y_2,y_1)m(dy_1)\frac{p(u_1;y_1,\beta)}{p(t;\alpha,\beta)}\\ &= p(u_1;\beta,y_1)m(dy_1)p(u_2-u_1;y_1,y_2)m(dy_2)\\ &\qquad \cdot\ldots\cdot p(u_n-u_{n-1};y_n,y_{n-1})m(dy_{n-1})\frac{p(t-u_n;y_n,\alpha)}{p(t;\alpha,\beta)}\\ &= \mathbf{P}^{\beta,t,\alpha}(\omega(u_1)\in dy_1,\dots,\omega(u_n)\in dy_n).\end{aligned}$$

□

Notice from the definition of $P_{v,u}^h$ that the finite dimensional distributions of $X^{\alpha,t,\beta}$ are obtained from the distributions of X by conditioning on $X_t = \beta$. Due to this and the next result the process $X^{\alpha,t,\beta}$ is called an *X-bridge from α to β having the length t.*

PROPOSITION 2. $\lim_{u\to t} X_u^{\alpha,t,\beta} = \beta$ *a.s.*.

PROOF. Clearly, by continuity of paths, $\lim_{u\to 0} X_u = \alpha$ a.s.. The measures $\mathbf{P}_\alpha$ and $\mathbf{P}^{\alpha,t,\beta}$ are equivalent when restricted to $\mathcal{F}_u := \sigma\{\omega(v) : 0 \leqq v \leqq u\}$, $u < t$, and, hence, a.s.

$$\lim_{u\to 0} X_u^{\alpha,t,\beta} = \alpha.$$

The claim follows now from Proposition 1. □

REMARKS. (a) It is assumed above that $\beta \in I$. However, it is also possible to take $\beta = r$, say, where r, the right-hand end point of I, is supposed not to be in I. Assume, furthermore, that $H_r < \infty$ with positive probability (strictly speaking, here $H_r := \inf\{u : X_{u-} = r\}$) and let

$$h(x,u) := \begin{cases} n_x(t-u;r), & 0 \leqq u < t,\ x \in I \\ 0, & \text{otherwise.}\end{cases}$$

From (1) it follows that h is space-time excessive for X, and we can construct $X^{\alpha,t,r}$. Intuitively, $X^{\alpha,t,r}$ is obtained from X by conditioning X to hit r at time t. Proposition 1 is true also in this case. The process $X^{r,t,\alpha}$ is governed by the measure obtained as a weak limit

$$\mathbf{P}^{r,t,\alpha} := \lim_{\beta\uparrow r} \mathbf{P}^{\beta,t,\alpha}.$$

Existence of the limit follows from the fact (Itô and McKean [9] p. 154)

$$n_x(t;r) = \lim_{y\uparrow r} \frac{p(t;x,y)}{s(r)-s(y)},$$

where s is the scale function of X. Notice that $s(r) < \infty$ because $H_r < \infty$ with positive probability.

(b) We give here a more informative construction of the process $X^{r,t,\alpha}$ introduced above. Let Z be the process obtained from X by conditioning X not to hit r. Then Z can be realized as an h-transform by taking

$$h(x) = \begin{cases} s(r)-s(x), & \text{if } H_r < \infty \text{ a.s.} \\ \mathbf{P}_x(H_r = \infty), & \text{otherwise.} \end{cases}$$

Hence, Z is a diffusion. It can be proved that r is an entrance-not-exit boundary point for Z. In particular, this means that Z can be started from r and it never returns there. The diffusion bridge $Z^{\alpha,t,\beta}$ can be constructed in the usual way, and straightforward computations show that

(3) $$X^{\alpha,t,\beta} \sim Z^{\alpha,t,\beta}.$$

Moreover, $Z^{r,t,\beta}$ is well defined and (3) is valid also for $\alpha = r$.

(c) The absolute continuity property of the measures $\mathbf{P}_\alpha$ and $\mathbf{P}^{\alpha,t,\beta}$ pointed out in the proof of Proposition 2 extends by standard arguments to be valid at stopping times. To formulate this let T be a stopping time and F_T an $\mathcal{F}_T$-measurable bounded random variable then

$$\mathbf{E}^{\alpha,t,\beta}(F_T; T<t) = \mathbf{E}_\alpha\Big(\frac{p(t-T;X_T,\alpha)}{p(t;\alpha,\beta)}\, F_T;\ T<t\Big).$$

3. Last exit decompositions and distributions

Let X be a regular diffusion and G_t^x the last exit time at x before a fixed time t. We start with by proving the *last exit decomposition* at a fixed time.

PROPOSITION 3. *For $t>0$*

$$\mathbf{P}_x(G_t^x \in du,\ X_t \in dy) = p(u;x,x)n_y(t-u,x)\,du\,m(dy).$$

PROOF. From Proposition 1 we obtain $u<t$

$$\begin{aligned}\mathbf{P}_x(G_t^x > u,\ X_t \in dy) &= \mathbf{P}_x(G_t^x > u \mid X_t = y)p(t;x,y)m(dy)\\ &= \mathbf{P}^{x,t,y}(G_t^x > u)p(t;x,y)m(dy)\\ &= \mathbf{P}^{y,t,x}(H_x < t-u)p(t;x,y)m(dy)\end{aligned}$$

$$= \mathbf{E}_y\Big(\frac{p(t-H_x;x,x)}{p(t;y,x)};\ H_x < t-u\Big)p(t;x,y)m(dy)$$

$$= m(dy)\int_0^{t-u} p(t-v;x,x)n_y(v,x)\,dv.$$

Differentiating with respect to u gives the claim. □

Let $\hat{X}$ be the diffusion obtained from X by killing at H_x and $\hat{X}^{x,u,y}$ the $\hat{X}$-bridge from x to y having the lenght $t-u$ (see Remarks (a) and (b) above). For Brownian motion the second statement in the next proposition can be found in Revuz and Yor [13] p. 454.

PROPOSITION 4. (a) *Conditionally on* $G_t^x = u$

$$\{X_s:\ 0 \leqq v < G_t^x\} \sim \{X_v^{x,u,x}:\ 0 \leqq v < u\}.$$

(b) *Conditionally on* $G_t^x = u$ *and* $X_t = y$

$$\{X_v:\ G_t^x \leqq v < t\} \sim \{\hat{X}_v^{x,t-u,y}:\ 0 \leqq v < t-u\}.$$

PROOF. (a) Consider for $0 < t_1 < \cdots < t_n$

$$\mathbf{P}_x(X_{t_1} \in dx_1, \ldots, X_{t_n} \in dx_n,\ t_n \leqq G_t^x,\ X_t \in I)$$

$$= \int_I \mathbf{P}^{x,t,y}(\omega(t_1) \in dx_1, \ldots, \omega(t_n) \in dx_n,\ t_n < G_t^x)p(t;x,y)m(dy)$$

$$= \int_I \mathbf{P}^{y,t,x}(\omega(t-t_n) \in dx_n, \ldots, \omega(t-t_1) \in dx_1,\ H_x < t-t_n)p(t;x,y)m(dy)$$

$$= \int_I m(dy)p(t;x,y) \int_0^{t-t_n} \mathbf{P}^{y,t,x}(H_x \in du)$$
$$\cdot \mathbf{P}^{y,t,x}_{x,u}(\omega(t-t_n) \in dx_n, \ldots, \omega(t-t_1) \in dx_1)$$

$$= \int_I m(dy)p(t;x,y) \int_0^{t-t_n} \mathbf{P}^{y,t,x}(H_x \in du)$$
$$\cdot \mathbf{P}^{x,t-u,x}(\omega(t-u-t_n) \in dx_n, \ldots, \omega(t-u-t_1) \in dx_1)$$

$$= \int_I m(dy)p(t;x,y) \int_0^{t-t_n} \mathbf{P}^{y,t,x}(H_x \in du)\mathbf{P}^{x,t-u,x}(\omega(t_1) \in dx_1, \ldots, \omega(t_n) \in dx_n)$$

$$= \int_I m(dy)p(t;x,y) \int_{t_n}^{t} \mathbf{P}^{x,t,y}(G_t^x \in du)\mathbf{P}^{x,u,x}(\omega(t_1) \in dx_1, \ldots, \omega(t_n) \in dx_n)$$

$$= \int_{t_n}^{t} \mathbf{P}_x(G_t^x \in du,\ X_t \in I)\mathbf{P}^{x,u,x}(\omega(t_1) \in dx_1, \dots, \omega(t_n) \in dx_n).$$

The second equality is based on Proposition 1, the third one on the strong Markov property, and the sixth again on Proposition 1. The claim (b) can be proved very much in the similar way, and we leave it to the reader. □

Assume now that X is transient and let ζ denote its life time. Define the last exit time at x:

$$G_\zeta^x := \sup\{u < \zeta :\ X_u = x\}.$$

Let k denote the killing measure of X, and recall the formula (see Itô and McKean [9] p. 184 or Borodin and Salminen [1])

$$\mathbf{P}_x(\zeta \in du,\ X_{\zeta-} \in A) = du \int_A p(u; x, y) k(dy), \quad A \in \mathcal{B}(I). \tag{4}$$

Below we give *last exit decompositions at* ζ. In the first one X dies inside the state space I, in the second one X dies at a killing boundary point or drifts toward a boundary point, and finally these results are combined to give the distribution of G_ζ^x.

PROPOSITION 5. *For* $y \in I$

$$\mathbf{P}_x(G_\zeta^x \in du,\ X_{\zeta-} \in dy) = p(u; x, x)\mathbf{P}_y(H_x < \infty)\, du\, k(dy).$$

PROOF. From (4) it is seen that X conditioned to have $\zeta = u$ and $X_{\zeta-} = y \in I$ can be realized as $X^{x,u,y}$. To prove the claim, we compute for $t > 0$ and $y \in I$ as follows

$$\begin{aligned}
\mathbf{P}_x(G_\zeta^x > t, X_{\zeta-} \in dy) &= \int_{v=t}^{\infty} \int_{u=v}^{\infty} \mathbf{P}_x(G_\zeta^x \in dv,\ \zeta \in du,\ X_{\zeta-} \in dy) \\
&= \int_{v=t}^{\infty} \int_{u=v}^{\infty} \mathbf{P}^{x,u,y}(G_u^x \in dv)\mathbf{P}_x(\zeta \in du, X_{\zeta-} \in dy) \\
&= \int_{u=t}^{\infty} \int_{v=t}^{u} \mathbf{P}^{x,u,y}(G_u^x \in dv)\, \mathbf{P}_x(\zeta \in du,\ X_{\zeta-} \in dy) \\
&= \int_{t}^{\infty} \mathbf{P}^{y,u,x}(H_x < u - t)\, \mathbf{P}_x(\zeta \in du,\ X_{\zeta-} \in dy)
\end{aligned}$$

$$= \int_t^\infty \mathbf{E}_y\Big(\frac{p(u-H_x;x,x)}{p(u;y,x)}; H_x<u-t\Big)\mathbf{P}_x(\zeta\in du, X_{\zeta-}\in dy)$$

$$= k(dy)\int_t^\infty du \int_0^{u-t} dv\, p(u-v;x,x)\, n_y(v,x)$$

$$= \mathbf{P}_y(H_x<\infty)k(dy)\int_t^\infty p(u;x,x)du. \qquad \square$$

Before proceeding recall that for $y \leqq x$

(5) $$\mathbf{P}_y(H_x<\infty) = \lim_{\alpha\downarrow 0}\mathbf{E}_y(e^{-\alpha H_x}) = \lim_{\alpha\downarrow 0}\frac{g_\alpha(x,y)}{g_\alpha(x,x)} = \frac{g_0(x,y)}{g_0(x,x)}.$$

PROPOSITION 6. (a) *Assume that the right-hand end point r is not in I, and that $H_r<\infty$ with positive probability or $H_r=\infty$ a.s. and $\lim_{u\to\zeta} X_u = r$ with positive probability. Then*

$$\mathbf{P}_x(G_\zeta^x\in ds,\ \lim_{u\to\zeta} X_u = r) = \frac{p(s;x,x)}{\psi_0(r)\varphi_0(x)}\,ds.$$

(b) *For the left-hand end point l we have similarly*

$$\mathbf{P}_x(G_\zeta^x\in ds,\ \lim_{u\to\zeta} X_u = l) = \frac{p(s;x,x)}{\varphi_0(l)\psi_0(x)}\,ds.$$

PROOF. We prove (a); the proof of (b) is similar. Consider first the case $H_r<\infty$ with probability 1. Let $\hat{\mathbf{P}}$ be the measure associated with the diffusion Z introduced in Remark (b) above. The speed and the scale measure of Z (see Borodin and Salminen [1]) are

(6) $$\hat{m}(dx) = h^2(x)m(dx) \quad \text{and} \quad \hat{s}(dx) = h^{-2}(x)m(dx),$$

respectively, where $h(x) := \varphi_0(x) = s(r) - s(x)$. Notice also that in this case $\psi_0 \equiv 1$. The transition density $\hat{p}$ with respect to m^h is given by

(7) $$\hat{p}(u;x,y) = \frac{p(u;x,y)}{h(x)h(y)},$$

and

$$\hat{p}(u;r,y) = \lim_{x\uparrow r}\frac{p(u;x,y)}{h(x)h(y)} = \frac{n_y(u;r)}{h(y)}.$$

Then, because $\zeta = H_r$,

$$\begin{aligned}
\mathbf{P}_x(G^x_\zeta > v,\ \lim_{u\to\zeta} X_u = r) &= \int_{s=v}^{\infty}\int_{u=s}^{\infty} \mathbf{P}_x(G^x_\zeta \in ds,\ \zeta\in du,\ \lim_{t\to\zeta} X_t = r) \\
&= \int_{s=v}^{\infty}\int_{u=s}^{\infty} \mathbf{P}^{x,u,r}(G^x_u \in ds)\,\mathbf{P}_x(\zeta\in du,\ \lim_{t\to\zeta} X_t = r) \\
&= \int_{u=v}^{\infty}\int_{s=v}^{u} \mathbf{P}^{x,u,r}(G^x_u \in ds)\, n_x(u,r)\,du \\
&= \int_{v}^{\infty} \hat{\mathbf{P}}^{r,u,x}(H_x < u-v)\, n_x(u,r)\,du \\
&= \int_{v}^{\infty} \hat{\mathbf{E}}_r\Big(\frac{\hat p(u-H_x;x,x)}{\hat p(u;r,x)};\ H_x < u-v\Big)\, n_x(u,r)\,du \\
&= \int_{u=v}^{\infty}\int_{s=0}^{u-v} \frac{\hat p(u-s;x,x)}{\hat p(u;r,x)}\hat{\mathbf{P}}_r(H_x\in ds)\, n_x(u,r)\,du \\
&= \int_{u=v}^{\infty}\int_{s=0}^{u-v} \frac{\hat p(u-s;x,x)}{\hat p(u;r,x)}\lim_{y\uparrow r}\hat{\mathbf{P}}_y(H_x\in ds)\, n_x(u,r)\,du.
\end{aligned}$$

Using the formulae for the transition density $\hat p$ given above, absolute continuity, and changing the order of integration give

$$\begin{aligned}
\mathbf{P}_x(G^x_\zeta > v,\ \lim_{u\to\zeta} X_u = r) &= \lim_{y\uparrow r}\frac{\mathbf{P}_y(H_x<\infty)}{s(r)-s(y)}\int_v^\infty p(u;x,x)\,du \\
&= \lim_{y\uparrow r}\frac{\varphi_0(y)}{(s(r)-s(y))\varphi_0(x)}\int_v^\infty p(u;x,x)\,du.
\end{aligned}$$

The second equality follows from (5). The proof of the first special case is now complete. Next assume that $0 < \mathbf{P}_x(H_r<\infty) < 1$ for every $x\in I$, and introduce

$$h^*(x) := \mathbf{P}_x(H_r<\infty) = \frac{\psi_0(x)}{\psi_0(r)}.$$

Then X conditioned by $\{H_r<\infty\}$ is a diffusion and can be realized as an h^*-transform of X. The speed and scale measure, m^* and s^*, respectively, are as in (6) and the transition density, p^*, as in (7) with h^* instead of h. Furthermore, for $x \leqq y$

$$\int_0^\infty p^*(t; x, y)\, dt = \frac{\psi_0(x)}{h(x)} \frac{\varphi_0(y)}{h(y)} = \psi_0^*(x)\, \varphi_0^*(y)$$

giving

$$\varphi_0^*(x) = s^*(r) - s^*(x) = \psi_0^2(r) \frac{\varphi_0(x)}{\psi_0(x)}.$$

Applying now the result in the special case above to the conditioned process we obtain

$$\begin{aligned}\mathbf{P}_x(G_\zeta^x \in dt,\ \lim_{u\to\zeta} X_u = r) &= \mathbf{P}_x(\lim_{u\to\zeta} X_u = r)\, \mathbf{P}_x(G_\zeta^x \in dt |\lim_{u\to\zeta} X_u = r) \\ &= h^*(x) p^*(t, x, x) \frac{1}{\varphi_0^*(x)}\, dt = \frac{p(t; x, x)}{\psi_0(r)\varphi_0(x)}\, dt\end{aligned}$$

as claimed. Next assume that X drifts to r with probability one, that is, a.s. $H_r = \infty$ and $\lim\limits_{u\to\zeta} X_u = r$. Then, of course, also a.s. $\zeta = \infty$. Moreover, $\psi_0 \equiv 1$, and $\varphi_0 = s(r) - s(\cdot)$. Let $\{r(n)\}$ be a sequence increasing to r as $n \to \infty$. For $\gamma > 0$ we have

$$\mathbf{E}_x(e^{-\gamma G_\zeta^x}) = \mathbf{E}_x(e^{-\gamma G_\zeta^x};\ \lim_{t\to\zeta} X_t = r) = \lim_{n\to\infty} \mathbf{E}_x(e^{-\gamma G_{\zeta(n)}^x};\ X_{\zeta(n)} = r(n)),$$

where $\zeta(n) = \zeta \wedge H_{r(n)} = H_{r(n)}$. Let $\varphi^{(n)}$ and $\psi^{(n)}$ be the fundamental decreasing and increasing, respectively, solutions associated to X when stopped at $\zeta(n)$:

$$\begin{aligned}\varphi_\gamma^{(n)}(x) &= \varphi_\gamma(x) - \frac{\varphi_\gamma(r(n))}{\psi_\gamma(r(n))} \psi_\gamma(x), \\ \psi_\gamma^{(n)}(x) &= \psi_\gamma(x),\end{aligned}$$

and especially for $\gamma = 0$

$$\begin{aligned}\varphi_0^{(n)}(x) &= \varphi_0(x) - \varphi_0(r(n)) = s(r(n)) - s(x), \\ \psi_0^{(n)}(x) &= \psi_0(x) = 1.\end{aligned}$$

Hence, using the first special case above and the fact that $\lim\limits_{n\to\infty} \psi_\alpha(r(n)) = \infty$ when $\alpha > 0$ (see Itô and McKean [9] or Borodin and Salminen [1]) we obtain

$$\mathbf{E}_x(e^{-\gamma G_\zeta^x};\ \lim_{t\to\zeta} X_t = r) = \lim_{n\to\infty} G_\gamma^{(n)}(x, x) \frac{1}{\varphi_0^{(n)}(x)} = G_\gamma(x, x) \frac{1}{\varphi_0(x)}$$

proving the claim. When $0 < \mathbf{P}_x(\lim_{t\to\zeta} X_t = r) < 1$ we can proceed similarly as in the second special case above, and we leave the details to the reader. □

PROPOSITION 7.

$$\mathbf{P}_x(G_\zeta^x \in dt) = \frac{p(t;x,x)}{g_0(x,x)}\,dt = \frac{p(t;x,x)}{\varphi_0(x)\psi_0(x)}\,dt.$$

PROOF. Because $\mathbf{P}_x(G_\zeta^x < \infty) = 1$ and

$$\int_0^\infty p(t;x,x)\,dt = g_0(x,x) = \varphi_0(x)\psi_0(x)$$

the claim follows from Propositions 5 and 6. □

REMARK. From Propositions 5, 6, and 7 we obtain under the made assumptions on the boundary behaviour the following identity for the Green function:

$$\frac{1}{g_0(x,x)} = \frac{1}{\psi_0(r)\varphi_0(x)} + \int_I \frac{g_0(x,y)}{g_0(x,x)}k(dy) + \frac{1}{\varphi_0(l)\psi_0(x)}.$$

To explain this analytically notice that

$$\begin{aligned}\int_I \frac{g_0(x,y)}{g_0(x,x)}k(dy) &= \frac{1}{\psi_0(x)}\int_l^x \psi_0(y)\,k(dy) + \frac{1}{\varphi_0(x)}\int_x^r \varphi_0(y)\,k(dy)\\ &= \frac{\psi_0^-(x)}{\psi_0(x)} - \frac{\psi_0^-(l+)}{\psi_0(x)} + \frac{\varphi_0^-(r-)}{\varphi_0(x)} - \frac{\varphi_0^-(x)}{\varphi_0(x)}\\ &= \frac{1}{g_0(x,x)} + \frac{\varphi_0^-(r-)}{\varphi_0(x)} - \frac{\psi_0^-(l+)}{\psi_0(x)}.\end{aligned}$$

We have used here the fact that ψ_0 and φ_0 satisfy a generalized differential equation (see Itô and McKean [9]); it is also assumed that k does not charge x. Consequently, we must have

$$w_0 = 1 = \varphi_0(l)\psi_0^-(l+) = \psi_0(r)\varphi_0^-(r-)$$

or, equivalently,

$$\varphi_0(r)\psi_0^-(r-) = \psi_0(l)\varphi_0^-(l+) = 0$$

which is not, perhaps, obvious when r, say, is exit-not-entrance or natural.

REFERENCES

[1] BORODIN, A. N. and SALMINEN, P., *Handbook of Brownian motion — facts and formulae*, Birkhäuser, Basel–Boston–Berlin, 1996.

[2] Chung, K.-L., Excursions in Brownian motion, *Ark. Mat.* **14** (1976), 155–177. *MR* **57** #7791

[3] Chung, K.-L., *Lectures on boundary theory for Markov chains*, Annals of Mathematics Studies, No. 65, Princeton University Press, Princeton, NJ, 1970. *MR* **42** #2546

[4] Getoor, R. K., Excursions of a Markov process, *Ann. Probability* **7** (1979), 244–266. *MR* **80j**:60103

[5] Getoor, R. K. and Sharpe, M. J., Last exit times and additive functionals, *Ann. Probability* **1** (1973), 550–569. *MR* **50** #5951

[6] Getoor, R. K. and Sharpe, M. J., Last exit decompositions and distributions, *Indiana Univ. Math. J.* **23** (1973), 377–404. *MR* **48** #12654

[7] Getoor, R. K. and Sharpe, M., Excursions of Brownian motion and Bessel processes, *Z. Wahrsch. Verw. Gebiete* **47** (1979), 83–106. *MR* **80b**:60104

[8] Getoor, R. K. and Sharpe, M., Excursions of dual processes, *Adv. in Math.* **45** (1982), 259–309. *MR* **84h**:60129

[9] Itô, K. and McKean, H. P., *Diffusion processes and their sample paths*, Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 125, Springer-Verlag, Berlin–New York, 1974. *MR* **49** #9963

[10] Maisonneuve, B., Exit systems, *Ann. Probability* **3** (1975), 399–411. *MR* **53** #4251

[11] McKean, H. P., Elementary solutions for certain parabolic partial differential equations, *Trans. Amer. Math. Soc.* **82** (1956), 519–548. *MR* **19**, 285b

[12] Pitman, J. and Yor, M., Bessel processes and infinitely divisible laws, *Stochastic integrals* (Proc. Sympos. Univ. Durham, Durham, 1980), Lecture notes in math., 851, Springer-Verlag, Berlin–Heidelberg, 1981, 285–370. *MR* **82j**:60149

[13] Revuz, D. and Yor, M., *Continuous martingales and Brownian motion,* Grundlehren der mathematischen Wissenschaften, 293, Springer-Verlag, Berlin–Heidelberg–New York, 1991. *MR* **92d**:60053

[14] Rogers, L. C. G., Smooth transition densities for one-dimensional diffusions, *Bull. London Math. Soc.* **17** (1985), 157–161. *MR* **87a**:60087

[15] Salminen, P., One-dimensional diffusions and their exit spaces, *Math. Scand.* **54** (1984), 209–220. *MR* **86g**:60091

[16] Salminen, P., On local times of a diffusion, *Séminaire de Probabilités XIX.* 1983/84, eds. Azéma, J. and Yor, M., Lecture notes in Mathematics, 1123, Springer-Verlag, Berlin–Heidelberg, 1985, 63–79. *MR* **88h**:60152

[17] Sharpe, M., *General theory of Markov processes*, Pure and applied mathematics, 133, Academic Press, Boston, MA, 1988. *MR* **89m**:60169

[18] Williams, D., *Diffusions, Markov processes, and martingales, Vol. I: Foundations*, Probability and mathematical statistics, John Wiley and Sons, Chichester–New York, 1979. *MR* **80i**:60100

(Received September 8, 1996)

ÅBO AKADEMI
MATHEMATICAL DEPARTMENT
DOMKYRKOTORGET 3
FIN–20500 ÅBO
FINLAND

paavo.salminen@abo.fi

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 263–277

DISTRIBUTIONS BASED ON SMIRNOV ONE-SIDED AND RELATED RANK ORDER STATISTICS

JAGDISH SARAN and M. K. SUKLA

Dedicated to Professor E. Csáki for his sixtieth birthday

Abstract

This paper deals with the two-sample problem and investigates the joint and marginal distributions of D^+_{mn}, the Smirnov one-sided statistic, $R^+_{mn}(i)$, the index where D^+_{mn} is achieved for the i^{th} time, $M^+_{mn}(i,j)$, the length of the interval between the i^{th} and the j^{th} maxima $(1 \leqq i \leqq j)$ and Q^+_{mn}, the number of times D^+_{mn} is achieved.

1. Introduction

Let $X_{(1)} < X_{(2)} < \cdots < X_{(m)}$ and $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ be the order statistics from two independent samples of i.i.d. random variables having continuous distribution functions F and G, respectively, and let $F_m(x)$ and $G_n(x)$ be the corresponding empirical distribution functions. Let $Z_1 < Z_2 < \cdots < Z_{m+n}$ denote the ordered combined sample and let R_i denote the rank of $X_{(i)}$ in the ordered combined sample. The statistical problem in question is to ascertain whether or not two samples are from the same population (i.e., $F = G$), and thus it is important to derive probability distributions of various statistics when $H_0 : F = G$ is true. The Smirnov one-sided statistic is given by

$$D^+_{mn} = \sup_t \{F_m(t) - G_n(t)\} = \left(\frac{1}{mn}\right) \max_{1 \leqq k \leqq m} (k(m+n) - mR_k).$$

This follows from Maag and Stephens [3] and also from Steck [7]. If $mnD^+_{mn} = d$, let $R^+_{mn}(i)$ denote the index of the point where $k(m+n) - mR_k = d$ for the i^{th} time $(i = 1, 2, \ldots)$, i.e., $R^+_{mn}(i)$ denotes the index where D^+_{mn} is attained for the i^{th} time. Let $M^+_{mn}(i,j)$ denote the length of the interval between the i^{th} and the j^{th} maximum $(1 \leqq i \leqq j)$, i.e.,

$$M^+_{mn}(i,j) = R^+_{mn}(j) - R^+_{mn}(i)$$

and let Q^+_{mn} denote the number of times D^+_{mn} is attained.

1990 *Mathematics Subject Classifications.* Primary 62G30.

Key words and phrases. Two-sample problem, lattice path, Smirnov one-sided statistic, index of the i^{th} maxima, length of the interval between the i^{th} and the j^{th} maxima, number of times the maxima is achieved, Steck determinants.

0081–6906/97/\$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

The distributions of $R^+_{mn}(i)$, D^+_{mn} and Q^+_{mn} have been discussed by several authors viz., Vincze [9], Sarkadi [6], Steck [7], Geller [2], Steck and Simmons [8], Saran and Sen [4] and Saran and Rani [5] in certain special cases. In this paper it is proposed to investigate, for finite m and n, the joint and marginal distributions of D^+_{mn}, $R^+_{mn}(i)$, $M^+_{mn}(i,j)$ and Q^+_{mn}, for $1 \leqq i \leqq j$, under the hypothesis $F = G$.

2. Path representation

The $(m+n)$ observations in the ordered combined sample $Z_1 < Z_2 < \cdots < Z_{m+n}$ are represented by a minimal lattice path from $(0,0)$ to (n,m) with the k^{th} step being one unit up or one unit to the right according as Z_k is an X or a Y observation, respectively. It can be observed that after the k^{th} step up, the path is at the point $(R_k - k, k)$ and that $k(m+n) - mR_k$ is m times the horizontal distance from $(R_k - k, k)$ to the diagonal $y = mx/n$. Thus mnD^+_{mn} is m times the maximum horizontal distance from the path to the diagonal $y = mx/n$. In the sequel we shall use the word 'distance' to denote 'horizontal distance'. Distances to the diagonal will be taken positive if the point is to the left of the diagonal and negative otherwise.

The results obtained in Sections 4 and 5 also have an interpretation in terms of the ballot problem. In that context A scores m votes, B scores n votes, all possible vote sequences are $\binom{m+n}{n}$ in number, and a path is interpreted as a vote sequence.

3. Some auxiliary results

The following two results, needed in the sequel, are quoted from Bizley [1] and Steck [7], respectively.

LEMMA 3.1. *Let p be the greatest common divisor (g.c.d.) of sample sizes m and n, i.e., $m = ap$ and $n = bp$ where a and b are coprime positive integers. Then the number of minimal lattice paths from $(0,0)$ to (kb, ka) having just t contacts with the line $y = mx/n$ (not counting $(0,0)$) and having no points above this line (where k is a positive integer) is given by $\phi_{k,t}$ where*

(3.1) $\phi_{k,t} = \text{coeff. of } y^k \text{ in the expansion of } \{1 - \exp(-A_1 y - A_2 y^2 - \dots)\}^t$

and

$$A_j = \frac{1}{j(a+b)}\binom{ja+jb}{ja}. \tag{3.2}$$

LEMMA 3.2. *Let* $b_1 \leqq b_2 \leqq \cdots \leqq b_m$ *and* $c_1 \leqq c_2 \leqq \cdots \leqq c_m$ *be sequences of integers such that* $i \leqq b_i \leqq c_i \leqq n+i$, $i = 1, 2, \ldots, m$. *Then*

$$(3.3) \qquad \binom{m+n}{n}\mathbf{P}[b_i \leqq R_i \leqq c_i, \text{ all } i] = \det\left[\binom{c_i - b_j + j - i + 1}{j - i + 1}_+\right]_{m\times m},$$

where

$$\binom{x}{r}_+ = \binom{\max(x,0)}{r} \quad \textit{and} \quad j = 1, 2, \ldots, m.$$

4. The joint distribution of $D^+_{mn}, R^+_{mn}(i), M^+_{mn}(i,j)$ and Q^+_{mn}

LEMMA 4.1. *The number of paths from* $(0,0)$ *to* (n,m) *through the points* (x_1, y_1), (x_2, y_2), (x_3, y_3) *and* (x_4, y_4), $x_i \leqq ny_i/m$, $i = 1, 2, 3, 4$; $x_1 \leqq x_2 \leqq x_3 \leqq x_4$, $y_1 \leqq y_2 \leqq y_3 \leqq y_4$, *that attain their maximum distance from the diagonal* $y = mx/n$ *for the first,* $i^{\text{th}}, j^{\text{th}}$ *and* l^{th} *(i.e., the last) time* $(1 \leqq i \leqq j \leqq l)$ *at* $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ *and* (x_4, y_4), *respectively, is the same as the number of paths from* $(0,0)$ *to* (n,m) *through the points* $(x_2 - x_1, y_2 - y_1)$, $(x_3 - x_1, y_3 - y_1), (x_4 - x_1, y_4 - y_1)$ *and* $(n - x_1, m - y_1)$ *that are never above the diagonal and, moreover, never touch the diagonal after* $(x_4 - x_1, y_4 - y_1)$ *except at* (n,m) *and having exactly* $(l-1)$ *contacts with the diagonal up to the point* $(x_4 - x_1, y_4 - y_1)$ *of which the first* $(i-1)$ *contacts occur up to the point* $(x_2 - x_1, y_2 - y_1)$ *and the next* $(j-i)$ *contacts occur up to the point* $(x_3 - x_1, y_3 - y_1)$.

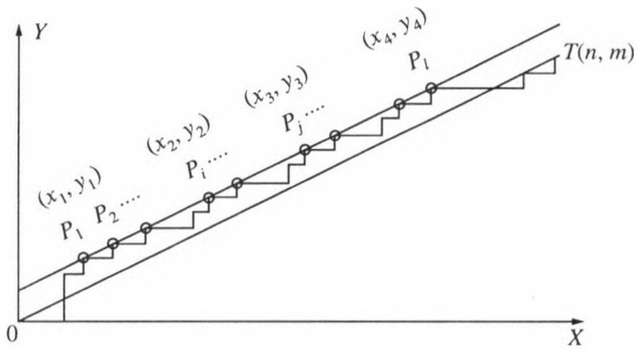


Fig. 1

PROOF. Let $OP_1P_2 \ldots P_i \ldots P_j \ldots P_lT$ (Fig. 1) be a lattice path from $(0,0)$ to (n,m) which attains its maximum distance from the diagonal $y = mx/n$ for the first, i^{th}, j^{th} and the l^{th} (i.e., the last) time at $P_1(x_1, y_1)$, $P_i(x_2, y_2)$, $P_j(x_3, y_3)$ and $P_l(x_4, y_4)$, respectively. Now we apply the following transformation to this path. The path segment OP_l is shifted up $(m-y_1)$ units and shifted right $(n-x_1)$ units. Then OP_l is transformed to a path from $(n-x_1, m-y_1)$ to (n,m) remaining entirely below the line $y=mx/n$ never touching it in-between (as shown by TP_l in Fig. 2). Similarly, let the path segment $P_1P_2\ldots P_i\ldots P_j\ldots P_lT$ of Fig. 1 be shifted down y_1 units and shifted left x_1 units. Then it is transformed to a path from $(0,0)$ to $(n-x_1, m-y_1)$ not rising above the line $y=mx/n$ and passing through the points $(x_2-x_1, y_2-y_1), (x_3-x_1, y_3-y_1)$ and (x_4-x_1, y_4-y_1), each point lying on the diagonal $y=mx/n$, and having in all $(l-1)$ contacts with the diagonal $y=mx/n$; the $(l-1)^{\text{st}}$ contact occurring at the point (x_4-x_1, y_4-y_1), the $(j-1)^{\text{st}}$ contact at the point (x_3-x_1, y_3-y_1) and the $(i-1)^{\text{st}}$ contact at the point (x_2-x_1, y_2-y_1) (as shown by $OP_2\ldots P_i\ldots P_j\ldots P_lT$ in Fig. 2). Thus the complete transformed path $OP_2\ldots P_i\ldots P_j\ldots P_lTP_l$ (Fig. 2) is in one-to-one correspondence with the original path in Fig. 1. This proves Lemma 4.1.

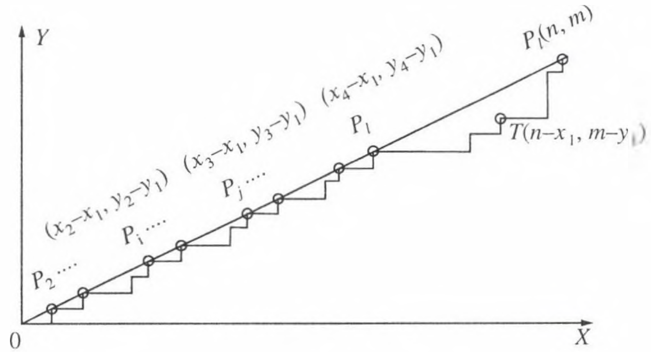


Fig. 2

To derive the joint distribution of $D^+_{mn}, R^+_{mn}(i), M^+_{mn}(i,j)$ and Q^+_{mn} we first obtain an expression for the probability

$$\mathbf{P}[mnD^+_{mn}=d,\ R^+_{mn}(1)=r,\ R^+_{mn}(i)=s,\ M^+_{mn}(i,j)=t,\ R^+_{mn}(l)=u,\ Q^+_{mn}=l]$$
$$=\mathbf{P}[mnD^+_{mn}=d, R^+_{mn}(1)=r, R^+_{mn}(i)=s, R^+_{mn}(j)=s+t, R^+_{mn}(l)=u, Q^+_{mn}=l]$$

for $r\leqq s\leqq u$, $0\leqq t\leqq u-s$, $1\leqq i\leqq j\leqq l$ and $d>0$. For this we consider a path from $(0,0)$ to (n,m) through the points (x_1,y_1), (x_2,y_2), (x_3,y_3) and (x_4,y_4), $x_i\leqq ny_i/m$, $i=1,2,3,4$; $x_1\leqq x_2\leqq x_3\leqq x_4$, $y_1\leqq y_2\leqq y_3\leqq y_4$ that attains its maximum distance from the diagonal $y=mx/n$ for the first, i^{th}, j^{th} and l^{th} (i.e., the last) time at $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ and (x_4,y_4), respectively (as in Fig. 1). This corresponds to a path for which $mnD^+_{mn}=ny_1-mx_1=ny_2-mx_2=ny_3-mx_3=ny_4-mx_4$, $R^+_{mn}(1)=x_1+y_1$, $R^+_{mn}(i)=x_2+y_2$, $R^+_{mn}(j)=x_3+y_3$ and $R^+_{mn}(l)=x_4+y_4$. By Lemma 4.1 and Fig. 2, the number of such paths is equal to the product of the following five factors, viz.

$F_1 =$ the number of paths from $(0,0)$ to $(x_2 - x_1, y_2 - y_1)$ that are never above the diagonal $y = mx/n$ and having exactly $(i-1)$ contacts with the diagonal, the $(i-1)^{\text{st}}$ contact occurring at $(x_2 - x_1, y_2 - y_1)$,

$F_2 =$ the number of paths from $(x_2 - x_1, y_2 - y_1)$ to $(x_3 - x_1, y_3 - y_1)$ that are never above the diagonal $y = mx/n$ and having exactly $(j-i)$ contacts with the diagonal, the $(j-i)^{\text{th}}$ contact occurring at $(x_3 - x_1, y_3 - y_1)$,

$F_3 =$ the number of paths from $(x_3 - x_1, y_3 - y_1)$ to $(x_4 - x_1, y_4 - y_1)$ that are never above the diagonal $y = mx/n$ and having exactly $(l-j)$ contacts with the diagonal, the $(l-j)^{\text{th}}$ contact occurring at $(x_4 - x_1, y_4 - y_1)$,

$F_4(x_4, y_4) =$ the number of paths from $(x_4 - x_1, y_4 - y_1)$ to $(n - x_1, m - y_1)$ that remain entirely below the diagonal $y = mx/n$ and never touch it in-between except at the initial point $(x_4 - x_1, y_4 - y_1)$,

and

$F_5(x_1, y_1) =$ the number of paths from $(n - x_1, m - y_1)$ to (n, m) that remain entirely below the line $y = mx/n$ and never touch it in-between except at the end point (n, m).

As each one of the points $(x_2 - x_1, y_2 - y_1)$, $(x_3 - x_1, y_3 - y_1)$ and $(x_4 - x_1, y_4 - y_1)$ lies on the diagonal $y = mx/n$, we suppose that $(x_2 - x_1, y_2 - y_1) \equiv (\lambda b, \lambda a)$, $(x_3 - x_1, y_3 - y_1) \equiv (\mu b, \mu a)$ and $(x_4 - x_1, y_4 - y_1) \equiv (\delta b, \delta a)$, where $\lambda = [(x_2 + y_2) - (x_1 + y_1)]/(a+b)$, $\mu = [(x_3 + y_3) - (x_1 + y_1)]/(a+b)$ and $\delta = [(x_4 + y_4) - (x_1 + y_1)]/(a+b)$ are all integers such that $\lambda \leqq \mu \leqq \delta$, $\lambda \geqq i-1$, $\mu \geqq j-1$ and $\delta \geqq l-1$. Then on using Lemma 3.1, we have

$$F_1 = \phi_{\lambda, i-1} \tag{4.1}$$

and on taking $(x_2 - x_1, y_2 - y_1)$ and $(x_3 - x_1, y_3 - y_1)$ as new origins, we have

$$F_2 = \phi_{\mu-\lambda, j-i} \tag{4.2}$$

and

$$F_3 = \phi_{\delta-\mu, l-j}, \tag{4.3}$$

where $\phi_{k,t}$ is given by (3.1). In what follows we shall use the symbols $\{x\}$ and $\langle x \rangle$ to denote, respectively, the smallest integer greater than x and the smallest integer greater than or equal to x.

Taking (x_4-x_1, y_4-y_1) as a new origin, $F_4(x_4, y_4)$ equals the number of paths from $(0,0)$ to $(n-x_4, m-y_4)$ for which $R_k - k > nk/m$, $k=1,2,\dots,m-y_4$. Thus $F_4(x_4,y_4)$ is given by Lemma 3.2 with sample sizes $m'=m-y_4$, $n'=n-x_4$, $c_i - i = n-x_4$ and $b_i - i = \{ni/m\}$, i=1, 2, ..., $m-y_4$. Hence, for $d>0$,

$$(4.4)\qquad F_4(x_4,y_4)=$$

$$\begin{vmatrix} \binom{n-x_4-\{n/m\}+1}{1} & \binom{n-x_4-\{2n/m\}+1}{2} & \cdots & \binom{n-x_4-\{(m-y_4)n/m\}+1}{m-y_4} \\ 1 & \binom{n-x_4-\{2n/m\}+1}{1} & \cdots & \binom{n-x_4-\{(m-y_4)n/m\}+1}{m-y_4-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \binom{n-x_4-\{(m-y_4)n/m\}+1}{1} \end{vmatrix}_{(m-y_4)\times(m-y_4)}.$$

To determine $F_5(x_1,y_1)$, we observe that the point $(n-x_1, m-y_1)$ is y_1 units below and x_1 units to the left of (n,m). Considering the reversed path (i.e., rotating the path from $(n-x_1,m-y_1)$ to (n,m) in Fig. 2 about its left end through 180° in the clockwise direction so that its starting point becomes the end point and vice-versa) and then taking (n,m) as a new origin, $F_5(x_1,y_1)$ equals the number of paths from $(0,0)$ to (x_1,y_1) that remain entirely above the line $y=mx/n$. Thus $F_5(x_1,y_1)$ is given by Lemma 3.2 with sample sizes $m'=y_1$, $n'=x_1$; $b_i - i = 0$, $i=1,2,\dots,y_1$; $c_1=1$ and $c_j - j + 1 \equiv W_j = \min\left(x_1+1, \left\langle \frac{n(j-1)}{m} \right\rangle\right)$, $j=2,3,\dots,y_1$. Hence, for $d>0$,

$$(4.5)\qquad F_5(x_1,y_1)=\begin{vmatrix} \binom{\langle n/m\rangle}{1} & \binom{\langle n/m\rangle}{2} & \binom{\langle n/m\rangle}{3} & \cdots & \binom{\langle n/m\rangle}{y_1-1} \\ 1 & \binom{\langle 2n/m\rangle}{1} & \binom{\langle 2n/m\rangle}{2} & \cdots & \binom{\langle 2n/m\rangle}{y_1-2} \\ 0 & 1 & \binom{\langle 3n/m\rangle}{1} & \cdots & \binom{\langle 3n/m\rangle}{y_1-3} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \binom{W_{y_1}}{1} \end{vmatrix}_{(y_1-1)\times(y_1-1)},$$

since the first row of the determinant in Lemma 3.2 becomes $(1,0,0,\dots,0)$.

The foregoing leads to the following theorem.

THEOREM 4.1. *Let* p=g.c.d.(m,n), *i.e.*, $m=ap$, $n=bp$ *with* g.c.d.$(a,b)=1$. *Let* $\lambda=(s-r)/(a+b)$, $\mu=(s+t-r)/(a+b)$, $\delta=(u-r)/(a+b)$, $\theta=r/(a+b)$, $\varepsilon=s/(a+b)$ *and* $\zeta=(s+t)/(a+b)$ *be all integers such that* $\lambda \geqq i-1$, $\mu \geqq j-1$, $\delta \geqq l-1$, $\theta \geqq 1$, $\varepsilon \geqq i$ *and* $\zeta \geqq j$ *where* $1 \leqq i \leqq j \leqq l$. *Then*

(a) *for* $d>0$, $r \leqq s \leqq u \leqq m+n$, $0 \leqq t \leqq u-s$,

(4.6)
$$\binom{m+n}{n}\mathbf{P}\Big[mnD^+_{mn}=d, R^+_{mn}(1)=r, R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, R^+_{mn}(l)=u, Q^+_{mn}=l\Big] = \begin{cases} \phi_{\lambda,i-1}\phi_{\mu-\lambda,j-i}\phi_{\delta-\mu,l-j}F_4(x_4,y_4)F_5(x_1,y_1), \\ 0, \end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_1-mx_1=ny_2-mx_2=ny_3-mx_3=ny_4-mx_4=d$, $x_1+y_1=r, x_2+y_2=s$, $x_3+y_3=s+t$, $x_4+y_4=u$ *such that* $0\leqq x_1\leqq x_2\leqq x_3\leqq x_4\leqq n$, $0\leqq y_1\leqq y_2\leqq y_3\leqq y_4\leqq m$, *and*

(b) *for* $d=0$, $\quad r\leqq s\leqq m+n, \quad 0\leqq t\leqq m+n-s,$

(4.7)
$$\binom{m+n}{n}\mathbf{P}\Big[D^+_{mn}=0, R^+_{mn}(1)=r, R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, R^+_{mn}(l)=m+n, Q^+_{mn}=l\Big] =$$
$$=\binom{m+n}{n}\mathbf{P}\Big[D^+_{mn}=0, R^+_{mn}(1)=r, R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, Q^+_{mn}=l\Big] = \begin{cases} \phi_{\theta,1}\phi_{\varepsilon-\theta,i-1}\phi_{\zeta-\varepsilon,j-i}\phi_{p-\zeta,l-j}, \\ 0, \end{cases}$$

according to whether or not θ *is an integer* $\geqq 1$, ε *is an integer* $\geqq i$ *and* ζ *is an integer* $\geqq j$.

REMARK. It may be pointed out that the joint distribution of D^+_{mn}, $R^+_{mn}(1)$, $R^+_{mn}(i)$, $M^+_{mn}(i,j)$, $R^+_{mn}(l)$ and Q^+_{mn} when $D^+_{mn}=0$, as given in (4.7), could be obtained directly without applying any transformation.

Theorem 4.1 gives immediately the joint distribution of D^+_{mn}, $R^+_{mn}(i)$, $M^+_{mn}(i,j)$ and Q^+_{mn} as follows:

COROLLARY 4.1. *Let* $p=$ g.c.d.(m,n), *i.e.*, $m=ap$, $n=bp$ *with* g.c.d.$(a,b)=1$. *Then*

(a) *for* $d>0$,

(4.8)
$$\binom{m+n}{n}\mathbf{P}\left[mnD^+_{mn}=d, R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, Q^+_{mn}=l\right] = \begin{cases} \sum_{x_1}\sum_{y_1}\sum_{x_4}\sum_{y_4}\phi_{\lambda,i-1}\phi_{\mu-\lambda,j-i}\phi_{\delta-\mu,l-j}F_4(x_4,y_4)F_5(x_1,y_1), \\ 0, \end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_2 - mx_2 = ny_3 - mx_3 = d$, $x_2 + y_2 = s$, $x_3 + y_3 = s + t$ *such that* $0 \leqq x_2 \leqq \leqq x_3 \leqq n, 0 \leqq y_2 \leqq y_3 \leqq m$. *The summation extends over all possible points* (x_1, y_1) *and* (x_4, y_4) *satisfying the following conditions:*

(i) x_1, y_1, x_4 *and* y_4 *are all integers with* $0 \leqq x_1 \leqq x_2 \leqq x_3 \leqq x_4 \leqq n$, $0 \leqq y_1 \leqq y_2 \leqq y_3 \leqq y_4 \leqq m$,

(ii) $ny_1 - mx_1 = ny_4 - mx_4 = d$,

(iii) $\lambda = (s - (x_1 + y_1))/(a + b)$ *is an integer* $\geqq i - 1$,

(iv) $\mu = (s + t - (x_1 + y_1))/(a + b)$ *is an integer* $\geqq j - 1$,

(v) $\delta - \mu = ((x_4 + y_4) - s - t)/(a + b)$ *is an integer* $\geqq l - j$, *and*

(b) *for* $d = 0$,

$$\binom{m+n}{n} \mathbf{P}\,[D^+_{mn} = 0, R^+_{mn}(i) = s, M^+_{mn}(i, j) = t, Q^+_{mn} = l] = \tag{4.9}$$
$$= \begin{cases} \phi_{\varepsilon,i}\phi_{\zeta-\varepsilon,j-i}\phi_{p-\zeta,l-j}, \\ 0, \end{cases}$$

according to whether or not $\varepsilon = s/(a+b)$ *is an integer* $\geqq i$ *and* $\zeta = (s+t)/(a+b)$ *is an integer* $\geqq j$.

PARTICULAR CASES. (A) Setting $i = 1$ in Theorem 4.1 implies $s = r$, $x_2 = x_1$, $y_2 = y_1$, $\lambda = 0$, $\mu = t/(a + b) = \beta$, say (where β is an integer $\geqq j - 1$), $\delta = (u - r)/(a + b)$, $\varepsilon = \theta = r/(a + b)$, $\zeta = (r + t)/(a + b) = \beta + \theta$, $F_1 = 1$, $F_2 = \phi_{\beta,j-1}$, $F_3 = \phi_{\delta-\beta,l-j}$ and $\phi_{\varepsilon-\theta,i-1} = 1$. Hence Theorem 4.1 reduces to

COROLLARY 4.2. (a) *For* $d > 0$,

$$\binom{m+n}{n} \mathbf{P}\,[mnD^+_{mn} = d, R^+_{mn}(1) = r, M^+_{mn}(1, j) = t, R^+_{mn}(l) = u, Q^+_{mn} = l] \tag{4.10}$$
$$= \begin{cases} \phi_{\beta,j-1}\phi_{\delta-\beta,l-j}F_4(x_4, y_4)F_5(x_1, y_1), \\ 0, \end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_1 - mx_1 = ny_3 - mx_3 = ny_4 - mx_4 = d$, $x_1 + y_1 = r$, $x_3 + y_3 = r + t$, $x_4 + y_4 = u$ *such that* $0 \leqq x_1 \leqq x_3 \leqq x_4 \leqq n$, $0 \leqq y_1 \leqq y_3 \leqq y_4 \leqq m$, *and*

(b) *for* $d = 0$,

$$\binom{m+n}{n} \mathbf{P}\,[D^+_{mn} = 0, R^+_{mn}(1) = r, M^+_{mn}(1, j) = t, R^+_{mn}(l) = m + n, Q^+_{mn} = l] =$$

(4.11)

$$= \binom{m+n}{n} \mathbf{P}\,[D^+_{mn} = 0, R^+_{mn}(1) = r, M^+_{mn}(1,j) = t, Q^+_{mn} = l] =$$

$$= \begin{cases} \phi_{\theta,1}\phi_{\beta,j-1}\phi_{p-\beta-\theta,l-j}, \\ 0, \end{cases}$$

according to whether or not $\theta = r/(a+b)$ *is an integer* $\geqq 1$ *and* $\beta = t/(a+b)$ *is an integer* $\geqq j-1$.

(B) Setting $j = l$ in Theorem 4.1 implies $u = s+t$, $x_3 = x_4$, $y_3 = y_4$, $\lambda = (s-r)/(a+b)$, $\delta = \mu = (s+t-r)/(a+b)$, $\theta = r/(a+b)$, $\varepsilon = s/(a+b)$, $\zeta = p$, $F_3 = 1$ and $\phi_{p-\zeta,l-j} = 1$. Hence Theorem 4.1 reduces to

COROLLARY 4.3. (a) *For* $d > 0$,

(4.12)

$$\binom{m+n}{n} \mathbf{P}\,[mnD^+_{mn} = d, R^+_{mn}(1) = r, R^+_{mn}(i) = s, M^+_{mn}(i,l) = t, Q^+_{mn} = l] =$$

$$= \begin{cases} \phi_{\lambda,i-1}\phi_{\mu-\lambda,l-i}F_4(x_4,y_4)F_5(x_1,y_1), \\ 0, \end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_1 - mx_1 = ny_2 - mx_2 = ny_4 - mx_4 = d$, $x_1 + y_1 = r$, $x_2 + y_2 = s$, $x_4 + y_4 = s + t$ *such that* $0 \leqq x_1 \leqq x_2 \leqq x_4 \leqq n$, $0 \leqq y_1 \leqq y_2 \leqq y_4 \leqq m$, *and*

(b) *for* $d = 0$,

(4.13)

$$\binom{m+n}{n} \mathbf{P}\,[D^+_{mn} = 0, R^+_{mn}(1) = r, R^+_{mn}(i) = s, M^+_{mn}(i,l) = t, Q^+_{mn} = l] =$$

$$= \begin{cases} \phi_{\theta,1}\phi_{\varepsilon-\theta,i-1}\phi_{p-\varepsilon,l-i}, \\ 0, \end{cases}$$

according to whether or not $\theta = r/(a+b)$ *is an integer* $\geqq 1$ *and* $\varepsilon = s/(a+b)$ *is an integer* $\geqq i$.

(C) Setting $j = 1$ in (4.10) and (4.11) implies $t = 0$, $x_3 = x_1$, $y_3 = y_1$, $\beta = 0$, $\delta = (u-r)/(a+b)$, $\theta = r/(a+b)$ and $\phi_{\beta,j-1} = 1$. Hence Corollary 4.2 reduces to

COROLLARY 4.4. (a) *For* $d > 0$,

$$\binom{m+n}{n} \mathbf{P}\,[mnD^+_{mn} = d, R^+_{mn}(1) = r, R^+_{mn}(l) = u, Q^+_{mn} = l] = \tag{4.14}$$

$$= \begin{cases} \phi_{\delta,l-1}F_4(x_4,y_4)F_5(x_1,y_1), \\ 0, \end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_1-mx_1=ny_4-mx_4=d$, $x_1+y_1=r$, $x_4+y_4=u$ *such that* $0\leqq x_1\leqq x_4\leqq n$, $0\leqq y_1\leqq y_4\leqq m$, *and*

(b) *for* $d=0$,

(4.15)
$$\binom{m+n}{n}\mathbf{P}\,[D^+_{mn}=0, R^+_{mn}(1)=r, R^+_{mn}(l)=m+n, Q^+_{mn}=l]$$
$$=\binom{m+n}{n}\mathbf{P}\,[D^+_{mn}=0, R^+_{mn}(1)=r, Q^+_{mn}=l]$$
$$=\begin{cases}\phi_{\theta,1}\phi_{p-\theta,l-1},\\ 0,\end{cases}$$

according to whether or not $\theta=r/(a+b)$ *is an integer* $\geqq 1$.

The same results (4.14) and (4.15) could also be deduced from Corollary 4.3 by setting therein $i=l$, in which case $t=0$, $x_2=x_4$, $y_2=y_4$, $\mu=\lambda=(s-r)/(a+b)$, $\theta=r/(a+b)$, $\varepsilon=p$, $\phi_{\mu-\lambda,l-i}=1$ and $\phi_{p-\varepsilon,l-i}=1$.

REMARK. Corollary 4.4 is in agreement with Corollary 2 of Saran and Rani [5].

(D) Setting $j=i$ in Theorem 4.1 implies $t=0$, $x_3=x_2$, $y_3=y_2$, $\mu=\lambda=(s-r)/(a+b)$, $\delta=(u-r)/(a+b)$, $\theta=r/(a+b)$, $\zeta=\varepsilon=s/(a+b)$, $F_2=1$, $F_3=\phi_{\delta-\lambda,l-i}$ and $\phi_{\zeta-\varepsilon,j-i}=1$. Hence Theorem 4.1 reduces to

COROLLARY 4.5. (a) *For* $d>0$,

(4.16)
$$\binom{m+n}{n}\mathbf{P}\,[mnD^+_{mn}=d, R^+_{mn}(1)=r, R^+_{mn}(i)=s, R^+_{mn}(l)=u, Q^+_{mn}=l]$$
$$=\begin{cases}\phi_{\lambda,i-1}\phi_{\delta-\lambda,l-i}F_4(x_4,y_4)F_5(x_1,y_1),\\ 0,\end{cases}$$

according to whether or not there exists an integer solution to the equations $ny_1-mx_1=ny_2-mx_2=ny_4-mx_4=d$, $x_1+y_1=r$, $x_2+y_2=s$, $x_4+y_4=u$ *such that* $0\leqq x_1\leqq x_2\leqq x_4\leqq n$, $0\leqq y_1\leqq y_2\leqq y_4\leqq m$, *and*

(b) *for* $d=0$,

(4.17)
$$\binom{m+n}{n}\mathbf{P}\,[D^+_{mn}=0, R^+_{mn}(1)=r, R^+_{mn}(i)=s, R^+_{mn}(l)=m+n, Q^+_{mn}=l]$$
$$=\binom{m+n}{n}\mathbf{P}\,[D^+_{mn}=0, R^+_{mn}(1)=r, R^+_{mn}(i)=s, Q^+_{mn}=l]=$$
$$=\begin{cases}\phi_{\theta,1}\phi_{\varepsilon-\theta,i-1}\phi_{p-\varepsilon,l-i},\\ 0,\end{cases}$$

according to whether or not $\theta = r/(a+b)$ *is an integer* $\geqq 1$ *and* $\varepsilon = s/(a+b)$ *is an integer* $\geqq i$.

REMARK. Corollary 4.5 is in agreement with Theorem 1 of Saran and Rani [5].

It may be noted that Corollary 4.4 could also be deduced from Corollary 4.5 by setting either $i=1$ (in which case $s=r, x_2=x_1, y_2=y_1$, $\lambda=0$, $\phi_{\lambda,i-1}=1$, $\varepsilon=\theta=r/(a+b)$, $\phi_{\varepsilon-\theta,i-1}=1$) or $i=l$ (in which case $s=u$, $x_2=x_4$, $y_2=y_4$, $\lambda=\delta=(u-r)/(a+b), \phi_{\delta-\lambda,l-i}=1$, $\theta=r/(a+b)$, $\varepsilon=p$, $\phi_{p-\varepsilon,l-i}=1$).

5. The joint distribution of $R^+_{mn}(i), M^+_{mn}(i,j)$ and Q^+_{mn}

LEMMA 5.1. *The number of paths from* $(0,0)$ *to* (n,m) *which attain their maximum distance from the diagonal* $y=mx/n$ *for the first,* $i^{\text{th}}, j^{\text{th}}$ *and* l^{th} *(i.e., the last) time* $(1 \leqq i \leqq j \leqq l)$ *on the* $r^{\text{th}}, s^{\text{th}}, (s+t)^{\text{th}}$ *and* u^{th} *steps* $(r \leqq s \leqq u,\ 0 \leqq t \leqq u-s)$, *respectively, is the same as the number of paths from* $(0,0)$ *to* (n,m) *that are never above the diagonal and, moreover, never touch the diagonal after the* $(u-r)^{\text{th}}$ *step except at* (n,m) *and having exactly* $(l-1)$ *contacts with the diagonal; the* $(l-1)^{\text{st}}$ *contact occurring on the* $(u-r)^{\text{th}}$ *step, the* $(j-1)^{\text{st}}$ *contact occurring on the* $(s+t-r)^{\text{th}}$ *step and the* $(i-1)^{\text{st}}$ *contact occurring on the* $(s-r)^{\text{th}}$ *step.*

PROOF. Consider all points $(x_1,y_1),(x_2,y_2),(x_3,y_3)$ and (x_4,y_4) such that $x_1+y_1=r$, $x_2+y_2=s$, $x_3+y_3=s+t$, $x_4+y_4=u$, $x_i \leqq ny_i/m$, $i=1,2,3,4$, $1 \leqq x_1 \leqq x_2 \leqq x_3 \leqq x_4 \leqq n$, $1 \leqq y_1 \leqq y_2 \leqq y_3 \leqq y_4 \leqq m$. The set of required paths is the union of the disjoint subsets of paths through each of the possible quadruple of points $\{(x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4)\}$. By Lemma 4.1, the paths in each of these subsets are in one-to-one correspondence with those in the disjoint sets of paths from $(0,0)$ to (n,m) through the points (x_2-x_1,y_2-y_1), (x_3-x_1,y_3-y_1) and (x_4-x_1,y_4-y_1) that are never above the diagonal $y=mx/n$ and, moreover, never touch the diagonal after (x_4-x_1,y_4-y_1) except at (n,m) and having exactly $(l-1)$ contacts with the diagonal up to the point (x_4-x_1,y_4-y_1) of which the first $(i-1)$ contacts occur up to the point (x_2-x_1,y_2-y_1) and the next $(j-i)$ contacts occur up to the point (x_3-x_1,y_3-y_1). Hence the elements in the set of paths in question are in one-to-one correspondence with the elements in the set of paths from $(0,0)$ to (n,m) that are never above the diagonal and, moreover, never touch the diagonal after the $(u-r)^{\text{th}}$ step except at (n,m) and having exactly $(l-1)$ contacts with the diagonal; the $(l-1)^{\text{st}}$ contact occurring on the $(u-r)^{\text{th}}$ step, the $(j-1)^{\text{st}}$ contact occurring on the $(s+t-r)^{\text{th}}$ step and the $(i-1)^{\text{st}}$ contact occurring on the $(s-r)^{\text{th}}$ step.

THEOREM 5.1. *Let* $p = \text{g.c.d.}(m,n)$, *i.e.*, $m = ap$ *and* $n = bp$ *with* $\text{g.c.d.}(a,b) = 1$. *Then for* $1 \leqq i \leqq j \leqq l, r \leqq s \leqq u$ *and* $0 \leqq t \leqq u-s$

(5.1)

$$\binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(1)=r, R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, R^+_{mn}(l)=u, Q^+_{mn}=l\right] = \\ = \begin{cases} \phi_{\lambda,i-1}\phi_{\mu-\lambda,j-i}\phi_{\delta-\mu,l-j}F_4(\delta b, \delta a+1), \\ 0, \end{cases}$$

according to whether or not $\lambda = (s-r)/(a+b)$ *is an integer* $\geqq i-1, \mu = (s+t-r)/(a+b)$ *is an integer* $\geqq j-1$ *and* $\delta = (u-r)/(a+b)$ *is an integer* $\geqq l-1$.

PROOF. The number of paths envisaged in the left-hand side of (5.1) is given by Lemma 5.1. According to Lemma 5.1, the $(i-1)^{\text{st}}$, the $(j-1)^{\text{st}}$ and the $(l-1)^{\text{st}}$ contacts with the diagonal $y=mx/n$ occur on the $(s-r)^{\text{th}}$, the $(s+t-r)^{\text{th}}$ and the $(u-r)^{\text{th}}$ steps, respectively. Hence each one of $(s-r)$, $(s+t-r)$ and $(u-r)$ should be integer multiple of $(a+b)$. Let us assume, as in Theorem 4.1, that $s-r=\lambda(a+b)$, $s+t-r=\mu(a+b)$ and $u-r=\delta(a+b)$ where λ, μ and δ are all integers such that $\lambda \leqq \mu \leqq \delta, \lambda \geqq i-1, \mu \geqq j-1$ and $\delta \geqq l-1$. Then the number of transformed paths in Lemma 5.1 is the same as the number of paths from $(0,0)$ to $(\lambda b, \lambda a)$ that are never above the diagonal $y=mx/n$ and having exactly $(i-1)$ contacts with $y=mx/n$, times the number of paths from $(0,0)$ to $((\mu-\lambda)b, (\mu-\lambda)a)$ that are never above the diagonal $y=mx/n$ and having exactly $(j-i)$ contacts with $y=mx/n$, times the number of paths from $(0,0)$ to $((\delta-\mu)b, (\delta-\mu)a)$ that are never above the diagonal $y=mx/n$ and having exactly $(l-j)$ contacts with $y=mx/n$, times the number of paths from $(0,0)$ to $(n-\delta b, m-\delta a) \equiv ((p-\delta)b, (p-\delta)a)$ that remain entirely below the line $y=mx/n$ and never touch it in-between. Call these numbers T_1, T_2, T_3 and T_4, respectively. By Lemma 3.1, the numbers T_1, T_2 and T_3 are given by

$$T_1 = \phi_{\lambda,i-1}, \tag{5.2}$$

$$T_2 = \phi_{\mu-\lambda,j-i} \tag{5.3}$$

and

$$T_3 = \phi_{\delta-\mu,l-j}, \tag{5.4}$$

respectively. The number T_4 corresponds to those paths for which $R_k - k > nk/m$, $1 \leq k \leq m-\delta a-1$; $R_{m-\delta a} - (m-\delta a) = n-\delta b$. Thus T_4 is given by Lemma 3.2 with sample sizes $m' = m-\delta a$, $n' = n-\delta b$; $c_i - i = n-\delta b$, $i=1,2,\ldots,m-\delta a$; $b_j - j = \{nj/m\}$, $j=1,2,\ldots,m-\delta a-1$ and $b_{m-\delta a}-$

$(m-\delta a)=n-\delta b$, i.e.,

(5.5)

$$T_4=\begin{vmatrix} \binom{n-\delta b-\{\frac{n}{m}\}+1}{1} & \binom{n-\delta b-\{\frac{2n}{m}\}+1}{2} & \cdots & \binom{n-\delta b-\{\frac{(m-\delta a-1)n}{m}\}+1}{m-\delta a-1} & 0 \\ 1 & \binom{n-\delta b-\{\frac{2n}{m}\}+1}{1} & \cdots & \binom{n-\delta b-\{\frac{(m-\delta a-1)n}{m}\}+1}{m-\delta a-2} & 0 \\ 0 & 1 & \cdots & \binom{n-\delta b-\{\frac{(m-\delta a-1)n}{m}\}+1}{m-\delta a-3} & 0 \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & \binom{n-\delta b-\{\frac{(m-\delta a-1)n}{m}\}+1}{1} & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{vmatrix}_{(m-\delta a)\times(m-\delta a)}$$

$$=F_4(\delta b, \delta a+1),$$

which is given in (4.4). This proves Theorem 5.1 which in turn gives the joint distribution of $R^+_{mn}(i)$, $M^+_{mn}(i,j)$ and Q^+_{mn} given below:

COROLLARY 5.1.

$$\binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(i)=s, M^+_{mn}(i,j)=t, Q^+_{mn}=l\right]= \tag{5.6}$$
$$=\sum_r\sum_u \phi_{\lambda,i-1}\phi_{\mu-\lambda,j-i}\phi_{\delta-\mu,l-j}F_4(\delta b,\delta a+1),$$

where the summations extend over all possible positive integer values of r *and* u *such that* $r\leqq s\leqq u\leqq(m+n)$, $0\leqq t\leqq u-s$ *and for which* $\lambda=(s-r)/(a+b)$ *is an integer* $\geqq i-1$, $\mu=(s+t-r)/(a+b)$ *is an integer* $\geqq j-1$ *and* $\delta-\mu=(u-s-t)/(a+b)$ *is an integer* $\geqq l-j$.

PARTICULAR CASES. (A) Setting $i=1$ in Theorem 5.1 implies $s=r$, $\lambda=0$, $\mu=t/(a+b)=\beta$, say, $\delta=(u-r)/(a+b)$. Thus Theorem 5.1 reduces to

COROLLARY 5.2.

$$\binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(1)=r, M^+_{mn}(i,j)=t, R^+_{mn}(l)=u, Q^+_{mn}=l\right]= \tag{5.7}$$
$$=\begin{cases}\phi_{\beta,j-1}\phi_{\delta-\beta,l-j}F_4(\delta b,\delta a+1),\\ 0,\end{cases}$$

according to whether or not $\beta=t/(a+b)$ *is an integer* $\geqq j-1$ *and* $\delta=$ $=(u-r)/(a+b)$ *is an integer* $\geqq l-1$.

(B) Setting $j=l$ in Theorem 5.1 implies $u=s+t$, $\lambda=(s-r)/(a+b)$ and $\delta=\mu=(s+t-r)/(a+b)$. Thus Theorem 5.1 reduces to

COROLLARY 5.3.

$$(5.8)\qquad \binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(1)=r, R^+_{mn}(i)=s, M^+_{mn}(i,l)=t, Q^+_{mn}=l\right]= \\ =\begin{cases}\phi_{\lambda,i-1}\phi_{\mu-\lambda,l-i}F_4(\delta b,\delta a+1),\\ 0,\end{cases}$$

according to whether or not $\lambda=(s-r)/(a+b)$ *is an integer* $\geqq i-1$ *and* $\mu=(s+t-r)/(a+b)$ *is an integer* $\geqq l-1$.

(C) Setting $j=1$ in (5.7) implies $t=0$, $\beta=0$ and $\delta=(u-r)/(a+b)$. Thus Corollary 5.2 reduces to

COROLLARY 5.4.

$$(5.9)\qquad \binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(1)=r, R^+_{mn}(l)=u, Q^+_{mn}=l\right]=\begin{cases}\phi_{\delta,l-1}F_4(\delta b,\delta a+1),\\ 0,\end{cases}$$

according to whether or not $\delta=(u-r)/(a+b)$ *is an integer* $\geqq l-1$.

The same result (5.9) could also be deduced from Corollary 5.3 by setting $i=l$, in which case $t=0$ and $\mu=\lambda=(s-r)/(a+b)$.

REMARK. Corollary 5.4 is in agreement with Corollary 4 of Saran and Rani [5].

(D) Setting $j=i$ in Theorem 5.1 implies $t=0, \mu=\lambda=(s-r)/(a+b)$ and $\delta=(u-r)/(a+b)$. Thus Theorem 5.1 reduces to

COROLLARY 5.5.

$$(5.10)\qquad \binom{m+n}{n}\mathbf{P}\left[R^+_{mn}(1)=r, R^+_{mn}(i)=s, R^+_{mn}(l)=u, Q^+_{mn}=l\right]= \\ =\begin{cases}\phi_{\lambda,i-1}\phi_{\delta-\lambda,l-i}F_4(\delta b,\delta a+1),\\ 0,\end{cases}$$

according to whether or not $\lambda=(s-r)/(a+b)$ *is an integer* $\geqq i-1$ *and* $\delta=(u-r)/(a+b)$ *is an integer* $\geqq l-1$.

REMARK. Corollary 5.5 is in agreement with Theorem 2 of Saran and Rani [5].

It may be noted that Corollary 5.4 could also be deduced from Corollary 5.5 by setting either $i=1$ (in which case $s=r$ and $\lambda=0$) or $i=l$ (in which case $s=u$ and $\lambda=\delta=(u-r)/(a+b)$).

REFERENCES

[1] BIZLEY, M. T. L., Derivation of a new formula for the number of minimal lattice paths from $(0,0)$ to (km,kn) having just t contacts with the line $my=nx$ and having no points above this line; and a proof of Grossman's formula for

the number of paths which may touch but do not rise above this line, *J. Inst. Actuar.* **80** (1954), 55–62. *MR* **15** #846c

[2] GELLER, N. L., Two limiting distributions for two-sample Kolmogorov–Smirnov type statistics, Report No. 5, Center for System Science, Case Western Reserve University, Cleveland, Ohio, 1971.

[3] MAAG, U. R. and STEPHENS, M. A., The V_{NM} two-sample test, *Ann. Math. Statist.* **39** (1968), 923–935. *MR* **37** #3710

[4] SARAN, J. and SEN, K., On the distributions of $R^+_{mn}(j)$ and $(D^+_{mn}, R^+_{mn}(j))$, *Aplikace Mat.* **27** (1982), 417–425. *MR* **85c**:62117

[5] SARAN, J. and RANI, S., On the distributions of $D^+_{mn}, R^+_{mn}(j)$ and Q^+_{mn}, *J. Nonparametr. Statist.* **2** (1993), 285–293. *MR* **94k**:62082

[6] SARKADI, K., On Galton's rank order test, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** (1961), 127–131. *MR* **26**#3139

[7] STECK, G. P., The Smirnov two-sample tests as rank tests, *Ann. Math. Statist.* **40** (1969), 1449–1466. *MR* **39** #7777

[8] STECK, G. P. and SIMMONS, G. J., On the distributions of R^+_{mn} and (D^+_{mn}, R^+_{mn}), *Studia Sci. Math. Hungar.* **8** (1973), 79–89. *MR* **49** #4133

[9] VINCZE, I., Einige zweidimensionale Verteilungs- und Grenzverteilungssätze in der Theorie der geordneten Stichproben, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **2** (1957), 183–209. *MR* **21**#3915

(Received February 20, 1995)

DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICAL SCIENCES
UNIVERSITY OF DELHI
IND–110007 DELHI
INDIA

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 279–298

INTEGRABILITY AND LOWER LIMITS OF THE LOCAL TIME OF ITERATED BROWNIAN MOTION

ZH. SHI and M. YOR

Dedicated to Endre Csáki on his sixtieth birthday

Summary

We study the extraordinarily large and small values of the local time of iterated Brownian motion. It is known that the local time has exponential moments for deterministic times. We prove that, taken at appropriate random times, the local time has finite α-th moment if and only if $\alpha < 2/3$. We also investigate the almost sure lower asymptotics of both the local time at a fixed level and the maximum local time. The critical rate functions for these two processes are obtained, which improves previous results of Csáki et al. [15]. Our approach essentially relies on Ray–Knight theorems and the general theory of stochastic calculus, notably some refined martingale inequalities.

1. Introduction

Let $\{W_1(t);\, t \geqq 0\}$ and $\{W_2(t);\, t \geqq 0\}$ be independent one-dimensional Brownian motions, starting from 0, and let

$$Z(t) \stackrel{\text{def}}{=} W_1(|W_2(t)|), \qquad t \geqq 0. \tag{1.1}$$

The process $\{Z(t);\, t \geqq 0\}$, which is often referred to as "iterated Brownian motion", a terminology coined by Burdzy [9]–[10], has received much research interest from many mathematicians. See for example Hu [23, Chap. III] for a detailed account of history, motivations, as well as many references prior to December 1995. We only mention some more recent publications and preprints by Arcones [1], Csáki et al. [15]–[16] and [19], Csáki and Földes [18] (extensions to more general iterated processes), Benachour et al. [4], Hochberg and Orsingher [22] (connections with partial differential equations), Bertoin and Shi [6] (one-sided small values), Khoshnevisan and Lewis [25] (stochastic calculus with respect to iterated Brownian motion), and Xiao [28] (local times).

The starting point of the present paper is the following theorem due to Csáki et al. [15] (see also Burdzy and Khoshnevisan [11] for a slightly different model): there exists a jointly continuous version of $\{L_Z(t;x);$

1991 *Mathematics Subject Classification.* Primary 60J55; Secondary 60J65.
Key words and phrases. Local time, iterated Brownian motion.

0081–6906/97/$ 5.00

$t \geqq 0, -\infty < x < \infty\}$, the local time process of Z. Moreover, it can be represented as

$$(1.2)\qquad \begin{aligned} L_Z(t;x) &= \int_0^\infty \overline{L}_2(t;u)\, d_u L_1(u;x) \\ &= \int_0^\infty \Big(L_2(t;u) + L_2(t;-u) \Big)\, d_u L_1(u;x), \end{aligned}$$

where L_1, L_2 and $\overline{L}_2$ denote, respectively, the local times of W_1, W_2 and $|W_2|$. For notational convenience, we write

$$(1.3)\qquad L_Z(t) \stackrel{\text{def}}{=} L_Z(t;0) = \int_0^\infty \overline{L}_2(t;u)\, d_u L_1(u;0).$$

We aim at studying extraordinarily large and small values of $L_Z(t)$. We first consider the large values of $L_Z(t)$. The following law of the iterated logarithm (LIL) is due to Csáki et al. [15] and Xiao [28]: there exist (finite) universal constants $c_1 > 0$ and $c_2 > 0$ such that

$$(1.4)\qquad c_1 \leqq \limsup_{t\to\infty} \frac{L_Z(t)}{t^{3/4}(\log\log t)^{3/4}} \leqq c_2 \qquad a.s..$$

(To be precise, the upper bound in (1.4) is proved in [28], and the lower bound in [15].) The LIL (1.4) tells that, as t tends to infinity, $L_Z(t)$ can infinitely often reach the level of (a constant times) $t^{3/4}(\log\log t)^{3/4}$. To get additional information about the large values of L_Z, we can for example investigate the integrability of L_Z. It is not hard to see that for any fixed $t > 0$, $L_Z(t)$ admits finite moments of all orders (and indeed, it even has exponential moments, a property largely exploited by Xiao [28].) However, the situation becomes considerably different, if t is replaced by some carefully chosen random times, say T. Our choice for T is

$$(1.5)\qquad \tau_2(r) \stackrel{\text{def}}{=} \inf\Big\{ t > 0 : L_2(t;0) > r \Big\}, \qquad r > 0,$$

the inverse local time at 0 of W_2. The reason for which we have chosen τ_2 is very simple: intuitively, in order to keep $L_Z(T)$ at a high level, Z has to cross 0 very frequently over $[0,T]$, which at least intuitively will be satisfied if T is a non-left isolated zero of Z. Since this is the case for $T = \tau_2(r)$, with respect to W_2 (thus to Z as well), and moreover since $\tau_2(r)$ is closely related to both L_Z and L_2, it has become our immediate candidate.

Here is our main result for the integrability of $L_Z(\tau_2(r))$.

THEOREM 1.1. *For any $r > 0$ and $\alpha \geqq 0$,*

$$\mathbf{E}\Big[\Big(L_Z(\tau_2(r))\Big)^\alpha\Big] < \infty, \tag{1.6}$$

if and only if $0 \leqq \alpha < 2/3$.

REMARK 1.2. Theorem 1.1 tells that $L_Z(\tau_2(r))$ has polynomial (rather than exponential) upper tails. From (1.6) and using standard arguments, we can easily deduce the almost sure upper asymptotics of $r \mapsto L_Z(\tau_2(r))$, which are rather different from those of L_Z taken at a deterministic time.

Concerning the lower asymptotics of $L_Z(t)$, Csáki et al. [15] establish the following interesting bounds: with probability one,

$$\liminf_{t\to\infty} \frac{\log t}{t^{3/4}(\log\log t)^{1/2}} L_Z(t) \leqq 1, \tag{1.7}$$

$$\liminf_{t\to\infty} \frac{(\log t)^\beta}{t^{3/4}} L_Z(t) = \infty, \qquad \text{for all } \beta > \frac{3}{2}. \tag{1.8}$$

There is an obvious gap between these two estimates. Our aim here is to close the gap, namely, to show that $t^{3/4}/\log t$ is, in some sense, the "critical level" for the lower asymptotic behaviour of $L_Z(t)$.

THEOREM 1.3. *With probability one,*

$$\liminf_{t\to\infty} \frac{(\log t)^\beta}{t^{3/4}} L_Z(t) = \begin{cases} 0 & \textit{if } \beta \leqq 1, \\ \infty & \textit{otherwise.} \end{cases}$$

Theorem 1.3 provides a clear image of the liminf asymptotics of the local time of Z at a fixed level. Not surprisingly, the situation is considerably different for the *maximum* local time. Indeed, it is proved by Csáki et al. [15] that for some universal constants $c_3 > 0$ and $c_4 > 0$,

$$\liminf_{t\to\infty} \frac{(\log\log t)^{1/4}}{t^{3/4}} \sup_{x\in\mathbb{R}} L_Z(t;x) \leqq c_3 \qquad \text{a.s.} \tag{1.9}$$

$$\liminf_{t\to\infty} \frac{(\log\log t)^{3/4}}{t^{3/4}} \sup_{x\in\mathbb{R}} L_Z(t;x) \geqq c_4 \qquad \text{a.s..} \tag{1.10}$$

It turns out that the lower bound (1.10) is optimal (up to multiplication by a constant). Our next result is Chung's form of the LIL for $\sup_{x\in\mathbb{R}} L_Z(t;x)$.

THEOREM 1.4. *There exist absolute constants $c_5 > 0$ and $c_6 > 0$ such that*

$$c_5 \leqq \liminf_{t\to\infty} \frac{(\log\log t)^{3/4}}{t^{3/4}} \sup_{x\in\mathbb{R}} L_Z(t;x) \leqq c_6 \qquad \textit{a.s..}$$

We say a few words about the proofs of the theorems. Those of Theorem 1.1 and the second part of Theorem 1.3 heavily rely on Ray–Knight theorems for Brownian local times and the general theory of stochastic calculus. More precisely, we apply some powerful martingale inequalities (Facts 2.10 and 2.11 below) to iterated Brownian motion. In the proofs of Theorem 1.4 and the first part of Theorem 1.3, we use an idea we have learnt from Csáki et al. [15], with some refinement. The first-order stochastic calculus plays an important role in our approach, though we certainly have not exhausted all its advantages. We feel that, to get a better understanding of the local time of iterated Brownian motion, it would be worth studying the two-parameter process $\overline{L}_2(t; G_1(v))\, W_1(v)$ (for the definition of G_1, cf. (3.4) below), whose local time at 0 is $\int_0^v \overline{L}_2(t;u)\, d_u L_1(u;0)$.

Section 2 is devoted to some preliminaries on Bessel processes and the general theory of martingales, especially the first-order stochastic calculus and martingale inequalities. Theorems 1.1, 1.3 and 1.4 are proved, respectively, in Sections 3–5.

2. Notation and preliminaries

The key ingredients in the proofs of Theorems 1.1 and 1.3 are: (a) Ray–Knight theorems for Brownian local times; (b) special properties of Bessel processes; (c) first-order stochastic calculus and martingale inequalities. We briefly recall some known results relative to (a)–(b) in the first half of the section, and to (c) in the second half.

2.1. Ray–Knight theorems and Bessel processes

A d-dimensional Bessel process ($d \geqq 0$) is a linear diffusion on $\mathbb{R}_+$ with generator $\mathcal{G}f(x) = \frac{1}{2}f''(x) + \frac{d-1}{2x}f'(x)$ (at least, if f has compact support in $(0,\infty)$), and in the particular case $d \geqq 1$ is an integer, it can be realized as the Euclidean modulus (in $\mathbb{R}^d$) of d-dimensional Brownian motion. We refer to Revuz and Yor [27, Chap. XI] for a detailed account of general properties of Bessel processes.

We keep the notation introduced in Section 1, and add the following:

(2.1) $\{U(t); t \geqq 0\}$ is a Bessel process of dimension 0, with $U(0) = 1$,

$$\zeta \stackrel{\text{def}}{=} \inf\{t > 0 : U(t) = 0\}, \tag{2.2}$$

(2.3) $\{R(t);\ t \geqq 0\}$ is a Bessel process of dimension 4, with $R(0) = 0$,

$$\mathcal{L} \stackrel{\text{def}}{=} \sup\{t > 0 : R(t) = 1\}. \tag{2.4}$$

In words, ζ denotes the life-time of U, and $\mathcal{L}$ the Last exit time of R from 1. We also define the first hitting times associated to W_2 and $|W_2|$: for $x \in \mathbb{R}$,

$$\mathcal{H}_2(x) \stackrel{\text{def}}{=} \inf\{t > 0 : W_2(t) = x\}, \tag{2.5}$$

$$\overline{\mathcal{H}}_2(x) \stackrel{\text{def}}{=} \inf\{t > 0 : |W_2(t)| = x\}, \qquad (\inf \emptyset \stackrel{\text{def}}{=} \infty) \tag{2.6}$$

with $\mathcal{H}$ for Hitting time. Throughout the paper, unless stated otherwise, the processes W_1, W_2, R and U are assumed to be *mutually independent.*

The next item is a collection of known results concerning Brownian and Bessel processes, which we shall need later on.

FACT 2.1 (First Ray–Knight theorem). *For $a > 0$, $\{L_2(\mathcal{H}_2(a); a - x);$ $x \geq 0\}$ is a continuous inhomogeneous Markov process. When $0 \leq x \leq a$, it is a squared Bessel process of dimension 2 starting from 0, and becomes a squared Bessel process of dimension 0 for $x \geq a$.*

FACT 2.2 (Second Ray–Knight theorem). *Fix $r > 0$ and let $\tau_2(r)$ be as in* (1.5). *Then $\{L_2(\tau_2(r); x);\ x \geq 0\}$ and $\{L_2(\tau_2(r); -x);\ x \geq 0\}$ are two independent squared Bessel processes of dimension 0, starting from r.*

FACT 2.3 (Normalized Bessel process). *Let $\{R(t);\ t \geq 0\}$ and $\mathcal{L}$ be as in* (2.3) *and* (2.4), *respectively. Define the normalized Bessel process*

$$\rho(s) \stackrel{\text{def}}{=} \frac{R(s\,\mathcal{L})}{\sqrt{\mathcal{L}}}, \qquad 0 \leq s \leq 1. \tag{2.7}$$

For any bounded functional F, we have

$$\mathbf{E}\Big[F\Big(\rho(s);\ 0 \leq s \leq 1\Big)\Big] = \mathbf{E}\Big[\frac{2}{R^2(1)}\, F\Big(R(s);\ 0 \leq s \leq 1\Big)\Big]. \tag{2.8}$$

EXPLANATION. The relation (2.8) may be deduced from the fact that, given $\mathcal{L} = a$, $\{R(t);\ 0 \leq t \leq \mathcal{L}\}$ is a 4-dimensional Bessel bridge, on the time interval $[0, a]$, starting from 0 at time 0, and ending at 1 at time a.

FACT 2.4 (Bessel time-reversal). *Recalling* (2.1)–(2.4), *we have*

$$\Big\{U(\zeta - t);\ 0 \leq t \leq \zeta\Big\} \stackrel{\text{law}}{=} \Big\{R(t);\ 0 \leq t \leq \mathcal{L}\Big\},$$

where "$\stackrel{\text{law}}{=}$" denotes identity in distribution. In words, a Bessel process of dimension 0, starting from 1, is the time reversal of a Bessel process of dimension 4, starting from 0, killed when exiting from 1 for the last time.

FACT 2.5 (Integration by parts). *Let $\{\Re(t);\ t \geq 0\}$ be a Bessel process of positive dimension, starting from 0. Let $0 \leq a \leq b < \infty$ and f, g: $[a, b] \mapsto \mathbb{R}_+$ two continuous functions, with f nonincreasing, and g nondecreasing,*

$$\int_a^b \Re^2(g(x))\, d(-f(x)) + f(b)\Re^2(g(b)) \stackrel{\text{law}}{=} g(a)\Re^2(f(a)) + \int_a^b \Re^2(f(x))\, dg(x).$$

AMPLIFICATION. The integration by parts formula remains true when f and g are two continuous, $\mathbb{R}_+$-valued stochastic processes, independent of $\Re$, with f nonincreasing, and g nondecreasing.

REMARK 2.6. The classical Ray–Knight theorems can be found in Revuz and Yor [27, Chap. XI]. See also Yor [31, Chap. III] for many extensions. The absolute continuity relation (2.8) is due to Yor [32, p. 52] who actually provides a proof for all transient Bessel processes. The time reversal theorem for Bessel processes, stated in Fact 2.4, can for example be found in Revuz and Yor [27, Exercise XI.1.23]. It actually holds for each couple of Bessel processes of dimensions d and $4-d$, respectively. The integration by parts formula in Fact 2.5 is found in Yor [31, Exercise 2.5].

2.2. Martingale theory

Let $\{X(t);\, t \geqq 0\}$ be a continuous local martingale, $\{L_X(t);\, t \geqq 0\}$ its local time process at 0, and $\{\langle X\rangle(t);\, t \geqq 0\}$ its increasing process. Write $X^*(t) \stackrel{\text{def}}{=} \sup_{0 \leqq s \leqq t} |X(s)|$. For each $t \geqq 0$, define

$$G_X(t) \stackrel{\text{def}}{=} \sup\Big\{s \leqq t \colon X(s) = 0\Big\}, \qquad (\sup \emptyset \stackrel{\text{def}}{=} 0)$$

the last zero of X before t.

FACT 2.7. *If K is a locally bounded predictable process,*

$$K(G_X(t))X(t) = K(0)X(0) + \int_0^t K(G_X(s))\, dX(s).$$

In particular, $\{K(G_X(t))X(t);\, t \geqq 0\}$ is a continuous local martingale, with local time at 0 equal to

$$\int_0^t |K(G_X(s))|\, d_s L_X(s).$$

FACT 2.8. *For any locally bounded function $\phi\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$,*

$$\phi(L_X(t))X(t) = \phi(0)X(0) + \int_0^t \phi(L_X(s))\, dX(s).$$

Consequently, $\{\phi(L_X(t))X(t);\, t \geqq 0\}$ is a continuous local martingale, whose increasing process is given by

$$\langle \phi(L_X)X\rangle(t) = \int_0^t \phi^2(L_X(s))\, d_s\langle X\rangle(s),$$

while its local time at 0 *equals* $\Phi(L_X(t))$, *where* $\Phi(a) \stackrel{\text{def}}{=} \int_0^a \phi(x)\,dx$.

REMARK 2.9. Facts 2.7 and 2.8 are special examples of first-order stochastic calculus. For a full story, cf. Azéma and Yor [2], and also Revuz and Yor [27, Chap. VI].

FACT 2.10. *Assume* $X(0)=0$ *and let* $0 \leqq p < 1$. *There exist finite positive constants* $c_7(p)$ *and* $c_8(p)$, *depending only on* p, *such that*

$$c_7(p)\,\mathbf{E}\Big[\,(X^*(\infty))^p\Big] \leqq \mathbf{E}\Big[\,(L_X(\infty))^p\Big] \leqq c_8(p)\,\mathbf{E}\Big[\,(X^*(\infty))^p\Big].$$

FACT 2.11. *If* $X(0)=0$, T *is a positive random time,* $\alpha>0$ *and* $p>1$, *there exists a finite positive constant* $c_9(\alpha,p)$, *depending only on* α *and* p, *such that*

$$\mathbf{E}\Big[\Big(\frac{1}{\langle X\rangle(T)}\Big)^{\alpha/2}\Big] \leqq c_9(\alpha,p)\,\Big[\mathbf{E}\Big(\Big(\frac{1}{X^*(T)}\Big)^{\alpha p}\Big)\Big]^{1/p}. \tag{2.9}$$

REMARK 2.12. Although (2.9) is reminiscent of the Burkholder–Davis–Gundy inequalities, one cannot take $p=1$ on the right-hand side, even for T varying among stopping times (cf. Barlow et al. [3]).

REMARK 2.13. Fact 2.10 (and much more) can be found in Lenglart [26] and Yor [29], and Fact 2.11 due to Barlow et al. [3] (cf. also Yor [33, Chap. 13]). Some weaker versions of (2.9) may also be obtained using the ratio inequalities in Yor [30], Gundy [21] or Dellacherie et al. [20, Chap. XXIII], followed by an application of Hölder's inequality. However, (2.9) is more powerful, and more "user friendly" in practical examples.

3. Proof of Theorem 1.1

For brevity, we shall write $L_1(t) \stackrel{\text{def}}{=} L_1(t;0)$.

PROOF OF THE "IF" PART IN THEOREM 1.1. By (1.3),

$$L_Z(\tau_2(r)) = \int_0^\infty L_2(\tau_2(r);u)\,d_uL_1(u) + \int_0^\infty L_2(\tau_2(r);-u)\,d_uL_1(u).$$

Since the two terms on the right-hand side have the same distribution, we only have to verify

$$\mathbf{E}\Big[\Big(\int_0^\infty L_2(\tau_2(r);u)\,d_uL_1(u)\Big)^\alpha\Big] < \infty, \tag{3.1}$$

for $0 \leqq \alpha < 2/3$. According to the second Ray–Knight theorem (Fact 2.2) and the scaling property, (3.1) is equivalent to:

$$\mathbf{E}\Big[\Big(\int_0^\infty U^2(u)\,d_uL_1(u)\Big)^\alpha\Big] < \infty. \tag{3.2}$$

To prove (3.2), let us recall (2.2) and observe that

$$\int_0^\infty U^2(u)\,d_uL_1(u) \leqq \sup_{0\leqq u\leqq \zeta} U^2(u)\,L_1(\zeta) \stackrel{\text{law}}{=} \sup_{0\leqq t\leqq \mathcal{L}} R^2(t)\,L_1(\mathcal{L}),$$

the last identity in distribution following from Bessel time-reversal (cf. Fact 2.4). By scaling, the last expression is distributed as $\sqrt{\mathcal{L}}\sup_{0\leqq t\leqq\mathcal{L}} R^2(t)\,L_1(1)$. Hence, in the notation of (2.7),

$$\begin{aligned}\mathbf{E}\Big[\Big(\int_0^\infty U^2(u)\,d_uL_1(u)\Big)^\alpha\Big] &\leqq \mathbf{E}\Big[\Big(\sqrt{\mathcal{L}}\sup_{0\leqq t\leqq\mathcal{L}} R^2(t)\,L_1(1)\Big)^\alpha\Big]\\ &= \mathbf{E}\Big[(\rho(1))^{-3\alpha}\big(\sup_{0\leqq s\leqq 1}\rho(s)\big)^{2\alpha}\,(L_1(1))^\alpha\Big]\\ &= c_{10}(\alpha)\,\mathbf{E}\Big[(\rho(1))^{-3\alpha}\big(\sup_{0\leqq s\leqq 1}\rho(s)\big)^{2\alpha}\Big],\end{aligned}$$

where $c_{10}(\alpha) \stackrel{\text{def}}{=} \mathbf{E}((L_1(1))^\alpha) \in (0,\infty)$. Applying the absolute continuity relation (2.8) yields

$$\mathbf{E}\Big[\Big(\int_0^\infty U^2(u)\,d_uL_1(u)\Big)^\alpha\Big] \leqq 2c_{10}(\alpha)\,\mathbf{E}\Big[\frac{1}{(R(1))^{2+3\alpha}}\big(\sup_{0\leqq s\leqq 1}R(s)\big)^{2\alpha}\Big]. \tag{3.3}$$

Since $\mathbf{E}(R^q(1)) < \infty$ for all $q > -4$, and $\mathbf{E}(\sup_{0\leqq s\leqq 1} R^q(s)) < \infty$ for all $q > 0$, it follows from the Hölder inequality that the expectation term on the right-hand side of (3.3) is finite once $2+3\alpha < 4$, which means $\alpha < 2/3$. This yields (3.2), hence the "if" part in Theorem 1.1. □

To prove the "only if" part in Theorem 1.1, we need the so-called "Brownian meander" process, which is introduced by Chung [13], and which turns out to be an important process in the study of many "usual" Brownian functionals, cf. Bertoin and Pitman [5], Biane and Yor [7]–[8], and Yor [32].

FACT 3.1 (Brownian meander). *For $t > 0$, define*

$$G_1(t) \stackrel{\text{def}}{=} \sup\Big\{s \leqq t : W_1(s) = 0\Big\}, \tag{3.4}$$

the last zero of W_1 before t. Let $T>0$ a.s. be a random time independent of W_1. The process

(3.5)
$$\Big\{ m_1(s) \stackrel{\text{def}}{=} m_1^T(s) \stackrel{\text{def}}{=} \frac{1}{\sqrt{T-G_1(T)}} \,\Big| W_1\Big(G_1(T)+s(T-G_1(T))\Big)\Big|;\ 0 \leqq s \leqq 1 \Big\},$$

which is independent of T and $\{W_1(u);\ 0\leqq u \leqq G_1(T)\}$, is called a Brownian meander.

Proof of the "only if" part in Theorem 1.1. It suffices to show that

$$\mathbf{E}\Big[\Big(\int_0^\infty U^2(u)\,d_uL_1(u)\Big)^{2/3}\Big] = \infty. \tag{3.6}$$

Recall $G_1(t)$ from (3.4). Consider the continuous local martingale

$$N(t) \stackrel{\text{def}}{=} U^2(G_1(t))\,W_1(t) = \int_0^t U^2(G_1(s))\,dW_1(s), \qquad t \geqq 0,$$

the second equality following from Fact 2.7, which moreover confirms that the local time at 0 of N, denoted by $L_N(\cdot)$, is

$$L_N(t) = \int_0^t U^2(u)\,d_uL_1(u), \qquad t\geqq 0.$$

Applying Fact 2.10 to $X \stackrel{\text{def}}{=} N$ and $p \stackrel{\text{def}}{=} 2/3$ yields (with $c_{11} \stackrel{\text{def}}{=} c_7(2/3)$)

$$\begin{aligned}\mathbf{E}\Big[\Big(\int_0^\infty U^2(u)\,d_uL_1(u)\Big)^{2/3}\Big] &\geqq c_{11}\,\mathbf{E}\Big[\Big(\sup_{t\geqq 0} U^2(G_1(t))\,|W_1(t)|\Big)^{2/3}\Big]\\ &\geqq c_{11}\,\mathbf{E}\Big[\Big(U^2(G_1(\zeta))\,|W_1(\zeta)|\Big)^{2/3}\Big].\end{aligned}$$

Writing $\Theta \stackrel{\text{def}}{=} U^2(G_1(\zeta))\,|W_1(\zeta)|$ for brevity, the proof of (3.6) is reduced to showing the following estimate:

$$\mathbf{E}(\Theta^{2/3}) = \infty. \tag{3.7}$$

Applying (3.5) to $T \stackrel{\text{def}}{=} \zeta$, we obtain, from the independence properties stated in Fact 3.1,

$$\begin{aligned}\mathbf{E}(\Theta^{2/3}) &= \mathbf{E}\Big[\Big(U^2(G_1(\zeta))\,\sqrt{\zeta-G_1(\zeta)}\,m_1(1)\Big)^{2/3}\Big]\\ &= c_{12}\,\mathbf{E}\Big[U^{4/3}(G_1(\zeta))\,(\zeta-G_1(\zeta))^{1/3}\Big],\end{aligned}$$

with $c_{12} \stackrel{\text{def}}{=} \mathbf{E}\left((m_1(1))^{2/3}\right) \in (0,\infty)$. According to Lévy's arc sine law, for any fixed $t>0$, the density function of $t-G_1(t)$ is: $\mathbf{P}(t-G_1(t)\in ds)/ds = (1/\pi\sqrt{s(t-s)}\,)\,\mathbf{1}_{\{0<s<t\}}$. By means of Bessel time-reversal (cf. Fact 2.4),

$$\begin{aligned}
\mathbf{E}(\Theta^{2/3}) &= c_{12}\,\mathbf{E}\left[\int_0^{\zeta} \frac{ds}{\pi\sqrt{s(\zeta-s)}}\, U^{4/3}(\zeta-s)\, s^{1/3}\right] \\
&= \frac{c_{12}}{\pi}\,\mathbf{E}\left[\int_0^{\zeta} \frac{ds}{s^{1/6}(\zeta-s)^{1/2}}\, U^{4/3}(\zeta-s)\right] \\
&= \frac{c_{12}}{\pi}\,\mathbf{E}\left[\int_0^{\mathcal{L}} \frac{ds}{s^{1/6}(\mathcal{L}-s)^{1/2}}\, R^{4/3}(s)\right].
\end{aligned}$$

Recalling (2.7)–(2.8), this leads to (writing $c_{13} \stackrel{\text{def}}{=} c_{12}/\pi$):

$$\begin{aligned}
\mathbf{E}(\Theta^{2/3}) &= c_{13}\,\mathbf{E}\left[\mathcal{L}\int_0^{1} \frac{ds}{s^{1/6}(1-s)^{1/2}}\, \rho^{4/3}(s)\right] \\
&= 2c_{13}\,\mathbf{E}\left[\frac{1}{R^4(1)}\int_0^{1} \frac{ds}{s^{1/6}(1-s)^{1/2}}\, R^{4/3}(s)\right] \\
&\geqq 2c_{13}\,\mathbf{E}\left[\frac{\mathbf{1}_{\{R(1)\leqq 1\}}}{R^4(1)}\int_0^{1} \frac{ds}{s^{1/6}(1-s)^{1/2}}\, R^{4/3}(s)\right].
\end{aligned}$$

Consider the function

$$h(r) \stackrel{\text{def}}{=} \mathbf{E}\left[\int_0^{1} \frac{ds}{s^{1/6}(1-s)^{1/2}}\, R^{4/3}(s)\,\middle|\, R(1)=r\right], \qquad 0\leqq r\leqq 1.$$

Via the Bessel bridge, it is easily checked that h is continuous and strictly positive over $[0,1]$. Hence, $\inf_{0\leqq r\leqq 1} h(r)>0$. Accordingly,

$$\mathbf{E}(\Theta^{2/3}) \geqq c_{14}\,\mathbf{E}\left[\frac{\mathbf{1}_{\{R(1)\leqq 1\}}}{R^4(1)}\right] = \infty,$$

proving (3.7). □

4. Proof of Theorem 1.3

We first study the integrability of a negative power of a two-dimensional iterated Brownian motion under the L^2-norm.

PROPOSITION 4.1. *Let $\{B(t);\ t \geqq 0\}$ be standard Brownian motion with local time L at 0, and $\{\Re(t);\ t \geqq 0\}$ a two-dimensional Bessel process starting from 0, independent of B. Then*

$$\mathbf{E}\Big[\Big(\int_0^1 \Re^2(L(s))\,ds\Big)^{-\mu}\Big] < \infty, \qquad \text{for all } \mu < 1, \tag{4.1}$$

while

$$\mathbf{E}\Big[\Big(\int_0^1 \Re^2(L(s))\,ds\Big)^{-1}\Big] = \infty. \tag{4.2}$$

PROOF. We begin with the proof of (4.1). Only the case of positive μ needs to be treated. By enlarging the filtration of B with $\sigma\{\Re(u); u \geqq 0\}$, it is seen, with the aid of Fact 2.8, that $\Re(L(t))\,B(t)$ is a continuous local martingale, with increasing process $\int_0^t \Re^2(L(s))\,ds$. Applying (2.9) to $X(t) \overset{\text{def}}{=} \Re(L(t))\,B(t)$, $\alpha = 2\mu$ and $T = 1$, we have, for all $p > 1$,

$$\mathbf{E}\Big[\Big(\int_0^1 \Re^2(L(s))\,ds\Big)^{-\mu}\Big] \leqq c_9(\alpha, p)\,\Big\{\mathbf{E}\Big[\Big(\sup_{0 \leqq s \leqq 1} |\Re(L(s))B(s)|\Big)^{-2\mu p}\Big]\Big\}^{1/p}.$$

It remains to show that for all $\nu < 2$,

$$\Delta \overset{\text{def}}{=} \mathbf{E}\Big[\Big(\sup_{0 \leqq s \leqq 1} |\Re(L(s))B(s)|\Big)^{-\nu}\Big] < \infty. \tag{4.3}$$

Let $s^* \in [1/2, 1]$ satisfy $|B(s^*)| = \sup_{1/2 \leqq s \leqq 1} |B(s)|$. Note that s^* is *independent* of the Bessel process $\Re$. Hence

$$\begin{aligned}\sup_{0 \leqq s \leqq 1} |\Re(L(s))B(s)| &\geqq \sup_{1/2 \leqq s \leqq 1} |\Re(L(s))B(s)| \\ &\geqq \Re(L(s^*))\,|B(s^*)| \\ &\overset{\text{law}}{=} \Re(1)\,\sqrt{L(s^*)}\,|B(s^*)|,\end{aligned}$$

by scaling. The last term being greater than (or equal to) $\Re(1)\,\sqrt{L(1/2)}\,\sup_{1/2 \leqq s \leqq 1} |B(s)|$, we obtain:

$$\Delta \leqq \mathbf{E}\Big(\frac{1}{\Re^\nu(1)}\Big)\,\mathbf{E}\Big(\frac{1}{(L(1/2))^{\nu/2}(\sup_{1/2 \leqq s \leqq 1} |B(s)|)^\nu}\Big).$$

Since

$$\mathbf{E}\Big(\Re^a(1)\Big) < \infty, \qquad \text{for all } a > -2,$$

$$\mathbf{E}\Big(L^a(1/2)\Big) < \infty, \qquad \text{for all } a > -1,$$

$$\mathbf{E}\Big[\Big(\sup_{1/2 \leqq s \leqq 1} |B(s)|\Big)^a\Big] < \infty, \qquad \text{for all } a \in \mathbb{R},$$

it follows from Hölder's inequality that $\Delta < \infty$. This proves (4.3).

The proof of (4.2) follows from the subsequent inequality and identity in law:

$$\int_0^1 \Re^2(L(s))\, ds \leqq \sup_{0 \leqq u \leqq L(1)} \Re^2(u) \overset{\text{law}}{=} L(1) \sup_{0 \leqq u \leqq 1} \Re^2(u),$$

and the fact that $\mathbf{E}(1/L(1)) = \infty$. □

COROLLARY 4.2. *For all* $\mu < 1$,

$$\mathbf{E}\Big[\Big(\int_0^1 \Re^2(1-s)\, dL(s)\Big)^{-\mu}\Big] < \infty. \tag{4.4}$$

PROOF OF COROLLARY 4.2. Taking $f(x) = 1 - x$, $g(x) = L(x)$, $a = 0$ and $b = 1$ in the integration by parts formula in the amplified form of Fact 2.5, we have

$$\int_0^1 \Re^2(L(s))\, ds \overset{\text{law}}{=} \int_0^1 \Re^2(1-s)\, dL(s),$$

which implies that (4.4) is equivalent to (4.1). □

REMARK 4.3. It would also be possible to show directly the integrability property (4.4) from the following argument:

The process $\{\int_0^t \Re^2(1-s)\, dL(s);\ 0 \leq t \leq 1\}$ is the local time at 0 of the martingale $\{\Re^2(1 - G(s))\, B(s);\ 0 \leq s \leq 1\}$ with respect to the enlarged filtration already considered in the proof of Proposition 4.1, where $G(s) \overset{\text{def}}{=} \sup\{u \leq s : B(u) = 0\}$. Moreover, if $u^* \in [0, 1/2]$ satisfies $|B(u^*)| = \sup_{0 \leqq u \leqq 1} |B(u)|$, we find:

$$\begin{aligned}\sup_{0 \leqq u \leqq 1} \Big(\Re^2(1 - G(u))\, |B(u)|\Big) &\geqq \Re^2(1 - G(u^*))\, |B(u^*)| \\ &\overset{\text{law}}{=} \Re^2(1)\, (1 - G(u^*))\, |B(u^*)| \\ &\geqq \frac{\Re^2(1)}{2} \sup_{0 \leqq u \leqq 1/2} |B(u)|.\end{aligned}$$

This implies that

$$\mathbf{E}\Big[\Big(\sup_{0\leqq u\leqq 1}\Re^2(1-G(u))\,|B(u)|\Big)^{-\nu}\Big]<\infty, \qquad \text{for } \nu<1.$$

To obtain (4.4), it then remains to use either the random time inequalities in Barlow et al. [3], which in this case yield:

$$\mathbf{E}\Big[\Big(\frac{1}{L_X(T)}\Big)^{\mu}\Big]\leqq c_{15}(\mu,p)\Big\{\mathbf{E}\Big[\Big(X^*(T)\Big)^{-\mu p}\Big]\Big\}^{1/p},$$

for $0<\mu<1$, $p>1/(1-\mu)$ and positive random time T ($c_{15}(\mu,p)$ being a constant depending only on μ and p), or the ratio inequalities as in Yor [30], Gundy [21] or Dellacherie et al. [20, Chap. XXIII].

The main ingredient in the proof of Theorem 1.3 is the following estimate of the lower tail of $L_Z(1)$, which is obtained by means of Corollary 4.2.

PROPOSITION 4.4. *Let $L_Z(t)$ be the local time at* 0 *of the iterated Brownian motion process Z defined in* (1.1). *For each $\mu<1$, there exists $c_{16}(\mu)\in(0,\infty)$, depending only on μ, such that*

$$\mathbf{P}\Big(L_Z(1)<\varepsilon\Big)\leqq c_{16}(\mu)\,\varepsilon^{\mu}, \qquad \textit{for all } \varepsilon>0. \tag{4.5}$$

PROOF. Note that $L_Z(t)$ inherits a scaling property from that of Brownian motion; precisely, for any given $c>0$,

$$\Big\{L_Z(ct);\, t\geqq 0\Big\}\overset{\text{law}}{=}\Big\{c^{3/4}L_Z(t);\, t\geqq 0\Big\}. \tag{4.6}$$

Let $r>0$, whose value will be chosen ultimately. Recall the definition of $\overline{\mathcal{H}}_2(r)$ from (2.6). It is easily seen that $L_Z(\overline{\mathcal{H}}_2(r))\overset{\text{law}}{=}r^{3/2}L_Z(\overline{\mathcal{H}}_2(1))$. Accordingly,

$$\begin{aligned}\mathbf{P}\Big(L_Z(1)<\varepsilon\Big)&\leqq\mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(r))<\varepsilon\Big)+\mathbf{P}\Big(\overline{\mathcal{H}}_2(r)>1\Big)\\&\leqq\mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(1))<\frac{\varepsilon}{r^{3/2}}\Big)+\mathbf{P}\Big(\sup_{0\leqq s\leqq 1}|W_2(s)|<r\Big)\\&\leqq\mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(1))<\frac{\varepsilon}{r^{3/2}}\Big)+2\exp\Big(-\frac{\pi^2}{8r^2}\Big).\end{aligned} \tag{4.7}$$

In the last inequality, we have used Chung's exact distribution function of Brownian motion under the sup-norm (cf. [12, p. 221]).

Let us treat the first term on the right-hand side of (4.7). Write $\delta\overset{\text{def}}{=}\varepsilon/r^{3/2}$. Starting from 0, when W_2 exits from $[-1,1]$ for the first time (at time $\overline{\mathcal{H}}_2(1)$, by definition), there are two possibilities: $W_2(\overline{\mathcal{H}}_2(1))$ equals either 1, or -1. In other words, $\overline{\mathcal{H}}_2(1) = \mathcal{H}_2(1) \wedge \mathcal{H}_2(-1)$. By symmetry,

$$\begin{aligned} \mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(1)) < \delta\Big) &\leqq 2\mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(1)) < \delta;\ \overline{\mathcal{H}}_2(1) = \mathcal{H}_2(1)\Big) \\ &\leqq 2\mathbf{P}\Big(L_Z(\mathcal{H}_2(1)) < \delta\Big) \\ &\leqq 2\mathbf{P}\Big(\int_0^1 L_2(\mathcal{H}_2(1); u)\, d_u L_1(u) < \delta\Big), \end{aligned} \tag{4.8}$$

using (1.2). According to the first Ray–Knight theorem (cf. Fact 2.1),

$$\int_0^1 L_2(\mathcal{H}_2(1); u)\, d_u L_1(u) \stackrel{\text{law}}{=} \int_0^1 \Re^2(1-u)\, dL(u),$$

in the notation of Proposition 4.1. By (4.4),

$$\mathbf{E}\Big[\Big(\int_0^1 L_2(\mathcal{H}_2(1); u)\, d_u L_1(u)\Big)^{-\mu}\Big] < \infty.$$

Using (4.8) and Chebyshev's inequality,

$$\mathbf{P}\Big(L_Z(\overline{\mathcal{H}}_2(1)) < \delta\Big) \leqq c_{17}(\mu)\, \delta^{\mu}.$$

Recall that $\delta = \varepsilon / r^{3/2}$. Going back to (4.7), and taking $r \stackrel{\text{def}}{=} 1/\sqrt{\log(1/\varepsilon)}$, we get the desired estimate (4.5). □

Proof of the "otherwise" part in Theorem 1.3. Pick $\beta > 1$. There exists small $\theta > 0$ such that $\beta > 1 + \theta$. Let $t_k \stackrel{\text{def}}{=} e^k$ for $k \geqq 1$. By (4.6) and (4.5) (with $\mu \stackrel{\text{def}}{=} 1 - \theta/2$),

$$\begin{aligned} \mathbf{P}\Big(L_Z(t_k) < \frac{t_k^{3/4}}{(\log t_k)^{1+\theta}}\Big) &= \mathbf{P}\Big(L_Z(1) < \frac{1}{(\log t_k)^{1+\theta}}\Big) \\ &\leqq \frac{c_{16}(\theta/2)}{(\log t_k)^{(1+\theta)(1-\theta/2)}}, \end{aligned}$$

which is summable for k. Applying the Borel–Cantelli lemma yields: almost surely for sufficiently large k, $L_Z(t_k) \geqq t_k^{3/4}/(\log t_k)^{1+\theta}$. Let $t \in [t_k, t_{k+1}]$. Then for large t,

$$L_Z(t) \geqq L_Z(t_k) \geqq \frac{t_k^{3/4}}{(\log t_k)^{1+\theta}} = e^{-3/4} \frac{t_{k+1}^{3/4}}{(\log t_k)^{1+\theta}} \geqq e^{-3/4} \frac{t^{3/4}}{(\log t)^{1+\theta}},$$

which implies

$$\liminf_{t\to\infty} \frac{(\log t)^{1+\theta}}{t^{3/4}} L_Z(t) \geqq e^{-3/4} \qquad \text{a.s..}$$

Since $\beta > 1+\theta$, we obtain

$$\lim_{t\to\infty} \frac{(\log t)^{\beta}}{t^{3/4}} L_Z(t) = \infty \qquad \text{a.s..}$$

This yields the desired lower bound in Theorem 1.3. □

Now let us recall Kesten's LIL for the maximum local time. For an extension in the form of an integral test, cf. Csáki [14].

FACT 4.5 (Kesten [24]). *With probability one,*

$$\limsup_{t\to\infty} \frac{1}{(2t\log\log t)^{1/2}} \sup_{x\in\mathbb{R}} L_1(t;x) = 1 \qquad a.s..$$

PROOF OF THE "IF" PART IN THEOREM 1.3. Only the case $\beta = 1$ needs to be treated. The proof, as well as that of Theorem 1.4 in Section 5, is based on an idea we have learnt from Csáki et al. [15], with some refinement. Fix $\varepsilon > 0$, and let $t_k \stackrel{\text{def}}{=} k^k$ for $k \geqq k_0$, where $k_0 \stackrel{\text{def}}{=} k_0(\varepsilon)$ is a sufficiently large initial value. Define the events:

$$D_k \stackrel{\text{def}}{=} \left\{ \sup_{2t_{k-1}^{1/2} \leqq s \leqq 2t_k^{1/2}} (W_1(s) - W_1(2t_{k-1}^{1/2})) \leqq \frac{\varepsilon\, t_k^{1/4}}{\log t_k} \right\},$$

$$E_k \stackrel{\text{def}}{=} \left\{ \sup_{x\in\mathbb{R}} \Big(L_2(t_k;x) - L_2(t_{k-1};x) \Big) \leqq t_k^{1/2};\right.$$

$$\left. \sup_{t_{k-1}\leqq s\leqq t_k} |W_2(s) - W_2(t_{k-1})| \leqq t_k^{1/2} \right\},$$

$$F_k \stackrel{\text{def}}{=} D_k \cap E_k.$$

Observe that

$$\mathbf{P}(D_k) \geqq \mathbf{P}\Big(\sup_{0\leqq t\leqq 1} W_1(t) \leqq \frac{\varepsilon}{2\log t_k} \Big) \geqq \frac{\varepsilon}{3\log t_k},$$

$$\mathbf{P}(E_k) \geqq \mathbf{P}\Big(\sup_{x\in\mathbb{R}} L_2(1;x) \leqq 1;\ \sup_{0\leqq t\leqq 1} |W_2(t)| \leqq 1 \Big) \stackrel{\text{def}}{=} c_{18} > 0,$$

which yields

$$\mathbf{P}(F_k) = \mathbf{P}(D_k)\,\mathbf{P}(E_k) \geqq \frac{c_{18}\,\varepsilon}{3k\log k}.$$

Since the events $(F_k)_{k\geqq k_0}$ are mutually independent, by the Borel–Cantelli lemma, almost surely there exist infinitely many k's such that F_k is realized.

On the other hand, by the usual and Kesten's LILs (for the latter, cf. Fact 4.5), for all $k \geqq k_1 \stackrel{\text{def}}{=} k_1(\varepsilon, \omega)$,

$$\Lambda_1 \leqq 3(t_{k-1}^{1/2} \log\log t_{k-1}^{1/2})^{1/2} \leqq \frac{\varepsilon\, t_k^{1/4}}{\log t_k},$$
$$\Lambda_2 \leqq 2(t_{k-1} \log\log t_{k-1})^{1/2} \leqq t_k^{1/2},$$

where

$$\Lambda_1 = \text{ either } W_1(2t_{k-1}^{1/2}), \text{ or } \sup_{0 \leqq s \leqq 2t_{k-1}^{1/2}} W_1(s),$$
$$\Lambda_2 = \text{ either } \sup_{x \in \mathbb{R}} L_2(t_{k-1}; x), \text{ or } |W_2(t_{k-1})|, \text{ or } \sup_{0 \leqq s \leqq t_{k-1}} |W_2(s)|.$$

Consequently, for those infinitely many k's for which F_k holds and such that $k \geqq k_1$, we have

$$\sup_{0 \leqq s \leqq 2t_k^{1/2}} W_1(s) \leqq \frac{2\varepsilon\, t_k^{1/4}}{\log t_k}, \tag{4.9}$$

$$\sup_{x \in \mathbb{R}} L_2(t_k; x) \leqq 2t_k^{1/2}, \tag{4.10}$$

$$\sup_{0 \leqq s \leqq t_k} |W_2(s)| \leqq 2t_k^{1/2}. \tag{4.11}$$

According to Lévy's identity, the supremum process $\{\sup_{0 \leqq s \leqq t} W_1(s);\, t \geqq 0\}$ is distributed as the local time $\{L_1(t);\, t \geqq 0\}$. Hence, there exist infinitely many k's such that (4.10), (4.11) and

$$L_1(2t_k^{1/2}) \leqq \frac{2\varepsilon\, t_k^{1/4}}{\log t_k}, \tag{4.12}$$

hold simultaneously. Now, assuming (4.10)–(4.12),

$$\begin{aligned} L_Z(t_k) &= \int_0^\infty \overline{L}_2(t_k; u)\, d_u L_1(u) \\ &= \int_0^{2t_k^{1/2}} \overline{L}_2(t_k; u)\, d_u L_1(u) \\ &\leqq 2 \sup_{x \in \mathbb{R}} L_2(t_k; x)\, L_1(2t_k^{1/2}) \\ &\leqq \frac{8\varepsilon\, t_k^{3/4}}{\log t_k}, \end{aligned}$$

which implies

$$\liminf_{t\to\infty} \frac{\log t}{t^{3/4}} L_Z(t) \leqq 8\varepsilon \qquad \text{a.s.}.$$

Sending ε to 0 gives the "if" part in Theorem 1.3. □

5. Proof of Theorem 1.4

In view of (1.10), only the upper bound needs to be proved. We first recall the following estimate due to Csáki and Földes [17]: for $0 < a \leqq 1$,

$$\mathbf{P}\Big(\sup_{x\in\mathbb{R}} L_1(1;x) \leqq a\Big) \geqq \exp\Big(-\frac{c_{19}}{a^2}\Big), \tag{5.1}$$

for some absolute constant $c_{19} > 0$.

Define $c_{20} \stackrel{\text{def}}{=} \sqrt{8\,c_{19}}$ and $t_k \stackrel{\text{def}}{=} k^k$ (for large k). Consider

$$s_k \stackrel{\text{def}}{=} 2\,(t_k \log\log t_k)^{1/2},$$

$$D_k \stackrel{\text{def}}{=} \Big\{ \sup_{x\in\mathbb{R}} \Big(L_1(s_k;x) - L_1(s_{k-1};x) \Big) \leqq \frac{c_{20}\, t_k^{1/4}}{(\log\log t_k)^{1/4}} \Big\},$$

$$E_k \stackrel{\text{def}}{=} \Big\{ \sup_{x\in\mathbb{R}} \Big(L_2(t_k;x) - L_1(t_{k-1};x) \Big) \leqq \frac{c_{20}\, t_k^{1/2}}{(\log\log t_k)^{1/2}} \Big\},$$

$$F_k \stackrel{\text{def}}{=} D_k \cap E_k.$$

We have, by means of (5.1),

$$\mathbf{P}(D_k) \geqq \mathbf{P}\Big(\sup_{x\in\mathbb{R}} L_1(1;x) \leqq \frac{c_{20}}{2(\log\log t_k)^{1/2}} \Big) \geqq \exp\Big(-\frac{1}{2}\log\log t_k\Big),$$

$$\mathbf{P}(E_k) \geqq \mathbf{P}\Big(\sup_{x\in\mathbb{R}} L_2(1;x) \leqq \frac{c_{20}}{(\log\log t_k)^{1/2}} \Big) \geqq \exp\Big(-\frac{1}{8}\log\log t_k\Big).$$

Hence

$$\mathbf{P}(F_k) = \mathbf{P}(D_k)\,\mathbf{P}(E_k) \geqq \frac{1}{(k\log k)^{5/8}},$$

which implies $\sum_k \mathbf{P}(F_k) = \infty$. Thanks to the independence of the F_k's, we can apply the Borel–Cantelli lemma to conclude that, almost surely there exist infinitely many k's for which F_k is realized. On the other hand, by Kesten's LIL (cf. Fact 4.5), for all large k,

$$\sup_{x\in\mathbb{R}} L_1(s_{k-1};x) \leqq 2(s_{k-1}\log\log s_{k-1})^{1/2} \leqq \frac{c_{20}\, t_k^{1/4}}{(\log\log t_k)^{1/4}},$$

$$\sup_{x\in\mathbb{R}} L_2(t_{k-1};x) \leqq 2(t_{k-1}\log\log t_{k-1})^{1/2} \leqq \frac{c_{20}\, t_k^{1/2}}{(\log\log t_k)^{1/2}}.$$

Therefore there exist infinitely many k's such that

$$\sup_{x\in\mathbb{R}} L_1(s_k;x) \leqq \frac{2c_{20}\, t_k^{1/4}}{(\log\log t_k)^{1/4}}, \tag{5.2}$$

$$\sup_{x\in\mathbb{R}} L_2(t_k;x) \leqq \frac{2c_{20}\, t_k^{1/2}}{(\log\log t_k)^{1/2}}. \tag{5.3}$$

For those k satisfying (5.2)–(5.3), we have, by the usual LIL,

$$\begin{aligned}\sup_{x\in\mathbb{R}} L_Z(t_k;x) &= \sup_{x\in\mathbb{R}} \int_0^\infty \overline{L}_2(t_k;u)\, d_u L_1(u;x)\\ &= \sup_{x\in\mathbb{R}} \int_0^{s_k} \overline{L}_2(t_k;u)\, d_u L_1(u;x)\\ &\leqq 2\sup_{y\in\mathbb{R}} L_2(t_k;y)\, \sup_{x\in\mathbb{R}} L_1(s_k;x)\\ &\leqq \frac{8c_{20}^2\, t_k^{3/4}}{(\log\log t_k)^{3/4}},\end{aligned}$$

proving the upper bound in Theorem 1.4. □

REFERENCES

[1] Arcones, M. A., On the law of the iterated logarithm for Gaussian processes, *J. Theoret. Probab.* **8** (1995), 877–903.

[2] Azéma, J. and Yor, M., En guise d'introduction, *Temps locaux, Astérisque* **52–53** (1978), 3–16. *MR* **81b**:60042

[3] Barlow, M. T., Jacka, S. D. and Yor, M., Inequalities for a pair of processes stopped at a random time, *Proc. London Math. Soc.* (3) **52** (1986), 142–172. *MR* **87e**:60071

[4] Benachour, S., Roynette, B. and Vallois, P., Une approche probabiliste de l'équation $u_t = -\frac{1}{8}\frac{\partial^4 u}{\partial x^4}$, 1996 (preprint).

[5] Bertoin, J. and Pitman, J. W., Path transformations connecting Brownian bridge, excursion and meander, *Bull. Sci. Math.* **118** (1994), 147–166. *MR* **95b**:60097

[6] Bertoin, J. and Shi, Z., Hirsch's integral test for the iterated Brownian motion, *Séminaire de probabilités XXX* (Eds.: J. Azéma, M. Emery, and M. Yor), Lecture Notes in Mathematics, **1626**, Springer, Berlin, 1996, 361–368.

[7] Biane, P. and Yor, M., Valeurs principales associées aux temps locaux browniens, *Bull. Sci. Math.* (2) **111** (1987), 23–101. *MR* **88g**:60188

[8] Biane, P. and Yor, M., Quelques précisions sur le méandre brownien, *Bull. Sci. Math.* (2) **112** (1988), 101–109. *MR* **89i**:60156

[9] Burdzy, K., Some path properties of iterated Brownian motion, *Seminar on stochastic processes 1992* (Seattle, WA, 1992), (Eds.: E. Çinlar, K.L. Chung, and M. Sharpe), Birkhäuser, Boston, MA, 1993, 67–87. *MR* **95c**:60075

[10] Burdzy, K., Variation of iterated Brownian motion, *Measure-valued processes, stochastic partial differential equations and interacting systems* (Montreal, PQ, 1992),

(Ed.: D.A. Dawson), CRM Proc. Lecture Notes, **5**, Amer. Math. Soc., Providence, RI, 1994, 35–53. *MR* **95h**:60123

[11] Burdzy, K. and Khoshnevisan, D., The level sets of iterated Brownian motion, *Séminaire de probabilités XXIX* (Eds.: J. Azéma, M. Emery, P.-A. Meyer and M. Yor), Lecture Notes in Mathematics, **1613**, Springer, Berlin, 1995, 231–236.

[12] Chung, K. L., On the maximum partial sums of sequences of independent random variables, *Trans. Amer. Math. Soc.* **64** (1948), 205–233. *MR* **10**, 132b

[13] Chung, K. L., Excursions in Brownian motion, *Ark. Mat.* **14** (1976), 155–177. *MR* **57**#7791

[14] Csáki, E., An integral test for the supremum of Wiener local time, *Probab. Theory Related Fields* **83** (1989), 207–217. *MR* **91a**:60206

[15] Csáki, E., Csörgő, M., Földes, A. and Révész, P., The local time of iterated Brownian motion, *J. Theoret. Probab.* **9** (1996), 717–743.

[16] Csáki, E., Csörgő, M., Földes, A. and Révész, P., On the occupation time of an iterated process having no local time, 1996 (preprint).

[17] Csáki, E. and Földes, A., How small are the increments of the local time of a Wiener process?, *Ann. Probab.* **14** (1986), 533–546. *MR* **87h**:60147

[18] Csáki, E. and Földes, A., On the logarithmic average of iterated processes, *Statist. Probab. Lett.* (to appear).

[19] Csáki, E., Földes, A. and Révész, P., Strassen theorems for a class of iterated processes, *Trans. Amer. Math. Soc.* (to appear).

[20] Dellacherie, C., Maisonneuve, B. and Meyer, P.-A., *Probabilités et potentiel V*, Hermann, Paris, 1992.

[21] Gundy, R. F., The density of the area integral, *Conference on harmonic analysis in honor of Antoni Zygmund*, Vol. I, II, (Chicago, IL, 1981), Wadsworth, Belmont, CA, 1983, 138–149. *MR* **85k**:42048

[22] Hochberg, K. J. and Orsingher, E., Composition of stochastic processes governed by higher-order parabolic and hyperbolic equations, *J. Theoret. Probab.* **9** (1996), 511–532.

[23] Hu, Y., Sur le mouvement brownien: calcul de lois; études asymptotiques; filtrations; relations avec certaines équations paraboliques, Thèse de Doctorat de l'Université Paris VI, Paris, 1996.

[24] Kesten, H., An iterated logarithm law for local time, *Duke Math. J.* **32** (1965), 447–456. *MR* **31**#2751

[25] Khoshnevisan, D. and Lewis, T. M., Stochastic calculus for Brownian motion on a Brownian fracture, 1996 (preprint).

[26] Lenglart, E., Relation de domination entre deux processus, *Ann. Inst. H. Poincaré Sect. B (N.S.)* **13** (1977), 171–179. *MR* **57** #10810

[27] Revuz, D. and Yor, M., *Continuous martingales and Brownian motion*, 2nd edition, Grundlehren der mathematischen Wissenschaften, **293**, Springer, Berlin, 1994. *MR* **95h**:60072

[28] Xiao, Y., Local times and related properties of multi-dimensional iterated Brownian motion, 1996 (preprint).

[29] Yor, M., Les inégalités de sous-martingales comme conséquences de la relation de domination, *Stochastics* **3** (1979), 1–15. *MR* **81i**:60052

[30] Yor, M., Application de la relation de domination à certains renforcements des inégalités de martingales, *Séminaire de probabilités XVI* (Eds.: J. Azéma and M. Yor), Lecture Notes in Mathematics, **920**, Springer, Berlin–New York, 1982, 221–233. *MR* **83j**:60044

[31] Yor, M., *Some aspects of Brownian motion. Part I. Some special functionals*, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 1992. *MR* **93i**:60155

[32] YOR, M., *Local times and excursions for Brownian motion: A concise introduction*, Lecciones en Matemáticas, Universidad Central de Venezuela, 1995.

[33] YOR, M., *Some aspects of Brownian motion. Part II. Some recent martingale problems*, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 1996 (to appear).

(Received October 9, 1996)

LABORATOIRE DE STATISTIQUE THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

shi@ccr.jussieu.fr

LABORATOIRE DE PROBABILITÉS
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

secret@proba.jussieu.fr

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 299–304

A DIRECT DECOMPOSITION OF THE CONVOLUTION SEMIGROUP OF PROBABILITY DISTRIBUTIONS

G. J. SZÉKELY and A. ZEMPLÉNI

Dedicated to Endre Csáki on his sixtieth birthday

Abstract

We show that the convolution semigroup of probability measures over the real line is the direct sum of the subsemigroup of normal distributions and the set (not subsemigroup!) of probability measures without nondegenerate normal convolution factor. For higher than one dimensional probability measures this kind of direct decomposition does not hold.

1. Introduction

Infinitely divisible probability distributions play a very important role in probability theory. In the language of algebra, infinitely divisible distributions are the divisible elements of the (commutative) convolution semigroup of probability distributions. For commutative (Abelian) divisible groups (where all elements are divisible) there are three basic results (see e.g. Fuchs [3]):

(i) One can embed every Abelian group into a divisible Abelian group.

(ii) If a divisible group D is a subgroup of an Abelian group G, then D is a direct summand, i.e. $G=D+H$ is a direct decomposition where H is a subgroup of G.

(iii) (structure theorem) Every divisible Abelian group is the direct sum of quasicyclic and full rational groups.

One can prove that (i) remains valid for semigroups, i.e. every commutative semigroup is embeddable into a divisible one. We plan to return to this result and its applications in probability theory. The extension of the structure theorem (iii) does not seem to be easy and we cannot state any structure theorem for semigroups. In this paper we plan to study property (ii), i.e. the direct decomposability of commutative semigroups and especially an interesting direct decomposition of the convolution semigroup of one dimensional probability measures.

1991 *Mathematics Subject Classification.* Primary 60E05; Secondary 20M14.

Key words and phrases. Convolution semigroups, direct decompositions, normal distributions.

Research supported by the Hungarian National Foundation for Scientific Research Grants No. 1405, 7210 and T016237

2. Direct decompositions of semigroups

Definition 1. A semigroup S is a direct sum of its subsets S_1 and S_2 if for all $s \in S$ we have $s = s_1 s_2$ where $s_1 \in S_1$, $s_2 \in S_2$ and this decomposition is unique. In this case S_1 and S_2 are called direct summands. If S has only trivial direct decompositions (one of the summands consists of the single unit element) then S is called direct indecomposable.

Sometimes it is possible to choose the summands so that they be subsemigroups. In this case we emphasize this extra property. E.g. in the above mentioned direct decomposition for Abelian groups (property (ii)) if D is not divisible then we cannot guarantee that in the direct decomposition $G = D + H$ the subset H is a subgroup but we can guarantee the existence of such a direct decomposition.

The difference between divisible groups and semigroups is not only the possible abundance of idempotents in the semigroup case. For the additive group $\mathbb{R}$ of real numbers there exists a direct decomposition into the direct sum of rationals while for the additive semigroup of nonnegative integers we have the following result:

Proposition 1. *The additive semigroup of nonnegative real numbers is direct indecomposable if the direct summands are supposed to be subsemigroups.*

Proof. Suppose indirectly that there were a direct decomposition with nontrivial direct summands A and B and select an arbitrary $a > 0$ from A and a $b > a$ from B. Then by the supposed direct decomposition there exists $a' \in A$, $b' \in B$ such that $b - a = a' + b'$ and thus $b = 0 + b = (a + a') + b'$. Since $a + a' \in A$ (semigroup property) and $a + a' \neq 0$, the decomposition of b is not unique.

If in S there is a unique prime factorization then S is clearly the direct sum of the subsemigroups $\{S_p = \{p^n\}_{n=0}^{\infty} : p$ is a prime element in $S\}$. The proof of Proposition 1 shows that these subsemigroups are direct indecomposable if the direct summands are supposed to be subsemigroups since a direct decomposition of the semigroup of nonnegative integers (in the exponent of p) into the direct sum of two subsemigroups is impossible. (If we drop the subsemigroups restriction then we may split further the cyclic semigroups that were indecomposable in the strict sense.)

It was proved by Ruzsa and Székely [7] (see also Ruzsa and Székely [8]) that in case of the convolution semigroup of probability measures there is no unique prime factorization, moreover, there are no prime elements at all in this semigroup (although there are many irreducibles, and there exists a decomposition into the product of irreducibles — with a possible remainder term, called antiirreducible — see e.g. Khinchin [5], Kendall [4], Ruzsa and Székely [8]) therefore it is an interesting problem to find direct decompositions in this convolution semigroup.

3. A direct decomposition of the convolution semigroup of probability measures

One of the most important subsemigroups of the convolution semigroup of probability measures is the semigroup of normal distributions. It was proved by R.A. Fisher and D. Dugué [2] that the set of distributions without normal convolution component — let us call them *antinormal distributions* — is not a semigroup (in other words, the convolution of two antinormal distributions is not necessarily antinormal). We still have the hope to get a direct decomposition into the subsemigroup of normal distributions and the set of antinormal ones.

Denote by $D(\mathbb{R}^n)$ the convolution semigroup of probability measures over the n-dimensional Euclidean space $\mathbb{R}^n$.

THEOREM 1. *$D(\mathbb{R})$ is the direct sum of the subsemigroup of normal distributions and the set of antinormal distributions (the unit element, i.e. the degenerate at 0 distribution, is considered both normal and antinormal). This kind of direct decomposition, however, does not exist for $D(\mathbb{R}^n)$ if $n \geqq 2$.*

PROOF. The first part of Theorem 1 follows from a general observation. Let us first recall the following definition (see Ruzsa and Székely [8]):

DEFINITION 2. A commutative topological semigroup S with a unity e and a Hausdorff topology is called Hungarian if

(i) the associate graph $\{(x,y) \in S \times S : x \sim y\}$ is closed ($x \sim y$ if $x \mid y$ and $y \mid x$);
(ii) in $S^* = S/\sim$ the set of divisors is compact for each element s^*;
(iii) $x \sim y$ implies the existence of a unit (associate of the unity) such that $x = uy$.

The family of Hungarian semigroups is quite wide, for example the convolution semigroup $D(G)$ of probability measures for a locally compact Abelian group G is always Hungarian.

PROPOSITION 2. *If S is a closed subsemigroup of a Hungarian semigroup T then S is a direct summand in T if*

(i) *for every pair of elements s, t in S either $s \mid t$ or $t \mid s$;*
(ii) *we can always cancel in T by elements from S (i.e. if $sv = sz$ for an $s \in S$ and $v, z \in T$ then $v = z$);*
(iii) *if $s \in S$ and $s \sim s_1$ then $s_1 \in S$.*

PROOF. The set $V = \{v \in T : v = st$ with $s \in S$ and $t \in T$ can hold only for $s = e\}$ (e is the unity in S) is the direct complement of S in T, i.e. T is the direct sum of S and V.

To see this first we recall that S fulfills the definition of the set of "atoms" in Ruzsa and Székely [8], thus any element $t \in T$ can be decomposed as $t = s \cdot v$ where $s \in S$ and $v \in V$. (The original decomposition theorem claims only that

$s = \prod_i s_i$ with $s_i \in S$ but for a closed subsemigroup the product belongs to S as well.)

Now let us observe that property (i) $s_1 v_1 = s_2 v_2$ implies $s s_2 v_1 = s_2 v_2$ (or $s_1 v_1 = s s_1 v_2$) and thus by (ii) $s v_1 = v_2$ (or $v_1 = s v_2$). Then by the definition of V $s = e$, $v_1 = v_2$ and hence $s_1 = s_2$. □

One dimensional normal distributions can clearly play the role of S in the convolution semigroup T of all probability distributions.

For simplicity we prove the second part of the theorem for the bivariate case only; it is straightforward to generalize the proof for higher dimensions. For the proof of this part of the theorem we need the following lemma (we omit the simple proof).

LEMMA 1. *A bivariate, absolutely continuous distribution with density function f, for which $\exists (x,y) \in \mathbb{R}^2 : f(x,y) = 0$, cannot have a nondegenerate normal component.*

Now let μ be the following signed measure: $\mu(0,0) = -\delta$,

$$\mu(1,1) = \mu(1,-1) = \mu(-1,1) = \mu(-1,-1) = \frac{1}{4}(1+\delta)$$

where δ is a small positive number such that $\delta < \frac{1}{3}$.

LEMMA 2. *Let (X,Y) be a bivariate nondegenerate normal random variable with zero expectation. Denote its distribution and density functions by $\Phi_{X,Y}$ and $\phi_{X,Y}$, respectively. Then $\exists c > 0$ for which $\mu * \Phi_{cX,cY} \in D(\mathbb{R}^2)$ (i.e. the convolution is a bivariate probability distribution).*

PROOF OF LEMMA 2. The density function h of the convolution of μ and Φ is the following:

$$h(x,y) = -\delta \cdot \phi(x,y) + (1/4+\delta) \sum_{\varepsilon=\pm 1, \eta=\pm 1} \phi(x+\varepsilon, y+\eta). \tag{1}$$

We have to show that $h(x,y) \geqq 0$ for all $(x,y) \in \mathbb{R}^2$. Let us distinguish two cases:

a) $0 \in N = [x-1, x+1] \times [y-1, y+1]$. In this case we have

$$h(x,y) \geqq -\delta \cdot \phi(0,0) + 4(1/4+\delta) I \tag{2}$$

where $I = \inf_{(x,y)\in N} \phi(x,y)$ by (1). Inequality (2) implies $h(x,y) \geqq 0$ for $c \geqq c_0$.

b) $0 \notin N$. As the same values of ϕ form a nondegenerate ellipsoid centered at the point $(0,0)$, the value $\phi(x,y)$ lies between $\phi(x-1, y-1)$ and $\phi(x+1, y+1)$, $h(x,y) > 0$ is a consequence of $\delta < 1/3$ for any positive c.

The suitable c's form a nonempty closed set, so there is a minimal c_0. The functional $\gamma(c) = \min_{(x,y)\in N} h(x,y)$ is continuous, and we see from the proof above that $\gamma(c_0) = 0$ by the minimality of c_0.

Now let us consider two different bivariate normal distributions. By Lemma 2 we have that the corresponding convolutions are probability distributions, and both $\mu * \Phi_1$ and $\mu * \Phi_2$ fulfill the condition of Lemma 1. By this statement we have that ν has two different direct-decompositions: $\nu = (\mu * \Phi_1) * \Phi_2 = (\mu * \Phi_2) * \Phi_1$. □

4. Further problems

In Ruzsa and Székely [6] (see also Ruzsa and Székely [8]) it is proved that there exists a homomorphism $\varphi\colon D(\mathbb{R}) \to \mathbb{R}$ which is an extension of the expectation. Hence $D(\mathbb{R}) = D_0 + \mathbb{R}$ is a direct decomposition where $D_0 = \{F\colon F \in D(\mathbb{R}), \varphi(F) = 0\}$.

CONJECTURE 1. *If $D(\mathbb{R})$ is the direct sum of two subsemigroups then one of them consists of degenerate distributions only.*

PROBLEM 1. Is it possible to decompose $D(\mathbb{R})$ into the direct sum of (direct) indecomposable subsets? (Not even every Abelian group is a direct decomposition of direct indecomposable subgroups.)

REMARK. If the word *direct* is replaced by *subdirect* then such a decomposition always exists for all semigroups, and also for more general algebraic structures, as it was proved by G. Birkhoff [1]. In fact, every commutative semigroup is the subdirect sum of a family of subdirect indecomposable subsemigroups.

PROBLEM 2. Characterize all commutative semigroups S that are direct summands in every commutative semigroup that contains S.

CONJECTURE 2. *If in S there exist two elements s_1 and s_2 such that $s_1 \nmid s_2$ and $s_2 \nmid s_1$, then one can always find a semigroup $T \supset S$ such that S is not a direct summand in T.*

REFERENCES

[1] BIRKHOFF, G., *Lattice theory*, Amer. Math. Soc. Coll. Publ., Vol. 25, Amer. Math. Soc., New York, 1948. *MR* **10**, 673a

[2] FISHER, R. A. and DUGUÈ, D., Un résultat assez inattendu d'arithmétique des lois des probabilités, *C. R. Acad. Sci. Paris* **227** (1948), 1205–1206. *MR* **10**, 463d

[3] FUCHS, L., *Abelian groups*, Publishing House of the Hungarian Academy of Sciences, Budapest, 1958. *MR* **21** #5672

[4] KENDALL, D. G., Delphic semi-groups, infinitely divisible regenerative phenomena, and the arithmetic of p-functions, *Z. Wahrsch. Verw. Gebiete* **9** (1968), 163–195. *MR* **37** #5320

[5] KHINCHIN, A. YA., On the arithmetic of distribution functions, *Bull. MGU. Sér. Int., Sect. A: Math. Mécan.* **1** (1937), 6–17 (in Russian). *Zbl* **18**, 316

[6] RUZSA, I. Z. and SZÉKELY, G. J., An extension of expectation, *Z. Wahrsch. Verw. Gebiete* **53** (1980), 17–19. *MR* **82a**:60017

[7] Ruzsa, I. Z. and Székely, G. J., No distribution is prime, *Z. Wahrsch. Verw. Gebiete* **70** (1985), 263–269. *MR* **86j**:60023

[8] Ruzsa, I. Z. and Székely, G. J., *Algebraic probability theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley, New York, 1988. *MR* **89j**:60006

(Received March 10, 1996)

BUDAPESTI MŰSZAKI EGYETEM
MATEMATIKAI INTÉZETE
STOCZEK U. 2
H–1521 BUDAPEST
HUNGARY

szjg@vma.bme.hu

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS
STATISZTIKAI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

zempleni@ludens.elte.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 305–320

WEIGHTED APPROXIMATIONS OF PARTIAL SUM PROCESSES IN $D[0,\infty)$. II

B. SZYSZKOWICZ

Dedicated to Endre Csáki on the occasion of his 60th birthday

Abstract

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables with $\mathbf{E}X_1=0$ and $\mathbf{E}X_1^2=1$. In Szyszkowicz [11] we obtained the weighted version of Donsker's theorem in $D[0,1]$ for the optimal class of weight functions, which is the same as the class of functions q for which $\lim_{t\downarrow 0}|W(t)|/q(t)=0$ a.s.. In this paper we show that there is no need for a similar assumption for the weak convergence of weighted partial sum processes in $D[1,\infty)$. Namely, we prove weighted approximations of $n^{-1/2}S(nt)$, $1\leqq t<\infty$, by a standard Wiener process $\{W(t),\ 0\leqq t<\infty\}$ in probability, and hence also weak convergence of the $n^{-1/2}S(nt)/h(t)$ processes in $D[1,\infty)$ for the largest possible class of weight functions which is the same as the class of functions h for which $\limsup_{t\to\infty}|W(t)|/h(t)<\infty$ a.s.. This paper is a continuation of Szyszkowicz [11] in that we improve on those results, for which additional conditions on weight functions were used in there. These two papers, together with Szyszkowicz [10], establish in probability approximations and weak convergence of partial sum processes in $D[0,\infty)$ in weighted supremum and L_p, $0<p<\infty$, metrics under the assumption of two moments only for X_1 and for the optimal classes of weight functions.

1. Introduction

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables (i.i.d.r.v.'s) with $\mathbf{E}X_1=0$, $\mathbf{E}X_1^2=1$, partial sums $S(n)=X_1+\cdots+X_n$ and let $\{W(t),\ 0\leqq t<\infty\}$ denote standard Wiener process starting at zero.

Let Q be the class of positive functions q on $(0,1]$ which are nondecreasing near zero, and let

$$I(q,c)=\int_0^1 t^{-1}\exp(-ct^{-1}q^2(t))\,dt,\quad 0<c<\infty.$$

It is well known (cf. discussion in Section 2) that for $q\in Q$

$$\lim_{t\downarrow 0}|W(t)|/q(t)=0\quad \text{a.s.} \tag{1.1}$$

1991 *Mathematics Subject Classification.* Primary 60F17, 60G25; Secondary 60G50.

Key words and phrases. Partial sums, Wiener processes, weighted approximations in probability, weak convergence in $D[0,\infty)$, supremum metrics, sup-functionals.

Research supported by an NSERC Canada grant at Carleton University, Ottawa.

if and only if

$$I(q,c) < \infty \quad \text{for all } c > 0,$$

and

(1.2)
$$\limsup_{t\downarrow 0} |W(t)|/q(t) < \infty \quad \text{a.s.}$$

if and only if

$$I(q,c) < \infty \quad \text{for some } c > 0.$$

In Szyszkowicz [11] we proved the following result which we state here for the completeness of our presentation.

THEOREM 1.A. *Let $X_1, X_2, \ldots$ be independent, identically distributed random variables such that*

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}X_1^2 = 1,$$

and for each $n \geqq 1$ let $S(nt) = \sum_{i=1}^{[nt]} X_i$. Then a standard Wiener process $\{W(t),\ 0 \leqq t < \infty\}$ can be constructed in such a way that the following statements hold true.

(a) *Let $q \in Q$. Then, as $n \to \infty$*

$$\sup_{0<t\leqq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) = o_P(1)$$

if and only if $I(q,c) < \infty$ for all $c > 0$.

(b) *Let $q \in Q$. Then, as $n \to \infty$,*

$$\sup_{0<t\leqq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) = O_P(1)$$

if and only if $I(q,c) < \infty$ for some $c > 0$.

As a corollary we obtain "weighted" Donsker's theorem for the optimal class of weight functions which is the same as the class of functions $q \in Q$ for which (1.1) holds. Namely, we have the following result (cf. Szyszkowicz [11]).

COROLLARY 1.A. *Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}X_1^2 = 1.$$

Let $q \in Q$ and $\{W(t), t \geqq 0\}$ be a standard Wiener process. We have, as $n \to \infty$,

$$n^{-1/2}S(nt)/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t) \quad \text{in } D[0,1]$$

if and only if $I(q,c)<\infty$ *for all* $c>0$.

Throughout this paper weak convergence statements on Skorohod spaces are stated as corollaries to approximations in probability. Naturally, when talking about weighted weak convergence on such spaces we will always assume that the weights are c.d.l.g. functions.

For motivation and elaboration on the history of these results we refer to [11].

In the first part of this paper we prove the following theorem, which improves Theorem 2.2 of [11] by dropping the condition of regular variation of weight functions.

THEOREM 1.1. *Let* $X_1, X_2, \dots$ *be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1$$

and let $\{W(t), t\geqq 0\}$ *be a standard Wiener process. Let* $q\in Q$ *be such that* $I(q,c)<\infty$ *for some* $c>0$. *Then, as* $n\to\infty$, *we have*

$$\sup_{1/n\leqq t\leqq 1}\left|n^{-1/2}(S(nt)-W(nt))\right|\Big/q(t)=o_P(1).$$

As a corollary we obtain the following convergence in distribution result for the sup-functional of weighted partial sums for the optimal class of weight functions which is the same as the class of functions $q\in Q$ for which (1.2) holds.

COROLLARY 1.1. *Let* $X_1, X_2, \dots$ *be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1.$$

Let $q\in Q$ *and* $\{W(t), t\geqq 0\}$ *be a standard Wiener process.*

(a) *We have, as* $n\to\infty$,

$$\sup_{0<t\leqq 1}|n^{-1/2}S(nt)|/q(t)\xrightarrow{\mathcal{D}}\sup_{0<t\leqq 1}|W(t)|/q(t)$$

if and only if $I(q,c)<\infty$ *for some* $c>0$.

(b) *We have, as* $n\to\infty$,

$$\sup_{0<t\leqq 1}n^{-1/2}S(nt)/q(t)\xrightarrow{\mathcal{D}}\sup_{0<t\leqq 1}W(t)/q(t)$$

if and only if $I(q,c)<\infty$ *for some* $c>0$.

Proofs of Theorem 1.1 and Corollary 1.1 will be given in Section 2.

Obviously, Corollary 1.A implies convergence in distribution of any continuous in sup-norm functional of $n^{-1/2}S(nt)/q(t)$ to the corresponding functional of $W(t)/q(t)$ with $q\in Q$ and such that (1.1) holds. However, for the sup-functionals Corollary 1.1 yields convergence in distribution for the optimal class of weight functions which is the same as the class of functions of $q \in Q$ for which (1.2) holds. Consequently, for the convergence of the sup-functional of weighted partial sum processes we do not have to assume the restriction that $\lim_{t \downarrow 0} |W(t)|/q(t) = 0$ a.s., which is the optimal condition for having Corollary 1.A. Such a phenomenon was first noticed and proved for weighted empirical and quantile processes by Csörgő, Csörgő, Horváth and Mason [2], and then by Csörgő and Horváth [3] for partial sums as well when assuming the existence of more than two moments for X_1. In [11] we proved Theorem 1.1 and Corollary 1.1 under the additional condition that $q(t)/t^{1/2}$ is slowly varying at zero (cf. Theorems 2.2 and 2.3 there). We note that the assumption of $q \in Q$, i.e., that q is nondecreasing near zero, is not really restrictive since if q decreases near zero then $\lim_{t \downarrow 0} |W(t)|/q(t) = 0$ a.s..

For weighted L_p-approximations of partial sum processes in $D[0,1]$ in terms of the optimal class of weight functions when only two moments are assumed to be finite for X_1, we refer to Szyszkowicz [10]. It is of interest to note that the latter class of functions is yet bigger than the one obtained in Corollary 1.1.

Obviously, all our results for $t \in [0,1]$ can be restated on $[0,T]$ for any $0 < T < \infty$. On the other hand, since $W(t) \to \infty$ a.s. as $t \to \infty$, there is no weak convergence of $n^{-1/2}S(nt)$ on $[1,\infty)$. In [11] we introduced weight functions $h(t)$, where $h(t) \to \infty$ as $t \to \infty$, in order to study such phenomena near infinity. In the case of assuming the existence of more than two moments, in the mentioned paper we obtained a complete solution of the problem in $D[1,\infty)$ (cf. Theorem 3.1 and Corollary 3.1 there). Namely, we obtained approximation in probability, and hence also weak convergence, of our weighted partial sum processes in $D[1,\infty)$, whenever

$$\limsup_{t \to \infty} |W(t)|/h(t) < \infty \quad \text{a.s.}.$$

We note that a similar result on $D[0,1]$ does not hold. Namely, we do not have the weak convergence of weighted partial sum processes in $D[0,1]$ for all weight functions such that (1.2) holds. Indeed, we have to assume (1.1) even if we assumed the existence of more than two moments.

In this paper we obtain a complete solution of this problem in $D[1,\infty)$, assuming the existence of two moments only.

Let $\mathcal{H}$ be the class of those positive functions h on $[1,\infty)$ for which $h(t)/t$ is nonincreasing in the neighbourhood of infinity, and let

$$I_\infty(h,c) = \int_1^\infty t^{-1} \exp(-ct^{-1}h^2(t))dt, \quad 0 < c < \infty.$$

It is known (cf. Section 3) that for $h \in \mathcal{H}$

$$\lim_{t \to \infty} |W(t)|/h(t) = 0 \quad \text{a.s.} \tag{1.3}$$

if and only if

$$I_\infty(h,c)<\infty \quad \text{for all } c>0,$$

and

(1.4)
$$\limsup_{t\to\infty}|W(t)|/h(t)<\infty \quad \text{a.s.}$$

if and only if

$$I_\infty(h,c)<\infty \text{ for some } c>0.$$

The main result in the second part of this paper is the following theorem.

THEOREM 1.2. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1.$$

Let $h\in\mathcal{H}$ and $I_\infty(h,c)<\infty$ for some $c>0$. Then a standard Wiener process $\{W(t),\ 0\leqq t<\infty\}$ can be constructed in such a way that, as $n\to\infty$, we have

$$\sup_{1\leqq t<\infty}|n^{-1/2}(S(nt)-W(nt))|/h(t)=o_P(1).$$

The above theorem implies weak convergence of weighted partial sum processes in $D[1,\infty)$. Namely we obtain the following result.

COROLLARY 1.2. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$\mathbf{E}X_1=0, \quad \mathbf{E}X_1^2=1.$$

Let $h\in\mathcal{H}$. Then the following three statements are equivalent:

(a) *There exists a standard Wiener process $\{W(t),\ 0\leqq t<\infty\}$ such that*

$$\sup_{1\leqq t<\infty}|n^{-1/2}(S(nt)-W(nt))|/h(t)=o_P(1)$$

as $n\to\infty$, and

$$\sup_{1\leqq t<\infty}|W(t)|/h(t)<\infty \quad \text{a.s.;}$$

(b) *For all measurable, bounded, continuous functions $g: D[1,\infty)\to\mathbb{R}$, we have*

$$g(n^{-1/2}S(n\cdot)/h(\cdot)) \xrightarrow{\mathcal{D}} g(W(\cdot)/h(\cdot)),$$

as $n\to\infty$, where $\{W(t),\ 0\leqq t<\infty\}$ is a standard Wiener process.

(c)

$$I_\infty(h,c)<\infty \quad \textit{for some} \quad c>0.$$

We note that the class of weight functions in Corollary 1.2 is the largest possible, since in order to have weak convergence at all, the limiting process has to be finite, i.e., we have to assume (1.4) to begin with, in any case. We emphasise again the *lack of analogy* with results on $D[0,1]$. Namely, in order to have weak convergence of $n^{-1/2}S(nt)/q(t)$ in $D[0,1]$ with $q \in Q$, we have to assume (1.1). There is no need for the similar assumption of (1.3) for the weak convergence of weighted partial sum process in $D[1,\infty)$.

In [11] we proved Theorem 1.2 under the additional condition that $h(t)/t^{1/2}$ is slowly varying at infinity (cf. Theorem 3.2 and Corollary 3.2 there). We note that the assumption that $h \in \mathcal{H}$, i.e., that $h(t)/t$ is nonincreasing near infinity, is not really restrictive, since if $h(t)/t$ is increasing there, then it follows from the strong law of large numbers for $W(t)$ that $\lim_{t\to\infty} |W(t)|/h(t) = 0$ a.s..

For optimal weighted L_p-approximations of partial sum processes on $[1,\infty)$, which are complete analogs of those on $[0,1]$, we refer to [11].

Theorem 1.2 will be proven in Section 3.

We wish to note that, even though some parts of the proofs of Theorems 1.1 and 1.2 are similar to parts of the proofs of Theorems 2.2 and 3.2 of [11], we decided to give the complete proofs of our results here for the convenience of the reader, as well as for the sake of clarity of presentation.

2. Proof of results on $[0,1]$

As in the Introduction, let Q be the class of functions q defined on $(0,1]$ which are positive, i.e.,

$$\inf_{\delta \leq t \leq 1} q(t) > 0 \qquad \text{for all } 0 < \delta < 1, \tag{2.1}$$

and nondecreasing in a neighbourhood of zero. Using terminology introduced in [6], such a function q will be called a local function of a standard Wiener process $\{W(t),\ 0 \leq t < \infty\}$ if (1.2) holds.

A local function q of a standard Wiener process W will be called a Chibisov–O'Reilly local function of W if (1.1) holds.

Introduce the following integrals:

$$\mathbf{E}(q,c) = \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1}q^2(t))\,dt,$$

and

$$I(q,c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t))\,dt,$$

for some constant $0 < c < \infty$.

The integral $\mathbf{E}(q,c)$ appeared in the works of Kolmogorov, Petrovski, Erdős and Feller. For details we refer to Itô and McKean ([7], Section 1.8).

The integral $I(q,c)$ appeared in the works of Chibisov [1] and O'Reilly [9].

For further comments on these two integrals, as well as for the proof of the next three theorems, we refer to [2], (cf. also [6]). We have (cf. Proposition 3.1, and Theorems 3.3 and 3.4, respectively, of [2]):

THEOREM 2.A. (i) *Whenever the integral* $I(q,c)<\infty$ *for* $q\in Q$, *then* $\mathbf{E}(q,c+\epsilon)<\infty$ *for every* $\epsilon>0$ *and* $q(t)/t^{1/2}\to\infty$ *as* $t\downarrow 0$.

(ii) *Whenever* $\mathbf{E}(q,c)<\infty$ *and* $q(t)/t^{1/2}\to\infty$ *as* $t\downarrow 0$ *for* $q\in Q$, *then* $I(q,c)<\infty$. □

THEOREM 2.B. *A function* $q\in Q$ *is a local function of a standard Wiener process starting at zero if and only if the integral* $I(q,c)<\infty$ *for some* $c>0$ *or, equivalently, if and only if the integral* $\mathbf{E}(q,c)<\infty$ *for some* $c>0$ *and* $\lim_{t\downarrow 0} q(t)/t^{1/2}=\infty$. □

THEOREM 2.C. *A function* $q\in Q$ *is a Chibisov–O'Reilly local function of a standard Wiener process if and only if the integral* $I(q,c)<\infty$ *for all* $c>0$ *or, equivalently, if and only if the integral* $\mathbf{E}(q,c)<\infty$ *for all* $c>0$ *and* $\lim_{t\downarrow 0} q(t)/t^{1/2}=\infty$. □

REMARK 2.1. Due to Theorem 2.A, the results in Theorem 1.A and Corollary 1.A, as well as those in Theorem 1.1 and Corollary 1.1, stated in terms of the integral $I(q,c)$ can be restated equivalently in terms of the integral $\mathbf{E}(q,c)$.

By Lemma 4.4.4 of Csörgő and Révész [5] (cf. also Section A.2 in Csörgő and Horváth [4]), we can assume without loss of generality that our probability space $(\Omega,\mathcal{A},\mathbf{P})$ accommodates all random variables and stochastic processes introduced so far and later on.

In the proof of Theorem 1.1 we will use the following result of Major [8].

THEOREM 2.D. *Let a distribution* $F(x)$ *be given with* $\int x\,dF(x)=0$, $\int x^2dF(x)=1$. *Define*

$$\sigma_k^2=\int_{-\sqrt{2^n}}^{\sqrt{2^n}} x^2dF(x)-\left(\int_{-\sqrt{2^n}}^{\sqrt{2^n}} x\,dF(x)\right)^2 \quad \text{if } 2^n\le k<2^{n+1},\ n=1,2,\dots .$$

A sequence of i.i.d.r.v.'s $X_1,X_2,\dots$ *with distribution function* $F(x)$ *and a sequence of independent normal random variables* $Y_1,Y_2,\dots$ *with* $\mathbf{E}Y_k=0$, $\mathbf{E}Y_k^2=\sigma_k^2$ *can be constructed in such a way that the partial sums* $S(n)=X_1+\dots+X_n$, $T(n)=Y_1+\dots+Y_n$, $n=1,2,\dots$ *satisfy the relation*

$$|S(n)-T(n)| \stackrel{a.s.}{=} o(n^{1/2}).$$

PROOF OF THEOREM 1.1. Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be as in Theorem 2.D and $\{W(t),\ 0 \leqq t < \infty\}$ be a Wiener process such that

$$W(n) = \sum_{i=1}^{n} Y_i/\sigma_i, \qquad n = 1, 2, \dots. \tag{2.2}$$

Let $T(nt) = \sum_{i=1}^{[nt]} Y_i$, $0 \leqq t \leqq 1$, and $q \in Q$. We have

$$\begin{aligned} &\sup_{1/n<t\leqq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) \\ &\leqq \sup_{1/n\leqq t<1} |n^{-1/2}(S(nt)-T(nt))|/q(t) + \sup_{1/n\leqq t<1} |n^{-1/2}(T(nt)-W(nt))|/q(t) \\ &= I_1(n) + I_2(n). \end{aligned} \tag{2.3}$$

By Theorem 2.D we have

$$|S(nt) - T(nt)| \overset{a.s.}{=} o((nt)^{1/2}), \qquad \text{as } nt \to \infty,$$

and, consequently,

$$\sup_{1\leqq nt<\infty} |S(nt) - T(nt)|/(nt)^{1/2} \overset{a.s.}{=} O(1), \qquad \text{as } n \to \infty.$$

Let $\delta \in (0,1)$ be fixed and n be such that $1/n < \delta$. Then, a.s., as $n \to \infty$,

$$\sup_{1/n\leqq t<\delta} |n^{-1/2}(S(nt) - T(nt))|/q(t) \leqq O(1) \sup_{0<t<\delta} t^{1/2}/q(t). \tag{2.4}$$

Using Theorem 2.D once again, we get

$$\sup_{\delta\leqq t<1} |n^{-1/2}(S(nt) - T(nt))|/q(t) \overset{a.s.}{=} o(1) \tag{2.5}$$

for any $\delta \in (0,1)$. Taking $\delta > 0$ arbitrarily small, by (2.4) and (2.5) we conclude

$$I_1(n) = \sup_{1/n\leqq t\leqq 1} |n^{-1/2}(S(nt) - T(nt))|/q(t) \overset{a.s.}{=} o(1) \tag{2.6}$$

for any $q \in Q$ such that

$$\lim_{t\downarrow 0} t^{1/2}/q(t) = 0. \tag{2.7}$$

Let $\delta > 0$ be small enough, so that q is already nondecreasing on $(0, \delta)$ and let n be such that $1/n < \delta$. By (2.2) we have

$$I_2(n) \leqq \sup_{1/n \leqq t < \delta} \left| n^{-1/2}\Big(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i\Big)\right| \Big/ q(t)$$
$$+ \sup_{\delta \leqq t \leqq 1} \left| n^{-1/2}\Big(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i\Big)\right| \Big/ q(t)$$

$$\leqq \sup_{1/n \leqq t \leqq q(1/n)/n^{1/2}} \left| n^{-1/2} \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| \Big/ q(t) \tag{2.8}$$
$$+ \sup_{q(1/n)/n^{1/2} < t < \delta} \left| n^{-1/2} \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| \Big/ q(t)$$
$$+ \sup_{\delta \leqq t \leqq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| \Big/ q(t)$$
$$= I_2^{(1)}(n) + I_2^{(2)}(n) + I_2^{(3)}(n).$$

Since $\sigma_i \to 1$ as $i \to \infty$, we have $\frac{1}{n} \sum_{i=1}^{n} (\sigma_i - 1)^2 \to 0$ as $n \to \infty$, and by Kolmogorov's inequality we obtain

$$\sup_{0 \leqq t \leqq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| = o_P(1),$$

which implies

$$I_2^{(3)}(n) = \sup_{\delta \leqq t \leqq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| \Big/ q(t) = o_P(1) \tag{2.9}$$

for any $\delta \in (0, 1)$.

In order to show that $I_2^{(1)}(n) = o_P(1)$, we note that due to $q(t)/t^{1/2} \to \infty$ as $t \downarrow 0$ (cf. Theorem 2.A) and $\sigma_i \to 1$ as $i \to \infty$, for any $\varepsilon > 0$ there is n large enough such that

$$\frac{1}{n^{1/2} q(1/n)} \sum_{i=1}^{n^{1/2} q(1/n)} (\sigma_i - 1)^2 \leqq \varepsilon^2.$$

Consequently, using again Kolmogorov's inequality, we have

$$\mathbf{P}\left\{ \sup_{1/n \leqq t \leqq q(1/n)/n^{1/2}} \left| \sum_{i=1}^{[nt]} \Big(1 - \frac{1}{\sigma_i}\Big) Y_i \right| \Big/ n^{1/2} q(t) > \varepsilon \right\}$$

$$\le \mathbf{P}\left\{\sup_{1/n\le t\le q(1/n)/n^{1/2}}\left|\sum_{i=1}^{[nt]}\left(1-\frac{1}{\sigma_i}\right)Y_i\right|\Big/ n^{1/2}q(1/n)>\varepsilon\right\}$$

$$\le \frac{\sum_{i=1}^{n^{1/2}q(1/n)}(\sigma_i-1)^2}{nq^2(1/n)\varepsilon^2}$$

$$\le \frac{\varepsilon^2}{n^{1/2}q(1/n)\varepsilon^2}$$

$$= \frac{1}{n^{1/2}q(1/n)}.$$

Since $q(t)/t^{1/2}\to\infty$ as $t\downarrow 0$, we obtain

$$I_2^{(1)}(n)=o_P(1) \tag{2.10}$$

as $n\to\infty$.

Next we show that $I_2^{(2)}(n)=o_P(1)$. Since $\sigma_i\to 1$ as $i\to\infty$, for any $\varepsilon>0$ there is a large enough n such that $\frac{1}{nt}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\le\varepsilon^2$ whenever $[nt]\ge$ $\ge n^{1/2}q(1/n)$, which gives $q\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\le q(t)$. Next we note that for each $n\ge 1$

$$\left\{n^{-1/2}\sum_{i=1}^{[nt]}\left(1-\frac{1}{\sigma_i}\right)Y_i,\ 0\le t\le 1\right\}\overset{\mathcal{D}}{=}\left\{W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right),\ 0\le t\le 1\right\}$$

as well as for each $\varepsilon>0$

$$\left\{\varepsilon W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right), 0\le t\le 1\right\}\overset{\mathcal{D}}{=}\left\{W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right), 0\le t\le 1\right\}.$$

Hence, we have

$$I_2^{(2)}(n)\overset{\mathcal{D}}{=}\sup_{q(1/n)/n^{1/2}<t<\delta}\left|W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|\Big/q(t)$$

$$\overset{\mathcal{D}}{=}\sup_{q(1/n)/n^{1/2}<t<\delta}\frac{\left|\varepsilon W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{q\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)}\frac{q\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)}{q(t)}$$

$$\leqq \varepsilon \sup_{q(1/n)/n^{1/2}<t<\delta} \frac{\left|W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{q\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)}$$

for any $\varepsilon>0$.

Consequently, by letting $\delta\to 0$ and hence also $n\to\infty$, and combining (1.2) with Theorem 2.B, we arrive at

$$\sup_{q(1/n)/n^{1/2}<t<\delta}\left|W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|\Big/q(t)=\varepsilon O_P(1).$$

Since $\varepsilon>0$ can be taken arbitrarily small, we have

$$I_2^{(2)}(n)=o_P(1). \tag{2.11}$$

Combining now (2.3) with (2.6) and also with (2.8)–(2.11), we get the result.

REMARK 2.2. Given now the embedding theorem, i.e. Theorem 1.1, as far as the intervals $(0,1/n)$ are being concerned, there remains only the Wiener process to be dealt with. But this is exactly what Theorems 2.B and 2.C can be used for. Hence Theorem 2.1 of [11] (cf. Theorem 1.A here) follows from Theorem 1.1 here.

Proof of Corollary 1.1 is similar to that of Theorem 2.3 in [11].

3. Proof of results on $[1,\infty)$

Let $X_1, X_2, \ldots$ be i.i.d.r.v.'s and for each $n\geqq 1$ let $S(nt)=\sum_{i=1}^{[nt]}X_i$, $0\leqq t<\infty$. Let $\{W(t),\ 0\leqq t<\infty\}$ be a standard Wiener process.

A function $h:[1,\infty)\to(0,\infty)$ will be called positive if $\inf_{1\leqq t\leqq K} h(t)>0$ for all $1<K<\infty$.

As in the Introduction, let $\mathcal{H}$ be the class of those positive functions h on $[1,\infty)$ for which $h(t)/t$ is non-increasing in a neighbourhood of infinity. A function $h\in\mathcal{H}$ will be called a global function of a standard Wiener process $\{W(t),\ 0\leqq t<\infty\}$ if

$$\limsup_{t\to\infty}|W(t)|/h(t)<\infty \quad \text{a.s.}.$$

Introduce the following integrals:

$$\mathbf{E}_\infty(h,c)=\int_1^\infty t^{-3/2}h(t)\exp(-ct^{-1}h^2(t))dt$$

and

$$I_\infty(h,c)=\int_1^\infty t^{-1}\exp(-ct^{-1}h^2(t))dt,$$

where $0<c<\infty$.

For a global description of the behaviour of a Wiener process near infinity, as well as for the following two results which are analogs of Theorem 2.B and 2.C for the case of $t\to\infty$, we refer to [6].

THEOREM 3.B*. *A function $h\in\mathcal{H}$ is a global function of a standard Wiener process if and only if the integral $I_\infty(h,c)<\infty$ for some $c>0$ or, equivalently, if and only if the integral $\mathbf{E}_\infty(h,c)<\infty$ for some $c>0$ and $\lim_{t\to\infty} h(t)/t^{1/2}=\infty$.*

THEOREM 3.C*. *Let $h\in\mathcal{H}$ and W be a standard Wiener process. Then*

$$\lim_{t\to\infty}|W(t)|/h(t)=0 \quad a.s.$$

if and only if the integral $I_\infty(h,c)<\infty$ for all $c>0$ or, equivalently, if and only if the integral $\mathbf{E}_\infty(h,c)<\infty$ for all $c>0$ and $\lim_{t\to\infty} h(t)/t^{1/2}=\infty$.

REMARK 3.1. If $q\in Q$ then $q(1/t)$ is well defined for $t\in[1,\infty)$, positive and non-increasing in t as $t\to\infty$. Hence $tq(1/t)\in\mathcal{H}$, and our results on $[0,1]$ and on $[1,\infty)$ can be stated in terms of the integral $I(q,c)$ (or $\mathbf{E}(q,c)$) for both cases (cf. also [6]).

PROOF OF THEOREM 1.2. Let $X_1,X_2,\ldots$ and $Y_1,Y_2,\ldots$ be as in Theorem 2.D and $\{W(t),\ 0\leqq t<\infty\}$ be a Wiener process such that

$$W(n)=\sum_{i=1}^n Y_i/\sigma_i, \qquad n=1,2,\ldots\ .$$

By Theorem 2.D, with $T(nt)=\sum_{i=1}^{[nt]} Y_i$, $0\leqq t<\infty$, we have

$$|S(nt)-T(nt)|=o((nt)^{1/2}) \quad \text{a.s.}$$

as $nt\to\infty$. Hence

$$\sup_{1\leqq t<\infty}|n^{-1/2}(S(nt)-T(nt))|/t^{1/2}=O(1) \quad \text{a.s.}$$

as $n \to \infty$, and for any $1 < K < \infty$, we have as $n \to \infty$

$$\sup_{1 \leqq t \leqq K} |n^{-1/2}(S(nt) - T(nt))|/h(t) = o(1) \quad \text{a.s.},$$

as well as

$$\sup_{K < t < \infty} |n^{-1/2}(S(nt) - T(nt))|/h(t) \leqq O(1) \sup_{K < t < \infty} t^{1/2}/h(t) \quad \text{a.s.}.$$

Consequently, taking K arbitrarily large, we obtain, as $n \to \infty$

$$\sup_{1 \leqq t < \infty} |n^{-1/2}(S(nt) - T(nt))|/h(t) = o(1) \quad \text{a.s.} \tag{3.1}$$

for any $h : [1, \infty) \to (0, \infty)$ which is positive and such that $\lim_{t \to \infty} t^{1/2}/h(t) = 0$. In particular for $h \in \mathcal{H}$ and such that $I_\infty(h, c) < \infty$ for some $c > 0$, the latter is true.

Next we have, for any $1 \leqq K < \infty$

$$\begin{aligned} &\sup_{1 < t < \infty} |n^{-1/2}(T(nt) - W(nt))|/h(t) \\ &\leqq \sup_{1 \leqq t \leqq K} |n^{-1/2}(T(nt) - W(nt))|/h(t) \\ &\quad + \sup_{K < t < \infty} |n^{-1/2}(T(nt) - W([nt]))|/h(t) \\ &\quad + \sup_{K < t < \infty} |n^{-1/2}(W(nt) - W([nt]))|/h(t) \\ &= \mathcal{I}_1(n) + \mathcal{I}_2(n) + \mathcal{I}_3(n). \end{aligned} \tag{3.2}$$

Since

$$T(nt) - W(nt) = \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i$$

and $\sigma_i \to 1$ as $i \to \infty$, by Kolmogorov's inequality, we have for any positive function $h(t) : [1, \infty) \to (0, \infty)$, as $n \to \infty$

$$\mathcal{I}_1(n) = o_P(1). \tag{3.3}$$

Let

$$\beta \overset{a.s.}{=} \limsup_{t \uparrow \infty} \frac{|W(t)|}{h(t)} = \lim_{T \to \infty} \sup_{T \leqq t < \infty} \frac{|W(t)|}{h(t)}, \tag{3.4}$$

due to

$$I_\infty(h, c) < \infty \quad \text{for some } c > 0.$$

We note that for each $n \geqq 1$

$$\left\{ n^{-1/2}\left(\sum_{i=1}^{[nt]}\left(1-\frac{1}{\sigma_i}\right)Y_i\right),\ 0 \leqq t < \infty \right\}$$

$$\stackrel{\mathcal{D}}{=} \left\{ W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right),\ 0 \leqq t < \infty \right\}.$$

Given $\varepsilon > 0$, however small, then on account of $\sigma_i \to 1$ as $i \to \infty$, we can take K large enough so that $\frac{1}{nt}\sum_{i=1}^{[nt]}(\sigma_i-1)^2 \leqq \varepsilon^2$ whenever $[nt] \geqq Kn$. Taking K even bigger if necessary, so that $h(t)$ is increasing for $\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2 < t < \infty$, and hence, in particular, $h\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right) \leqq h(t)$, we arrive at

$$\sup_{K<t<\infty}\left| n^{-1/2}\left(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i\right)\right| \Big/ h(t)$$

$$\stackrel{\mathcal{D}}{=} \sup_{K<t<\infty}\left| W\left(\frac{1}{n}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right| \Big/ h(t)$$

$$\stackrel{\mathcal{D}}{=} \sup_{K<t<\infty} \frac{\left|\varepsilon W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{h\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)} \frac{h\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)}{h(t)}$$

$$\leqq \varepsilon \sup_{K<t<\infty} \frac{\left| W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right|}{h\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)}.$$

Consequently, using (3.4), we have for any $\varepsilon > 0$

$$\sup_{K<t<\infty}\left|\varepsilon W\left(\frac{1}{n\varepsilon^2}\sum_{i=1}^{[nt]}(\sigma_i-1)^2\right)\right| \Big/ h(t) \stackrel{a.s.}{=} \beta\varepsilon O(1)$$

as $K \to \infty$.

Since $\varepsilon > 0$ can be taken arbitrarily small, we obtain

$$\mathcal{I}_2(n) = o_P(1). \tag{3.5}$$

On account of having, as $nt \to \infty$,

$$|W(nt) - W([nt])| = O((\log nt)^{1/2}) \quad \text{a.s.}$$

we have also

$$\sup_{1 \leq t < \infty} |n^{-1/2}(W(nt) - W([nt]))|/t^{1/2} = o(1) \quad \text{a.s.}$$

as $n \to \infty$. Consequently, for any $1 < K < \infty$, we get as $n \to \infty$

$$\begin{aligned}\sup_{K < t < \infty} & |n^{-1/2}(W(nt) - W([nt]))|/h(t) \\ & \leq o(1) \sup_{K < t < \infty} t^{1/2}/h(t) \qquad \text{a.s.,}\end{aligned}$$

which gives

$$\mathcal{I}_3(n) = o(1) \quad \text{a.s.} \tag{3.6}$$

for any h positive and such that $\limsup_{t \to \infty} t^{1/2}/h(t) < \infty$.

Combining now (3.3), (3.5) and (3.6), we obtain the result.

ACKNOWLEDGEMENTS. The author gratefully expresses her gratitude to Professor Miklós Csörgő for his valuable comments and suggestions.

REFERENCES

[1] CHIBISOV, D., Some theorems on the limiting behaviour of empirical distribution functions, *Selected Transl. Math. Statist. Prob.* **6** (1964), 147–156.

[2] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[3] CSÖRGŐ, M. and HORVÁTH, L., Nonparametric methods for changepoint problems, *Handbook of Statistics*, Elsevier Science Publishers B.V., North-Holland, 1988, 403–425.

[4] CSÖRGŐ, M. and HORVÁTH, L., *Weighted approximations in probability and statistics*, Wiley series in probability and mathematical statistics: Probability and mathematical statistics, Wiley, Chichester, 1993. *MR* **94c**:60060

[5] CSÖRGŐ, M. and RÉVÉSZ, P., *Strong approximations in probability and statistics*, Academic Press, New York, 1981. *MR* **84d**:60050

[6] CSÖRGŐ, M., SHAO, Q. M. and SZYSZKOWICZ, B., A note on local and global functions of a Wiener process and some Rényi-type statistics, *Studia Sci. Math. Hungar.* **26** (1991), 239–259. *MR* **93i**:60078

[7] ITÔ, K. and MCKEAN, H. P. JR., *Diffusion processes and their sample paths*, Academic Press, New York; Springer-Verlag, Berlin, 1965. *MR* **33** #8031

[8] MAJOR, P., An improvement of Strassen's invariance principle, *Ann. Probab.* **7** (1979), 55–61. *MR* **82f**:60069

[9] O'REILLY, N., On the weak convergence of empirical processes in sup-norm metrics, *Ann. Probab.* **2** (1974), 642–651. *MR* **52** #4367

[10] SZYSZKOWICZ, B., L_p-approximation of weighted partial sum processes, *Stochastic Process Appl.* **45** (1993), 295–308. *MR* **94c**:60063

[11] SZYSZKOWICZ, B., Weighted approximations of partial sum processes in $D[0, \infty)$. I., *Studia Sci. Math. Hungar.* **31** (1996), 323–353.

(Received February 20, 1995)

DEPARTMENT OF MATHEMATICS AND STATISTICS
4302 HERZBERG BUILDING
CARLETON UNIVERSITY
1125 COLONEL BY DRIVE
OTTAWA, ONTARIO
K1S 5B6
CANADA

gmurray@math.carleton.ca

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 321–337

LIMIT THEOREMS FOR WEAKLY REINFORCED RANDOM WALKS ON $\mathbb{Z}$

B. TÓTH

Dedicated to Endre Csáki on his 60-th birthday

Abstract

The weakly reinforced random walk (WRRW) on the one-dimensional integer lattice $\mathbb{Z}$ starts from the origin of the lattice and at each step it jumps to a neighbouring site, the probability of jumping along a bond being proportional to w (number of previous jumps along that lattice bond), where $w:\mathbb{N}\to\mathbb{R}_+$, with $w(n)\sim n^\alpha$ for large n, and $\alpha\in(0,1)$ is a fixed parameter. We prove that the properly scaled local time process of WRRW converges *in probability* to a deterministic function. Using this result we also prove a limit theorem for the position of the random walker at late times.

1. Introduction

We continue to investigate the long time asymptotic behaviour of self-interacting random walks on the one-dimensional integer lattice $\mathbb{Z}$. The walk X_i, $i=0,1,2,\dots$ starts from the origin of the lattice and at time $i+1$ it jumps to one of the two neighbouring sites of X_i, so that the probability of jumping along a bond of the lattice is proportional to

$$w \text{ (number of previous jumps along that bond)}$$

where

$$w:\mathbb{N}\to\mathbb{R}_+$$

is a weight function to be specified later. Formally, for a nearest neighbour walk $\underline{x}_0^i=(x_0,x_1,\dots,x_i)$ we define

$$(1.1)\qquad r(\underline{x}_0^i)=\#\{0\leqq j<i:\,(x_j,x_{j+1})=(x_i,x_i+1)\text{ or }(x_i+1,x_i)\}$$

$$(1.2)\qquad l(\underline{x}_0^i)=\#\{0\leqq j<i:\,(x_j,x_{j+1})=(x_i,x_i-1)\text{ or }(x_i-1,x_i)\}\,.$$

1991 *Mathematics Subject Classification.* Primary 60F05, 60J15, 60J55; Secondary 82C41.

Key words and phrases. Reinforced random walk, local time, limit theorems.

Work partially supported by the Hungarian National Foundation for Scientific Research Grants No. 7275 and 9002.

That is: the number $r(\underline{x}_0^i)$ (respectively, $l(\underline{x}_0^i)$) shows how many times has the walk $\underline{x}_0^i$ visited the edge adjacent from the right (respectively, from the left) to the terminal site x_i. The random walk X_i is governed by the law:

$$\begin{aligned}\mathbf{P}\Big(X_{i+1}=X_i+1\Big\|\underline{X}_0^i=\underline{x}_0^i\Big) &= \frac{w(r(\underline{x}_0^i))}{w(r(\underline{x}_0^i))+w(l(\underline{x}_0^i))} \\ &= 1-\mathbf{P}\Big(X_{i+1}=X_i-1\Big\|\underline{X}_0^i=\underline{x}_0^i\Big).\end{aligned} \tag{1.3}$$

The long time asymptotic behaviour of the random walk X_i depends strongly on the choice of the weight function $w(\cdot)$. In three previous papers we analyzed the following cases:

(1) The so-called *'true' self-avoiding walk*, with $w(n)=\exp(-g\cdot n)$, $g>0$, was studied in [6]. There we showed that for long times X_n scales as $n^{2/3}$ and we proved a limit theorem for $A^{-2/3}X_{\theta_{s/A}}$, as $A\to\infty$, where $\theta_{s/A}$ is a geometrically distributed random variable with distribution $\mathbf{P}(\theta_{s/A}=n)=(1-\exp\{-s/A\})\exp\{-ns/A\}$, independent of the random walk X_n.

(2) The *generalized 'true' self-avoiding walk*, a generalization of the previous model, with subexponential self-repulsion $w(n)=\exp(-g\cdot n^\kappa)$, $g>0$, $\kappa\in(0,1)$ was investigated in [5]. In this case we found that X_n scales as $n^{(\kappa+1)/(\kappa+2)}$ and we proved a limit theorem for $A^{-(\kappa+1)/(\kappa+2)}X_{\theta_{s/A}}$, as $A\to\infty$.

(3) Finally, in [7] weight functions with power-law asymptotics were considered: the so-called *polynomially self-repelling walks*, with $w(n)\sim n^{-\alpha}$, $\alpha>0$, respectively, the *asymptotically free walks*, with $w(n)=1+O(n^{-1})$ asymptotically, for $n\gg 1$. In these cases the correct scaling of X_n was $n^{-1/2}$ (as for ordinary random walks) but the scaling limit was *not* gaussian. A particular case of asymptotically free walks, the *once reinforced random walk* or *random walk partially reflected/attracted at its extrema* has also been considered in [2].

In the present paper we consider self-interacting random walks with *polynomial self-attraction*. That is: we assume that the weight function $w:\mathbb{N}\to(0,\infty)$ is *monotone increasing* and for large values of $n\in\mathbb{N}$ it obeys the asymptotics

$$w(n)=(1-\alpha)^{-1}\left(\frac{n}{2}\right)^{\alpha}-B(1-\alpha)^{-2}\left(\frac{n}{2}\right)^{\alpha-1}+O\left(n^{\alpha-2}\right), \tag{1.4}$$

or, equivalently

$$w(n)^{-1}=(1-\alpha)\left(\frac{n}{2}\right)^{-\alpha}+B\left(\frac{n}{2}\right)^{-1-\alpha}+O\left(n^{-2-\alpha}\right), \tag{1.5}$$

where $\alpha\in(0,1)$ and $B\in\mathbb{R}$ are fixed constant parameters. Since in the definition (1.3) of jump probabilities only ratios of w-s play any role, the constant factor in front of the leading term is chosen for convenience only. Note that the next-to-leading term is assumed asymptotically 'smooth'.

We call these walks *weakly reinforced random walks* (WRRW) since the self-attraction of trajectories is slightly weaker than in the linearly reinforced case (with $w(n) = 1 + Bn$, $B > 0$). According to Davis [1] self-attracting random walks on $\mathbb{Z}$ are recurrent if and only if $\sum_{n=0}^{\infty} w(n)^{-1} = \infty$, otherwise the random walker eventually sticks to one (randomly selected) edge of the lattice, jumping back and forth on it indefinitely. Pemantle [4] proved that the linearly reinforced random walk has an asymptotic distribution on $\mathbb{Z}$ without any scaling. These two remarks explain why we confine our investigations to $\alpha \in (0,1)$ in (1.4), (1.5).

The paper is organized as follows: In Section 2 we formulate our main results: Theorem 1 describes the asymptotics of the local time process of WRRW, Theorem 2 is a limit theorem for the position of the WRRW at late times. In Section 3 we give a representation of the local time process of our random walks in terms of generalized Pólya Urn Schemes. Section 4 is devoted to the proof of Theorem 1. As the proof of Theorem 2 is identical to a similar proof in [7], we do not repeat those details here.

In order to keep the paper self-contained we had to include parts of our previous paper [7]. These overlapping parts are typed with petite and thus, they are clearly distinguishable from the genuinely new parts.

2. Results

The present section is divided in two subsections: in subsection 2.1 we formulate the limit theorems referring to the local time processes and hitting times of the WRRW. In subsection 2.2 we formulate the limit theorems for the position of the SIRW at late times.

2.1. The local time process and hitting times

We define the following (bond) local time process:

(2.1.1) $$L(l,i) = \#\{0 \leqq j < i : X_j = l,\ X_{j+1} = l-1\}, \quad l \in \mathbb{Z},\ i \in \mathbb{N}$$

and stopping times

(2.1.2) $T^{>}_{k,-1} = 0$, $T^{>}_{k,m} = \inf\{i > T^{>}_{k,m-1} : X_{i-1} = k-1,\ X_i = k\}$ $k > 0$, $m \geqq 0$.

(2.1.3) $T^{<}_{k,0} = 0$, $T^{<}_{k,m} = \inf\{i > T^{<}_{k,m-1} : X_{i-1} = k+1,\ X_i = k\}$ $k \geqq 0$, $m \geqq 1$.

In plain words: $L(l,i)$ is the number of leftwards jumps on the bond $l \to l-1$ performed by the random walk up to time i. $T^{>}_{k,m}$ is the time of the $m+1$-th arrival to the lattice site k coming from left, $T^{<}_{k,m}$ is the time of the m-th arrival to the lattice site k coming from right.

In formula (2.1.4) below and thereafter the superscript $*$ stands for either $<$ or $>$. We consider the following shifted (bond) local time processes of the walk stopped at $T^*_{k,m}$:

$$S^*_{k,m}(l) = L(k-l, T^*_{k,m}). \tag{2.1.4}$$

$S^*_{k,m}(l)$ is roughly half of the total number of jumps across the bond $\{k-l-1, k-l\}$:

$$\#\{0 \leqq j < T^*_{k,m} : \{X_j, X_{j+1}\} = \{k-l-1, k-l\}\} = 2S^*_{k,m}(l) + \mathbf{1}_{[0,k)}(l). \tag{2.1.5}$$

Denote

$$\omega^{*-}_{k,m} = \omega^{-}\left(S^*_{k,m}\right) = \inf\{l \leqq 0 : S^*_{k,m}(l) > 0\}, \tag{2.1.6}$$

$$\omega^{*+}_{k,m} = \omega^{+}\left(S^*_{k,m}\right) = \sup\{l \geqq k : S^*_{k,m}(l) > 0\}. \tag{2.1.7}$$

In plain words: $k - \omega^{*+}_{k,m}$, respectively $k - \omega^{*-}_{k,m} - 1$, is the leftmost, respectively rightmost, site visited by the stopped walk $\underline{X}_0^{T^*_{k,m}}$.

From (2.1.5) it clearly follows that

$$T^*_{k,m} = 2 \sum_{l=\omega^{*-}_{k,m}}^{\omega^{*+}_{k,m}} S^*_{k,m}(l) + k = 2 \sum_{l=-\infty}^{\infty} S^*_{k,m}(l) + k. \tag{2.1.8}$$

Looking at the formal definitions only, in principle, these local times or hitting times might be infinite, i.e. it could happen that the site $k \in \mathbb{Z}$ is never hit. From the results of Davis [1] it follows that in case of WRRW-s considered in the present paper, with probability one, this does *not* happen: all the random variables defined above are finite almost surely.

The following theorem and its corollary describes the precise asymptotics of the *local time processes* $S^*_{k,m}(\cdot)$ and *hitting times* $T^*_{k,m}$ of WRRW:

THEOREM 1. *The sum*

$$D = \sum_{j=0}^{\infty} \left(\frac{1}{w(2j)} - \frac{1}{w(2j+1)} \right) \tag{2.1.9}$$

exists and $D \in (0, \infty)$.

Let $x \in [0, \infty)$, $h \geqq 0$ *and* $* = <$ *or* $>$ *be fixed.*

$$A^{-1} \omega^{*-}_{[Ax],[A^{1/(1-\alpha)}h]} \xrightarrow{\mathbf{P}} -D^{-1}h^{1-\alpha} \tag{2.1.10}$$

$$A^{-1} \omega^{*+}_{[Ax],[A^{1/(1-\alpha)}h]} \xrightarrow{\mathbf{P}} 2x + D^{-1}h^{1-\alpha} \tag{2.1.11}$$

(2.1.12)

$$\sup_y \left| A^{-1/(1-\alpha)} S^*_{[Ax],[A^{1/(1-\alpha)}h]}([Ay]) - \{h^{1-\alpha} + D(x - |y-x|)\}_+^{1/(1-\alpha)} \right| \xrightarrow{\mathbf{P}} 0$$

as $A \to \infty$.

REMARKS. Note that the non-trivial scaling of the local time process provides *convergence in probability to a deterministic function rather than convergence in distribution to a genuinely stochastic process.*

From the previous theorem and (2.1.8) it follows immediately:

COROLLARY 1. *Let x, h and $*$ and D be as in Theorem 1.*

(2.1.13)
$$\frac{1}{2}A^{-(2-\alpha)/(1-\alpha)}T^*_{[Ax],[A^{1/(1-\alpha)}h]} \xrightarrow{\mathbf{P}} \frac{2-2\alpha}{2-\alpha}D\Big(x+(D^{-1}h)^{1-\alpha}\Big)^{(2-\alpha)/(1-\alpha)}$$

as $A \to \infty$.

2.2. Limit theorem for the position at late times

The second result concerns the limiting distribution of the WRRW X_n for late times. We denote by $P(n,k)$, $n \in \mathbb{N}$, $k \in \mathbb{Z}$ the distribution of our WRRW at time n:

$$P(n,k) = \mathbf{P}\Big(X_n = k\Big) \tag{2.2.1}$$

and by $R(s,k)$, $s \in \mathbb{R}_+$, $k \in \mathbb{Z}$ the distribution of the walk observed at an independent random time θ_s, of geometric distribution

$$\mathbf{P}\Big(\theta_s = n\Big) = \left(1-e^{-s}\right)e^{-sn}, \tag{2.2.2}$$

$$R(s,k) = \mathbf{P}\Big(X_{\theta_s} = k\Big) = \left(1-e^{-s}\right)\sum_{n=0}^{\infty} e^{-sn}P(n,k). \tag{2.2.3}$$

We define the following rescaled 'densities' of the above distributions

$$\pi_A(t,x) = A^{(1-\alpha)/(2-\alpha)}P([At],[A^{(1-\alpha)/(2-\alpha)}x]) \tag{2.2.4}$$

$$\widehat{\pi}_A(s,x) = A^{(1-\alpha)/(2-\alpha)}R(A^{-1}s,[A^{(1-\alpha)/(2-\alpha)}x]) \tag{2.2.5}$$

$t, s \in \mathbb{R}_+$, $x \in \mathbb{R}$.

THEOREM 2. *For any $s \in \mathbb{R}_+$ and $x \in \mathbb{R}$*

$$\widehat{\pi}_A(s,x) \to \widehat{p}_D^{(\alpha)}(s,x) \tag{2.2.6}$$

as $A \to \infty$, where

$$\widehat{p}_D^{(\alpha)}(s,x) = s\int_0^{\infty} e^{-st}p_D^{(\alpha)}(t,x)dt \tag{2.2.7}$$

and

(2.2.8)
$$p_D^{(\alpha)}(t,x)=\frac{1}{2-2\alpha}\left(\frac{2-2\alpha}{2-\alpha}\frac{D}{t}\right)^{1/(2-\alpha)}\left\{\left(\frac{2-\alpha}{2-2\alpha}\frac{t}{D}\right)^{(1-\alpha)/(2-\alpha)}-|x|\right\}_+^{\alpha/(1-\alpha)}.$$

This is of course a *local limit theorem* for the WRRW, observed at an independent random time $\theta_{s/A}$ of geometric distribution with mean $e^{-s/A}\times$ $\times\left(1-e^{-s/A}\right)^{-1}\sim A/s$. In particular the (integral) limit law follows:

$$\mathbf{P}\left(A^{-(1-\alpha)/(2-\alpha)}X_{\theta_{s/A}}<x\right)\to\int_{-\infty}^{x}\widehat{p}_D^{(\alpha)}(s,y)dy. \tag{2.2.9}$$

This is a little bit short of stating the limit theorem for *deterministic time*:

$$\mathbf{P}\left(A^{-(1-\alpha)/(2-\alpha)}X_{[At]}<x\right)\to\int_{-\infty}^{x}p_D^{(\alpha)}(t,y)dy. \tag{2.2.10}$$

But, of course, we can conclude that if the sequence $A^{-(1-\alpha)/(2-\alpha)}X_{[At]}$, with $t\in\mathbb{R}_+$ fixed and $A\to\infty$, converges in distribution then (2.2.10) also holds.

REMARK. On the other hand we have good reason to expect that the sequence of *random processes* $t\mapsto X^{(A)}(t)=A^{-(1-\alpha)/(2-\alpha)}X_{[At]}$ is not tight in the function space $D[0,1]$ and there is no continuous limit *process*.

Given Corollary 1, the proof of Theorem 2 is formally identical to the proof of Theorem 3 in [7]. We omit the repetition of those details here.

3. Representation of the local time process in terms of Pólya urns

3.1. Generalized Pólya urn schemes

Given two weight functions

$$r:\mathbb{N}\to\mathbb{R}_+ \tag{3.1.1}$$
$$b:\mathbb{N}\to\mathbb{R}_+, \tag{3.1.2}$$

a generalized Pólya Urn Scheme is a Markov chain (ρ_i,β_i) on $\mathbb{N}\times\mathbb{N}$ with transition probabilities

$$\mathbf{P}\Big((\rho_{i+1},\beta_{i+1})=(k+1,l)\,\Big\|\,(\rho_i,\beta_i)=(k,l)\Big)=\frac{r(k)}{r(k)+b(l)}, \tag{3.1.3}$$

$$\mathbf{P}\Big((\rho_{i+1},\beta_{i+1})=(k,l+1)\,\Big\|\,(\rho_i,\beta_i)=(k,l)\Big)=\frac{b(l)}{r(k)+b(l)}, \tag{3.1.4}$$

and no other transitions allowed. Usually the initial values $(\rho_0, \beta_0) = (0,0)$ are assumed and β_i and ρ_i are interpreted as the number of blue, respectively red marbles drawn from the urn up to time i. Denote by τ_m the time when the m-th red marble is drawn and by $\mu(m)$ the number of blue marbles drawn before the m-th red one:

$$\tau_m = \min\{\, i \mid \rho_i = m \,\}, \tag{3.1.5}$$

$$\mu(m) = \beta_{\tau_m}. \tag{3.1.6}$$

The functions defined below are essential in the study of the Pólya Urn Scheme defined above:

$$R_p(n) = \sum_{j=0}^{n-1} (r(j))^{-p}, \quad p \in \mathbb{N} \tag{3.1.7}$$

$$B_p(n) = \sum_{j=0}^{n-1} (b(j))^{-p}, \quad p \in \mathbb{N}. \tag{3.1.8}$$

We shall be particularly interested in $p = 1, 2$.

LEMMA 1. *For any $m \in \mathbb{N}$ and $\lambda < \min\{r(j) : 0 \leqq j \leqq m-1\}$ the following identity holds:*

$$\mathbf{E}\Big(\prod_{j=0}^{\mu(m)-1} \Big(1 + \frac{\lambda}{b(j)}\Big)\Big) = \prod_{j=0}^{m-1} \Big(1 - \frac{\lambda}{r(j)}\Big)^{-1}. \tag{3.1.9}$$

In particular,

$$\mathbf{E}\Big(B_1(\mu(m))\Big) = R_1(m) \tag{3.1.10}$$

$$\mathbf{E}\Big(\Big[B_1(\mu(m)) - \mathbf{E}B_1(\mu(m))\Big]^2\Big) = R_2(m) + \mathbf{E}\Big(B_2(\mu(m))\Big). \tag{3.1.11}$$

PROOF. The proof of (3.1.9) follows from standard martingale considerations, using the representation of the generalized Pólya Urn Scheme in terms of two independent renewal processes with exponentially distributed waiting times (see e.g. the Appendix of [1]). Expanding (3.1.9) to second order in λ yields (3.1.10) and (3.1.11). We leave the standard details of this proof as an exercise for the reader. □

3.2. The local time process

For sake of definiteness we consider the case of superscript $>$, i.e. we stop the WRRW at the hitting time $T^{>}_{k,m}$. The case of superscript $<$ is done in a very similar way, with straightforward slight changes.

Let $(\rho_i^{(l)}, \beta_i^{(l)})$, $l \in \mathbb{Z}$ be *independent* Pólya Urn Schemes with weight functions

$$r^{(l)}(j) = w(2j+1) \quad b^{(l)}(j) = w(2j) \qquad \text{for } l \in (-\infty, 0] \cup [k+1, \infty) \tag{3.2.1}$$

$$r^{(l)}(j) = w(2j) \qquad b^{(l)}(j) = w(2j+1) \quad \text{for } l \in [1, k-1] \tag{3.2.2}$$

$$r^{(l)}(j) = w(2j) \qquad b^{(l)}(j) = w(2j) \qquad \text{for } l = k. \tag{3.2.3}$$

Denote by $\mu^{(l)}(m)$ the random variables defined in (3.1.6), the superscript l showing to which of the Urn Schemes it belongs.

The extension to *self-interacting walks* of F. Knight's description [3] of the local time process $S^{>}_{k,m}(l)$, $l \in \mathbb{Z}$ *as a Markov chain* is formally exhaustively explained in [6]. According to these arguments $S^{>}_{k,m}(l)$, $l \in \mathbb{Z}$ is obtained by patching together three homogeneous Markov chains in the following way:

(I) In the interval $l \in (0, k-2)$, that is steps $0 \to 1, 1 \to 2, \ldots, (k-2) \to (k-1)$:

$$S^{>}_{k,m}(0) = m, \; S^{>}_{k,m}(l+1) = \mu^{(l+1)}\left(S^{>}_{k,m}(l) + 1\right), \; l = 0, 1, \ldots, k-2. \tag{3.2.4}$$

(II) The single step $(k-1) \to k$ is exceptional

$$S^{>}_{k,m}(k-1) = \text{ given by (3.2.4)}, \quad S^{>}_{k,m}(k) = \mu^{(k)}\left(S^{>}_{k,m}(k-1) + 1\right). \tag{3.2.5}$$

(III) In the intervals $l \in (-\infty, 0)$, respectively $l \in (k+1, \infty)$, that is steps $0 \to -1, -1 \to -2, -2 \to -3, \ldots$, respectively $k \to (k+1)$, $(k+1) \to (k+2)$, $(k+2) \to (k+3), (k+3) \to (k+4), \ldots$:

$$S^{>}_{k,m}(0) = m, \quad S^{>}_{k,m}(l-1) = \mu^{(l)}\left(S^{>}_{k,m}(l)\right), \quad l = 0, -1, -2, \ldots \tag{3.2.6}$$

respectively

$$S^{>}_{k,m}(k) = \text{ given by (3.2.5)}, \quad S^{>}_{k,m}(l+1) = \mu^{(l+1)}\left(S^{>}_{k,m}(l)\right), \quad l = k, k+1, k+2, \ldots. \tag{3.2.7}$$

Due to (3.2.1) these last two Markov chains have the same transition laws.

4. Proof of Theorem 1

4.1. Preparations

As suggested by the representation of the local times given in the previous section, we consider two homogeneous Markov chains $\mathcal{Z}(l)$ and $\widetilde{\mathcal{Z}}(l)$, $l = 0, 1, 2, \ldots$ on the state space $\mathbb{N}$, defined as follows:

$$\mathcal{Z}(l+1) = \mu^{(l+1)}(\mathcal{Z}(l) + 1), \qquad \widetilde{\mathcal{Z}}(l+1) = \widetilde{\mu}^{(l+1)}(\widetilde{\mathcal{Z}}(l)) \tag{4.1.1}$$

where the processes $\left\{\mu^{(l)}(\cdot)\right\}_{l \in \mathbb{N}}$ are those defined in (3.1.5)-(3.1.6), belonging to i.i.d. Pólya Urn Schemes $\left\{(\rho_i^{(l)}, \beta_i^{(l)})\right\}_{l \in \mathbb{N}}$, with weight functions

$$r(j) = w(2j), \quad b(j) = w(2j+1) \tag{4.1.2}$$

and similarly, the processes $\left\{\widetilde{\mu}^{(l)}(\cdot)\right\}_{l \in \mathbb{N}}$ belong to i.i.d. Pólya Urn Schemes $\left\{(\widetilde{\rho}_i^{(l)}, \widetilde{\beta}_i^{(l)})\right\}_{l \in \mathbb{N}}$ with weight functions

$$\widetilde{r}(j) = w(2j+1), \quad \widetilde{b}(j) = w(2j). \tag{4.1.3}$$

We shall also need the hitting time

(4.1.4) $$\widetilde{\sigma}_0 = \widetilde{\sigma}_0\left(\widetilde{\mathcal{Z}}(\cdot)\right) = \inf\{l : \widetilde{\mathcal{Z}}(l) = 0\}.$$

From (4.1.1) and (3.1.5)–(3.1.6) we see that $\widetilde{\sigma}_0$ is actually the extinction time of $\widetilde{\mathcal{Z}}(\cdot)$:

(4.1.5) $$\widetilde{\mathcal{Z}}(l) \equiv 0 \quad \text{for } l \geqq \widetilde{\sigma}_0.$$

Lemma 1 suggests the introduction of the following functions:

(4.1.6) $$U_p(n) = \sum_{j=0}^{n-1} (w(2j))^{-p}, \qquad p = 1, 2,$$

(4.1.7) $$V_p(n) = \sum_{j=0}^{n-1} (w(2j+1))^{-p}, \quad p = 1, 2.$$

Using formulas (3.1.10) and (3.1.11) of Lemma 1 and the functions introduced above we get the following identities:

(4.1.8) $$\mathbf{E}(V_1(\mathcal{Z}(l+1)) \,\|\, \mathcal{Z}(l) = n) = U_1(n+1)$$

(4.1.9) $$\mathbf{D}^2(V_1(\mathcal{Z}(l+1)) \,\|\, \mathcal{Z}(l) = n) = U_2(n+1) + \mathbf{E}(V_2(\mathcal{Z}(l+1)) \,\|\, \mathcal{Z}(l) = n)$$

(4.1.10) $$\mathbf{E}(U_1(\widetilde{\mathcal{Z}}(l+1)) \,\|\, \widetilde{\mathcal{Z}}(l) = n) = V_1(n)$$

(4.1.11) $$\mathbf{D}^2(U_1(\widetilde{\mathcal{Z}}(l+1)) \,\|\, \widetilde{\mathcal{Z}}(l) = n) = V_2(n) + \mathbf{E}(U_2(\widetilde{\mathcal{Z}}(l+1)) \,\|\, \widetilde{\mathcal{Z}}(l) = n).$$

As both functions $n \mapsto U_1(n)$ and $n \mapsto V_1(n)$ are *bijections* between $\mathbb{N}$ and their ranges it is more convenient to consider the Markov chains

(4.1.12) $$\mathcal{Y}(l) = V_1(\mathcal{Z}(l)), \qquad \widetilde{\mathcal{Y}}(l) = U_1(\widetilde{\mathcal{Z}}(l)), \quad l = 0, 1, 2 \dots$$

instead of $\mathcal{Z}(l)$, respectively $\widetilde{\mathcal{Z}}(l)$. With this change of variable formulas (4.1.8)–(4.1.11) transform as follows:

(4.1.13)
$$\mathbf{E}(\mathcal{Y}(l+1) \| \mathcal{Y}(l) = x) = U_1\left(V_1^{-1}(x) + 1\right)$$

(4.1.14)
$$\mathbf{D}^2(\mathcal{Y}(l+1) \| \mathcal{Y}(l) = x) = U_2\left(V_1^{-1}(x) + 1\right) + \mathbf{E}(V_2 \circ V_1^{-1}(\mathcal{Y}(l+1)) \| \mathcal{Y}(l) = x)$$

(4.1.15)
$$\mathbf{E}(\widetilde{\mathcal{Y}}(l+1) \| \widetilde{\mathcal{Y}}(l) = x) = V_1 \circ U_1^{-1}(x)$$

(4.1.16)
$$\mathbf{D}^2(\widetilde{\mathcal{Y}}(l+1) \| \widetilde{\mathcal{Y}}(l) = x) = V_2 \circ U_1^{-1}(x) + \mathbf{E}(U_2 \circ U_1^{-1}(\widetilde{\mathcal{Y}}(l+1)) \| \widetilde{\mathcal{Y}}(l) = x).$$

We introduce the functions F, $G : \mathrm{Ran}(V_1) \to \mathbb{R}$ and $\widetilde{F}$, $\widetilde{G} : \mathrm{Ran}(U_1) \to \mathbb{R}$ defined below

(4.1.17) $$\begin{aligned} F(x) &= \mathbf{E}\Big(\mathcal{Y}(l+1) \Big\| \mathcal{Y}(l) = x\Big) - x \\ &= U_1\left(V_1^{-1}(x) + 1\right) - x, \end{aligned}$$

$$G(x) = \mathbf{E}\Big(\big[\mathcal{Y}(l+1) - \mathbf{E}(\mathcal{Y}(l+1) \| \mathcal{Y}(l) = x)\big]^2 \Big\| \mathcal{Y}(l) = x\Big) = U_2\left(V_1^{-1}(x)+1\right) + \mathbf{E}\Big(V_2 \circ V_1^{-1}(\mathcal{Y}(1)) \Big\| \mathcal{Y}(0) = x\Big), \tag{4.1.18}$$

$$\widetilde{F}(x) = \mathbf{E}\Big(\widetilde{\mathcal{Y}}(l+1) \Big\| \widetilde{\mathcal{Y}}(l) = x\Big) - x = V_1 \circ U_1^{-1}(x) - x, \tag{4.1.19}$$

$$\widetilde{G}(x) = \mathbf{E}\Big(\big[\widetilde{\mathcal{Y}}(l+1) - \mathbf{E}(\widetilde{\mathcal{Y}}(l+1) \| \widetilde{\mathcal{Y}}(l) = x)\big]^2 \Big\| \widetilde{\mathcal{Y}}(l) = x\Big) = V_2 \circ U_1^{-1}(x) + \mathbf{E}\Big(U_2 \circ U_1^{-1}(\widetilde{\mathcal{Y}}(1)) \Big\| \widetilde{\mathcal{Y}}(0) = x\Big). \tag{4.1.20}$$

Since $\mathcal{Y}(\cdot)$ and $\widetilde{\mathcal{Y}}(\cdot)$ are *Markov chains*, from (4.1.13)–(4.1.20) it follows that the processes

$$\mathcal{M}(l) = \mathcal{Y}(l) - \mathcal{Y}(0) - \sum_{j=0}^{l-1} F(\mathcal{Y}(j)), \qquad \widetilde{\mathcal{M}}(l) = \widetilde{\mathcal{Y}}(l) - \widetilde{\mathcal{Y}}(0) - \sum_{j=0}^{l-1} \widetilde{F}(\widetilde{\mathcal{Y}}(j)) \tag{4.1.21}$$

are *martingales* with quadratic variation processes

$$\langle \mathcal{M}, \mathcal{M}\rangle(l) = \sum_{j=0}^{l-1} G(\mathcal{Y}(j)), \qquad \langle \widetilde{\mathcal{M}}, \widetilde{\mathcal{M}}\rangle(l) = \sum_{j=0}^{l-1} \widetilde{G}(\widetilde{\mathcal{Y}}(j)). \tag{4.1.22}$$

4.2. Asymptotics of the relevant functions

In the present subsection we give the asymptotics of the relevant functions, F, G, $\widetilde{F}$, $\widetilde{G}$ to be used in the proof of Theorem 1. All formulas are valid for *large values* of the variable and are obtained from (1.4) and (1.5) in a straightforward way.

From (1.5) we get

$$U_1(n) = n^{1-\alpha} + u + O(n^{-\alpha}), \tag{4.2.1}$$

$$V_1(n) = n^{1-\alpha} + v + O(n^{-\alpha}), \tag{4.2.2}$$

$$V_2(n), U_2(n) = \begin{cases} O\left(n^{1-2\alpha}\right) & \text{if} \quad 0 < \alpha < \frac{1}{2}, \\ O\left(\log n\right) & \text{if} \quad \alpha = \frac{1}{2}, \\ O\left(1\right) & \text{if} \quad \frac{1}{2} < \alpha < 1, \end{cases} \tag{4.2.3}$$

u and v in (4.2.1) and (4.2.2) are two real constants. We define

$$D = \lim_{n\to\infty}\Big(U_1(n) - V_1(n)\Big) = u - v. \tag{4.2.4}$$

Clearly,

$$\begin{aligned} D &= \sum_{j=0}^{\infty}\left(\frac{1}{w(2j)} - \frac{1}{w(2j+1)}\right) \\ &= \frac{1}{w(0)} - \sum_{j=1}^{\infty}\left(\frac{1}{w(2j-1)} - \frac{1}{w(2j)}\right) \end{aligned} \tag{4.2.5}$$

and hence, due to (1.4),

$$0 < D < w(0)^{-1} < \infty. \tag{4.2.6}$$

The asymptotics of the functions F, $\widetilde{F}$, G, and $\widetilde{G}$ is given in the next Lemma:

LEMMA 2. *The following asymptotics hold for* $x \gg 1$:

$$F(x) = D + O(x^{-\alpha/(1-\alpha)} \vee x^{-1}) \tag{4.2.7}$$

$$\widetilde{F}(x) = -D + O(x^{-\alpha/(1-\alpha)} \vee x^{-1}) \tag{4.2.8}$$

$$G(x), \widetilde{G}(x) = \begin{cases} O\left(x^{(1-2\alpha)/(1-\alpha)}\right) & \textit{if} \quad 0 < \alpha < \frac{1}{2} \\ O(\log x) & \textit{if} \quad \alpha = \frac{1}{2} \\ O(1) & \textit{if} \quad \frac{1}{2} < \alpha < 1. \end{cases} \tag{4.2.9}$$

PROOF. Note first that (4.2.1) and (4.2.2) imply

$$U_1^{-1}(x) = x^{1/(1-\alpha)} - \frac{u}{1-\alpha}x^{\alpha/(1-\alpha)} + O(1) \tag{4.2.10}$$

and

$$V_1^{-1}(x) = x^{1/(1-\alpha)} - \frac{v}{1-\alpha}x^{\alpha/(1-\alpha)} + O(1), \tag{4.2.11}$$

respectively. Inserting (4.2.1) and (4.2.11) into (4.1.17) [respectively, (4.2.2) and (4.2.10) into (4.1.19)] we readily get (4.2.7) [respectively, (4.2.8)].

In order to prove (4.2.9) we note first that, due to (4.2.10), (4.2.11) and (4.2.3) we have:

$$\begin{aligned} &U_2(V_1^{-1}(x)+1),\ V_2\circ V_1^{-1}(x),\ V_2\circ U_1^{-1}(x),\ U_2\circ U_1^{-1}(x) = \\ &= \begin{cases} O\left(x^{(1-2\alpha)/(1-\alpha)}\right) & \text{if} \quad 0 < \alpha < \frac{1}{2} \\ O(\log x) & \text{if} \quad \alpha = \frac{1}{2} \\ O(1) & \text{if} \quad \frac{1}{2} < \alpha < 1. \end{cases} \end{aligned} \tag{4.2.12}$$

Inserting these into (4.1.18) and (4.1.20), and applying Jensen's inequality (note that the functions $x \mapsto x^{(1-2\alpha)/(1-\alpha)}$ and $x \mapsto \log x$ are *concave*), we get eventually (4.2.9). □

Note also that the functions $x \mapsto F(x)$ and $x \mapsto \widetilde{F}(x)$ are monotone decreasing, with

$$(4.2.13)\qquad \begin{aligned} D = \lim_{x\to\infty} F(x) \leq F(x) \leq F(0) = \frac{1}{w(0)}, \\ -D = \lim_{x\to\infty} \widetilde{F}(x) \leq \widetilde{F}(x) \leq \widetilde{F}(0) = 0. \end{aligned}$$

4.3. Scaling

The proper scaling of the processes $\mathcal{Y}(\cdot)$ and $\widetilde{\mathcal{Y}}(\cdot)$ is determined by the dominant terms in the asymptotics of the functions F, G, respectively $\widetilde{F}, \widetilde{G}$. The scaling of the processes $\mathcal{Z}(\cdot)$ and $\widetilde{\mathcal{Z}}(\cdot)$ is determined by the functional relations (4.1.12).

(4.2.9)–(4.2.11) suggest the following scaling:

$$(4.3.1)\qquad Y_A(t) = A^{-1}\mathcal{Y}([At]), \qquad \widetilde{Y}_A(t) = A^{-1}\widetilde{\mathcal{Y}}([At]).$$

The rescaled martingales $M_A(\cdot)$, $\widetilde{M}_A(\cdot)$ and their quadratic variation processes will be

$$(4.3.2)\qquad M_A(t) = A^{-1}\mathcal{M}([At]) = Y_A(t) - Y_A(0) - \int_0^{A^{-1}[At]} F\left(AY_A(s)\right) ds,$$

$$(4.3.3)\qquad \widetilde{M}_A(t) = A^{-1}\widetilde{\mathcal{M}}([At]) = \widetilde{Y}_A(t) - \widetilde{Y}_A(0) - \int_0^{A^{-1}[At]} \widetilde{F}\left(A\widetilde{Y}_A(s)\right) ds,$$

$$(4.3.4)\qquad \langle M_A, M_A\rangle(t) = A^{-2}\langle\mathcal{M},\mathcal{M}\rangle([At]) = \int_0^{A^{-1}[At]} A^{-1}G\left(AY_A(s)\right) ds,$$

$$(4.3.5)\qquad \langle \widetilde{M}_A, \widetilde{M}_A\rangle(t) = A^{-2}\langle\widetilde{\mathcal{M}},\widetilde{\mathcal{M}}\rangle([At]) = \int_0^{A^{-1}[At]} A^{-1}\widetilde{G}\left(A\widetilde{Y}_A(s)\right) ds.$$

The functional relations (4.1.12), the asymptotics (4.2.1), respectively (4.2.2), and the scaling (4.3.1) determine the proper scaling of the processes $\mathcal{Z}(\cdot)$ and $\widetilde{\mathcal{Z}}(\cdot)$:

$$(4.3.6)\qquad Z_A(t) = A^{-1/(1-\alpha)}\mathcal{Z}([At]), \qquad \widetilde{Z}_A(t) = A^{-1/(1-\alpha)}\widetilde{\mathcal{Z}}([At]).$$

4.4. Convergence of the processes

We assume that the initial conditions converge in probability to the deterministic constants y_0, respectively $\widetilde{y}_0$: denoting the events

$$\mathcal{A}_{\delta,A} = \Big\{ |Y_A(0) - y_0| < \delta \Big\}, \quad \widetilde{\mathcal{A}}_{\delta,A} = \Big\{ \Big|\widetilde{Y}_A(0) - \widetilde{y}_0\Big| < \delta \Big\} \tag{4.4.1}$$

we have for any fixed $\delta > 0$

$$\mathbf{P}\Big(\mathcal{A}_{\delta,A}\Big) \to 1, \quad \mathbf{P}\Big(\widetilde{\mathcal{A}}_{\delta,A}\Big) \to 1. \tag{4.4.2}$$

First we show that the martingales $M_A(\cdot)$ and $\widetilde{M}_A(\cdot)$ converge to zero in probability, uniformly on compact intervals $s \in [0, t]$. Indeed:

$$\begin{aligned} \mathbf{E}\Big(\langle M_A, M_A\rangle(t)\Big) &= \int_0^{A^{-1}[At]} A^{-1}\mathbf{E}\Big(G\left(AY_A(s)\right)\Big)ds \\ &\leqq \int_0^{A^{-1}[At]} A^{-1}\mathbf{E}\Big(C_1\left(AY_A(s)\right)^{1-\alpha} + C_2\Big)ds \\ &\leqq A^{-\alpha}C_1 \int_0^{A^{-1}[At]} \Big(\mathbf{E}Y_A(s)\Big)^{1-\alpha} ds + A^{-1}C_2 t \\ &\leqq A^{-\alpha}C_1 \int_0^{A^{-1}[At]} \left(Y_A(0) + C_3 s\right)^{1-\alpha} ds + A^{-1}C_2 t \to 0. \end{aligned} \tag{4.4.3}$$

In the first inequality the asymptotics (4.2.9) of the function G is used, the second one follows from Jensen's inequality, finally in the last inequality we have used (4.3.2) and the fact that the function F is bounded. Define the events

$$\mathcal{B}_{t,\delta,A} = \Big\{ \sup_{0 \leqq s \leqq t} |M_A(s)| < \delta \Big\}, \quad \widetilde{\mathcal{B}}_{\delta,A} = \Big\{ \sup_{0 \leqq s \leqq D^{-1}\widetilde{y}_0} \Big|\widetilde{M}_A(s)\Big| < \delta \Big\}. \tag{4.4.4}$$

From (4.4.3) and a similar argument applied to the martingale $\widetilde{M}_A(\cdot)$ we conclude that for any $t \in [0, \infty)$ and any $\delta > 0$ fixed

$$\mathbf{P}\Big(\mathcal{B}_{t,\delta,A}\Big) \to 1, \quad \mathbf{P}\Big(\widetilde{\mathcal{B}}_{\delta,A}\Big) \to 1 \tag{4.4.5}$$

as $A \to \infty$. Due to (4.3.2) (respectively, (4.3.3)) and (4.2.13), on the sets $\mathcal{A}_{\delta,A} \cap \mathcal{B}_{t,\delta,A}$ (respectively, on the sets $\widetilde{\mathcal{A}}_{\delta,A} \cap \widetilde{\mathcal{B}}_{\delta,A}$) we have for $s \in [0,t]$ (respectively, for $s \in [0, D^{-1}\widetilde{y}_0]$)

$$Y_A(s) \geqq \{y_0 + Ds - 2\delta\}_+, \tag{4.4.6}$$

respectively,

$$\widetilde{Y}_A(s) \geqq \{\widetilde{y}_0 - Ds - 2\delta\}_+. \tag{4.4.7}$$

Consequently, given any $t \in [0,\infty)$ fixed, on the set $\mathcal{A}_{\delta,A} \cap \mathcal{B}_{t,\delta,A}$

$$Y_A(s) \geqq \begin{cases} 0 & \text{for} \quad 0 \leqq s \leqq D^{-1}\{3\delta - y_0\}_+ \wedge t \\ \delta & \text{for} \quad D^{-1}\{3\delta - y_0\}_+ \wedge t < s \leqq t. \end{cases} \tag{4.4.8}$$

On the other hand, on the set $\widetilde{\mathcal{A}}_{\delta,A} \cap \widetilde{\mathcal{B}}_{\delta,A}$

$$\widetilde{Y}_A(s) \geqq \begin{cases} \delta & \text{for} \quad 0 \leqq s < D^{-1}\{\widetilde{y}_0 - 3\delta\}_+ \\ 0 & \text{for} \quad D^{-1}\{\widetilde{y}_0 - 3\delta\}_+ \leqq s \leqq D^{-1}\widetilde{y}_0. \end{cases} \tag{4.4.9}$$

Now, choose A big enough to have

$$F(A\delta) - D < \delta, \quad \widetilde{F}(A\delta) + D < \delta. \tag{4.4.10}$$

From (4.3.2) and (4.4.8) it follows that for any $t \in [0,\infty)$, on $\mathcal{A}_{\delta,A} \cap \mathcal{B}_{t,\delta,A}$

$$\begin{aligned} \sup_{0 \leqq s \leqq t} &|Y_A(s) - (y_0 + Ds)| \\ &\leqq |Y_A(0) - y_0| + \int_0^{A^{-1}[At]} (F(AY_A(s)) - D)\, ds + \sup_{0 \leqq s \leqq t} |M_A(s)| \\ &\leqq \delta + 3(w(0)^{-1} - D)D^{-1}\delta + (t + A^{-1})\delta + \delta \\ &\leqq \left(t + 3(w(0)D)^{-1}\right)\delta, \end{aligned} \tag{4.4.11}$$

and hence for any $t \in [0,\infty)$

$$\sup_{0 \leqq s \leqq t} |Y_A(s) - (y_0 + Ds)| \xrightarrow{\mathbf{P}} 0. \tag{4.4.12}$$

On the other hand, from (4.3.3) and (4.4.9) it follows that, on the set $\widetilde{\mathcal{A}}_{\delta,A} \cap \widetilde{\mathcal{B}}_{\delta,A}$

$$\begin{aligned}
&\sup_{0\leqq s\leqq D^{-1}\widetilde{y}_0} |\widetilde{Y}_A(s) - (\widetilde{y}_0 - Ds)| \\
(4.4.13)\qquad &\leq \left|\widetilde{Y}_A(0) - \widetilde{y}_0\right| + \int_0^{A^{-1}[AD^{-1}\widetilde{y}_0]} \left(\widetilde{F}(A\widetilde{Y}_A(s)) + D\right) ds \\
&\quad + \sup_{0\leqq s\leqq D^{-1}\widetilde{y}_0} \left|\widetilde{M}_A(s)\right| \\
&\leqq \delta + 3\delta + (t + A^{-1})\delta + \delta \leqq \left(D^{-1}\widetilde{y}_0 + 6\right)\delta
\end{aligned}$$

and hence:

$$(4.4.14)\qquad \sup_{0\leqq s\leqq D^{-1}\widetilde{y}_0} \left|\widetilde{Y}_A(s) - (\widetilde{y}_0 - Ds)\right| \overset{\mathbf{P}}{\to} 0.$$

4.5. Convergence of the extinction time

The forthcoming argument is a repeat of the proof presented in subsection 5.7/A1 of [7].

For $x \in \mathbb{R}_+$ we denote

$$(4.5.1)\qquad \widetilde{\sigma}_x = \inf\{l \geqq 0 : \widetilde{\mathcal{Y}}(l) \leqq x\}$$

$$(4.5.2)\qquad \widetilde{\sigma}_{x,A} = \inf\{t \geqq 0 : \widetilde{Y}_A(t) \leqq x\}.$$

We prove now that for any $\eta > 0$:

$$(4.5.3)\qquad \lim_{y\to 0}\lim_{A\to\infty} \mathbf{P}\left(\widetilde{\sigma}_{0,A} > \eta \middle\| \widetilde{Y}_A(0) = y\right) = 0,$$

which is equivalent to

$$(4.5.4)\qquad \lim_{y\to 0}\lim_{A\to\infty} \mathbf{P}\left(\widetilde{\sigma}_0 > A\eta \middle\| \widetilde{\mathcal{Y}}(0) = Ay\right) = 0.$$

From (4.2.13) it follows that there exists an $x_0 < \infty$ such that for $x \geqq x_0$

$$(4.5.5)\qquad \widetilde{F}(x) \leqq -\frac{D}{2} < 0$$

and thus

$$(4.5.6)\qquad \mathcal{N}(l) = \widetilde{\mathcal{Y}}(l) + \frac{D}{2} l$$

is *supermartingale* as long as $\mathcal{Y}(l) \geqq x_0$. Applying the optional sampling theorem to the supermartingale $\mathcal{N}(l)$ we get for $y > x_0$

$$\mathbf{E}\left(\widetilde{\sigma}_{x_0} \middle\| \widetilde{\mathcal{Y}}(0) = y\right) \leqq \frac{2}{D} y. \tag{4.5.7}$$

Now, we prove (4.5.4):

$$\begin{aligned} &\lim_{y\to 0} \lim_{A\to\infty} \mathbf{P}\left(\widetilde{\sigma}_0 > A\eta \middle\| \widetilde{\mathcal{Y}}(0) = Ay\right) \leqq \\ &\leqq \lim_{y\to 0} \lim_{A\to\infty} \mathbf{P}\left(\widetilde{\sigma}_{x_0} > A\eta/2 \middle\| \widetilde{\mathcal{Y}}(0) = Ay\right) + \\ &\lim_{A\to\infty} \sup_{0\leqq x\leqq x_0} \mathbf{P}\left(\widetilde{\sigma}_0 > A\eta/2 \middle\| \widetilde{\mathcal{Y}}(0) = x\right). \end{aligned} \tag{4.5.8}$$

Applying Markov's inequality and (4.5.7) we get

$$\lim_{y\to 0} \lim_{A\to\infty} \mathbf{P}\left(\widetilde{\sigma}_{x_0} > A\eta/2 \middle\| \widetilde{\mathcal{Y}}(0) = Ay\right) \leqq \lim_{y\to 0} \frac{4y}{D\eta} = 0. \tag{4.5.9}$$

On the other hand, since x_0 is constant independent of A, the second limit on the right-hand side of (4.5.8) clearly vanishes. Hence (4.5.4), or equivalently (4.5.3), follows.

From (4.5.3) and (4.4.14) it follows that

$$\widetilde{\sigma}_{A,0} \xrightarrow{\mathbf{P}} D^{-1}\widetilde{y}_0. \tag{4.5.10}$$

4.6. End of the proof

Collecting the results of subsections 4.1–4.5 we conclude that, provided that $Z_A(0) \xrightarrow{\mathbf{P}} z_0$, for any fixed $t \in [0, \infty)$:

$$\sup_{0\leqq s\leqq t} \left| Z_A(s) - \{z_0^{1-\alpha} + Ds\}^{1/(1-\alpha)} \right| \xrightarrow{\mathbf{P}} 0 \tag{4.6.1}$$

and, provided that $\widetilde{Z}_A(0) \xrightarrow{\mathbf{P}} \widetilde{z}_0$,

$$\widetilde{\sigma}_{0,A} \xrightarrow{\mathbf{P}} \widetilde{D}^{-1}\widetilde{z}_0^{1-\alpha} \tag{4.6.2}$$

and

$$\sup_{0\leqq s\leqq \widetilde{D}^{-1}\widetilde{z}_0^{1-\alpha}} \left| \widetilde{Z}_A(s) - \{\widetilde{z}_0^{1-\alpha} - Ds\}^{1/(1-\alpha)} \right| \xrightarrow{\mathbf{P}} 0. \tag{4.6.3}$$

Given the representation of the local time process described in subsection 3.2, Theorem 1 follows directly from (4.6.1)-(4.6.3), after noting that due to (3.1.11) it is easily seen that the single exceptional step (3.2.5) does not spoil the continuity of the limit process at $y=x$. □

Corollary 1 follows directly from Theorem 1. Note that the *joint convergence* of the processes $S^*_{[Ax],[A^{1/(1-\alpha)}h]}([A\cdot])/A^{1/(1-\alpha)}$ and extinction times $\omega^{*\pm}_{[Ax],[A^{1/(1-\alpha)}h]}/A$ is needed in this proof. □

REFERENCES

[1] DAVIS, B., Reinforced random walk, *Probab. Theory Related Fields* **84** (1990), 203–229. *MR* **91a**: 60179

[2] DAVIS, B., Weak limits of perturbed random walks and the equation $Y_t = B_t + \alpha \sup_{s\leq t} Y_s + \beta \inf_{s\leq t} Y_s$, *Ann. Probab.* **24** (1996), 2007–2023.

[3] KNIGHT, F. B., Random walks and a sojourn density process of Brownian motion, *Trans. Amer. Math. Soc.* **109** (1963), 56–86. *MR* **27** #4286

[4] PEMANTLE, R., Phase transition in reinforced random walk and RWRE on trees, *Ann. Probab.* **16** (1988), 1229–1241. *MR* **89g**:60220

[5] TÓTH, B., 'True' self-avoiding walk with generalized bond repulsion on Z, *J. Statist. Phys.* **77** (1994), 17–33. *MR* **95k**:82017

[6] TÓTH, B., The 'true' self-avoiding walk with bond repulsion on Z: limit theorems, *Ann. Probab.* **23** (1995), 1723–1756.

[7] TÓTH, B., Generalized Ray–Knight theory and limit theorems for self-interacting random walks on Z, *Ann. Probab.* **24** (1996), 1324–1367.

(Received February 4, 1997)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

balint@math-inst.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 339–343

GENERALIZED MEANDERS AS LIMITS OF WEIGHTED BESSEL PROCESSES, AND AN ELEMENTARY PROOF OF SPITZER'S ASYMPTOTIC RESULT ON BROWNIAN WINDINGS

M. YOR

Dedicated to Professor E. Csáki on his sixtieth birthday

Let $(B_t, t \geqq 0)$ be a 1-dimensional Brownian motion starting from 0, and $g_1 = \sup\{t < 1, B_t = 0\}$. Recall the definition of the Brownian meander (due to Chung; see e.g. [2]):

$$m_u = \frac{1}{\sqrt{1-g_1}} |B_{g_1+u(1-g_1)}| \qquad u \leqq 1,$$

and, more generally, of the Brownian meander of length t:

$$m_u^{(t)} = \sqrt{t} m_{u/t}, \qquad u \leqq t.$$

It is well known (Imhof [3], Biane–Yor [1], [2]) that M_t, the law of $m^{(t)}$, as defined on $C([0,t]; \mathbb{R}_+)$ satisfies:

$$(1) \qquad M_t = \left(\sqrt{\frac{\pi t}{2}} \frac{1}{X_t}\right) P_0^3\Big|_{\mathcal{F}_t},$$

where P_0^3 denotes the law of a 3-dimensional Bessel process starting from 0. More generally, we may define a two-parameter family $M_t^{d,d'}$ of distributions on $C([0,t]; \mathbb{R}_+)$ by

$$(1)_{d,d'} \qquad M_t^{d,d'} = \left(c_{d,d'} \frac{t^{d/2}}{X_t^d}\right) P_0^{d+d'}\Big|_{\mathcal{F}_t},$$

where P_0^δ denotes the law of a δ-dimensional Bessel process starting from 0; and $c_{d,d'}$ is a normalizing constant.

For some important properties of these distributions, see Yor [4], and Pitman–Yor [6].

Amongst these laws, the laws of the meanders associated to Bessel processes with dimension $\delta < 2$ are found: precisely, if $\delta = 2(1+\mu)$, with $-1 < \mu < 0$, and $(R_t^{(\mu)}, t \geqq 0)$ denotes a Bessel process starting from 0, with dimension δ, then the law of

$$m_{(\mu)}(u) \equiv \frac{1}{\sqrt{1-g_\mu}} R_{g_\mu + u(1-g_\mu)}, \qquad u \leqq 1$$

1991 *Mathematics Subject Classification.* Primary 60J65.
Key words and phrases. Meanders, Brownian windings.

0081–6906/97/$ 5.00

or rather $M_t^{(\mu)}$, the law of $\sqrt{t}m_{(\mu)}(u/t)$, $u \leqq t$, satisfies

$$M_t^{(\mu)} = \left(\frac{t^{-\mu}}{X_t^{(-2\mu)}} \right) P_0^{(-\mu)} \Big|_{\mathcal{F}_t}.$$

In this note, we show that some of those generalized meander laws $M_t^{d,d'}$ may be obtained as weak limits as $r \to 0$ of

$$P_r^{(\mu),\lambda} \Big|_{\mathcal{F}_t} \stackrel{\text{def}}{=} \frac{\exp\left(-\frac{\lambda^2}{2} \int\limits_0^t \frac{ds}{X_s^2}\right)}{e_{\mu,\lambda}} P_r^{(\mu)} \Big|_{\mathcal{F}_t},$$

where

$$e_{\mu,\lambda} \stackrel{\text{def}}{=} E_r^{(\mu)} \left[\exp\left(-\frac{\lambda^2}{2} \int\limits_0^t \frac{ds}{X_s^2} \right) \right],$$

and we found it more convenient to refer to the law P_r^δ as $P_r^{(\mu)}$, where μ is the *index* associated with δ by: $\delta = 2(1+\mu)$.

We shall use in an essential way the following absolute continuity relationship:

PROPOSITION 1. *Let $\mu > -1$, and let $\Phi_t \geqq 0$ be $\mathcal{F}_t$ measurable. Then, for $r > 0$, and $\lambda \neq 0$, one has*

$$E_r^{(\mu)} \left[\Phi_t \exp\left(-\frac{\lambda^2}{2} \int\limits_0^t \frac{ds}{X_s^2} \right) \right] = E_r^{(\nu)} \left[\Phi_t \left(\frac{r}{X_t} \right)^{\nu-\mu} \right], \tag{2}$$

where $\nu = (\mu^2 + \lambda^2)^{1/2}$.

Remark that, if $\mu < 0$, then: $P_r^{(\mu)}(t < T_0) < 1$, where $T_0 = \inf\{t : X_t = 0\}$, and we obtain from (2), by letting $\lambda \to 0$,

$$E_r^{(\mu)}[\Phi_t 1_{(t<T_0)}] = E_r^{(-\mu)} \left[\Phi_t \left(\frac{r}{X_t} \right)^{-2\mu} \right]. \tag{2_0}$$

These different results are found, e.g., in [5], where they play an important role in the computation of the laws of exponential functionals of Brownian motion.

1. Weak limits for fixed $\lambda \neq 0$, as $r \to 0$

We have the following

THEOREM 1. *Fix $\lambda \neq 0$. Then, if Φ_t is a bounded, $(\mathcal{F}_t)$ measurable, continuous functional, one has*

$$E_r^{(\mu)}\left[\Phi_t \exp\left(-\frac{\lambda^2}{2}\int_0^t \frac{ds}{X_s^2}\right)\right] \underset{(r\to 0)}{\sim} r^{\nu-\mu} E_0^{(\nu)}\left[\Phi_t \frac{1}{X_t^{\nu-\mu}}\right]. \tag{3}$$

Consequently, $P_r^{(\mu),\lambda}\Big|_{\mathcal{F}_t} \xrightarrow[r\to 0]{(w)} M_t^{\delta,\delta'}$, where $\delta = \nu - \mu$ and $\delta' = 2 + \nu + \mu$.

The proof of (3) follows easily from (2); note that in particular

$$E_r^{(\mu)}\left[\exp\left(-\frac{\lambda^2}{2}\int_0^t \frac{ds}{X_s^2}\right)\right] \sim r^{\nu-\mu} E_0^{(\nu)}\left(\frac{1}{X_t^{\nu-\mu}}\right) \tag{4}$$

and we have

$$E_0^{(\nu)}\left(\frac{1}{X_t^{\nu-\mu}}\right) = \frac{1}{t^{\frac{\nu-\mu}{2}}} E_0^{(\nu)}\left(\frac{1}{X_1^{\nu-\mu}}\right) = (2t)^{\frac{\mu-\nu}{2}} \frac{\Gamma(1+\frac{\nu+\mu}{2})}{\Gamma(1+\nu)},$$

where, to obtain the last formula, we have used the fact that X_1^2 is distributed, under $P_0^{(\nu)}$ as: $2Z_{1+\nu}$, with Z_α denoting a gamma (α) variable.

2. Weak limits for $\lambda = \lambda_r \underset{r\to 0}{\longrightarrow} 0$

In this situation, we need to discuss separately the cases when $\mu \geqq 0$, or $\mu < 0$.

THEOREM 2. *Let $\mu \geqq 0$. Then, if $\lambda_r \to 0$, and if Φ_t is a bounded, $(\mathcal{F}_t)$-measurable, continuous functional, one has*

$$E_r^{(\mu)}\left[\Phi_t \exp\left(-\frac{\lambda_r^2}{2}\int_0^t \frac{ds}{X_s^2}\right)\right] \underset{r\to 0}{\sim} (r^{\nu_r-\mu}) E_0^{(\mu)}(\Phi_t), \tag{5}$$

where

$$\nu_r = (\mu^2 + \lambda_r^2)^{1/2}.$$

Consequently,

$$P_r^{(\mu),\lambda_r}\Big|_{\mathcal{F}_t} \xrightarrow[r\to 0]{(w)} P_0^{(\mu)}.$$

THEOREM 3. *Let $\mu < 0$. Then, if $\lambda_r \underset{r\to 0}{\longrightarrow} 0$, and if Φ_t is a bounded continuous functional, one has*

$$(6)\qquad E_r^{(\mu)}\left[\Phi_t \exp\left(-\frac{\lambda_r^2}{2}\int_0^t \frac{ds}{X_s^2}\right)\right] \underset{r\to 0}{\sim} r^{\nu_r-\mu} E_0^{(-\mu)}\left[\Phi_t \frac{1}{X_t^{-2\mu}}\right].$$

Consequently,

$$P^{(\mu),\lambda_r}\Big|_{\mathcal{F}_t} \underset{r\to 0}{\overset{(w)}{\longrightarrow}} M_t^{(\mu)}.$$

Again, the proofs of Theorem 2 and 3 follow easily from (2). Moreover, it is interesting to look at the equivalences in (5) and (6) when $\Phi_t \equiv 1$.

For instance, if we take in (5) $\mu = 0$, we obtain:

$$E_r^{(0)}\left[\exp\left(-\frac{\lambda_r^2}{2}\int_0^t \frac{ds}{X_s^2}\right)\right] \sim \exp\left(-\lambda_r\left(\log\frac{1}{r}\right)\right),$$

so that, if $\lambda_r = \frac{\lambda}{(\log 1/r)}$ with $\lambda \geqq 0$, one obtains

$$E_r^{(0)}\left[\exp\left(-\frac{\lambda^2}{2(\log 1/r)^2}\int_0^t \frac{ds}{X_s^2}\right)\right] \underset{r\to 0}{\longrightarrow} \exp(-\lambda).$$

Using the scaling property of Brownian motion, it is easily seen that this result is equivalent to Spitzer's result about the asymptotics of the winding number θ_t of planar Brownian motion, as $t \to \infty$, precisely:

$$(7)\qquad \frac{2\theta_t}{\log t} \underset{t\to\infty}{\overset{(\text{law})}{\longrightarrow}} C_1,$$

where C_1 denotes a standard Cauchy variable.

REFERENCES

[1] BIANE, PH. and YOR, M., Valeurs principales associées aux temps locaux browniens, *Bull. Sci. Math.* (2) **111** (1987), 23–101. *MR* **88g**:60188

[2] BIANE, PH. and YOR, M., Quelques précisions sur le méandre brownien, *Bull. Sci. Math.* (2) **112** (1988), 101–109. *MR* **89i**:60156

[3] IMHOF, J. P., Density factorizations for Brownian motion, meander and the three-dimensional Bessel process, and applications, *J. Appl. Probab.* **21** (1984), 500–510. *MR* **85j**:60152

[4] YOR, M., *Some aspects of Brownian motion. Part 1. Some special functionals*, Lectures in Mathematics, ETH Zürich, Birkhäuser-Verlag, Basel, 1992. *MR* **93i**:60155

[5] YOR, M., On some exponential functionals of Brownian motion, *Adv. in Appl. Probab.* **24** (1992), 509–531. *MR* **94b**:60095

[6] PITMAN, J. W. and YOR, M., Quelques identités en loi pour les processus de Bessel, *Astérisque* **236**, Hommage à P. A. Meyer et J. Neveu, Paris, 1996, 249–276.

(Received July 23, 1996)

LABORATOIRE DE PROBABILITÉS
UNIVERSITÉ PARIS VI
TOUR 56–3ème ÉTAGE
4, PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

secret@proba.jussieu.fr

PRINTED IN HUNGARY
Akadémiai Nyomda, Martonvásár

RECENTLY ACCEPTED PAPERS

Roszak, B., On the pointwise approximation by Borel and Euler means

Barbe, P. and Broniatowski, M., Deviation principle for set indexed processes with independent increments

Pielichowski, W., A nonlinear eigenvalue problem related to Gabriella Bognár's conjecture

Bayasgalan, Ts., Fundamental reducibility of normal operators on Krein space

Moszyńska, M., Remarks on the minimality rings of convex bodies

Stanimirović, P. S., Computing minimum and basic solutions of linear systems using the hyper-power method

Mala, J., Finitely generated quasi-proximities

Grätzer, G. and Schmidt, E. T., On finite automorphism groups of simple arguesian lattices

Kiss, Gy., Generalized cotangency sets in projective spaces

Manuscripts should be submitted in duplicate, typed in double spacing on only one side of the paper with wide margins. Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only upon receipt of the signed copy of the agreement.

Authors are encouraged to submit their papers electronically. All common dialects of TeX are welcome. The electronic file of an article should always be accompanied by a hardcopy, the printed version exactly corresponding to the electronic one.

Figures should be submitted on separate sheets, drawn either in India ink at their expected final size, or as printouts and matching files processed electronically, preferably as encapsulated PostScript (EPS) ones.

CONTENTS

HU ISSN 0081-6906

Studia Scientiarum Mathematicarum Hungarica

VOLUME 33
NUMBER 4
1997

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian. It is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
H-1117 Budapest, Prielle Kornélia u. 19-35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H-1364 Budapest

Tel.: +36 1 118 2875 Fax: +36 1 117 7166
e-mail: merza @ math-inst.hu

Subscription information

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H-1519 Budapest

For 1998 volume 34 is scheduled for publication. The subscription price is $ 144.00, air delivery plus $ 20.00.

Coden: SSMHAX November, 1997
Vol: 33 **Pages**: 345-478
Number: 4 **Whole**: 65

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 345–350

PERIODIC SOLUTIONS OF CERTAIN THIRD ORDER NONLINEAR DIFFERENTIAL EQUATIONS

B. MEHRI and D. SHADMAN

Abstract

Periodic solutions of second order nonlinear ordinary differential equations have been considered by many authors (see for example [1], [2], [3] and [4]). Third order equations have also been the subject of many investigations. The differential equation $x'''+\psi(x')x''+\phi(x)x'+f(x)=p(t)$. $p(t)\equiv p(t+\omega)$ has been treated by Reissig [5] and several other authors. In [5] the author treats the cases $\phi(x)=k^2$ and $\psi(x')=c$, respectively. In this paper we make use of the method used in [5] to obtain sufficient conditions for the existence of an ω-periodic solution for the general case of the differential equation cited above.

We consider the third order differential equation

$$x'''+\psi(x')x''+\phi(x)x'+f(t,x)=p(t), \qquad p(t+\omega)\equiv p(t) \tag{1}$$

where the functions $\psi(y)$, $\phi(x)$, $f(t,x)$ and $p(t)$ are assumed to be continuous and in addition f is assumed to be ω-periodic in t.

THEOREM 1. *The differential equation* (1) *admits at least one ω-periodic solution if*

(i) $$\int_0^\omega p(t)dt=0 \quad \left(i.e.,\ P(t)=\int_0^t p(t)dt \ \text{is } \omega\text{-periodic}\right),$$

(ii) $$|\Psi(y)|\leqq G, \quad \left(\Psi(y)=\int_0^y \psi(\eta)d\eta\right),$$

(iii) $$\frac{|\Phi(x)|}{|x|}\to 0, \quad x\to\infty \quad \left(\Phi(x)=\int_0^x \phi(\xi)d\xi\right),$$

(iv) $$\frac{|f(t,x)|}{|x|}\to 0, \quad x\to\infty \quad (\text{uniformly in } t),$$

(v) $$f(t,x)\operatorname{sgn} x\geqq 0, \quad |x|\geqq\rho.$$

1991 *Mathematics Subject Classification.* Primary 34C25.
Key words and phrases. Differential equations, nonlinear, periodic solutions.

For the proof we make use of the Leray–Schauder principle. First we look at the following differential equation containing a parameter μ, $0 \leqq \mu \leqq 1$,

$$x''' + ax' + bx = \mu[p(t) - f(t,x) - \phi(x)x' - \psi(x')x'' + ax' + bx], \tag{2}$$

where a and b are positive but otherwise arbitrary constants.

We notice that for $\mu = 1$, (2) is identical with (1) and for $\mu = 0$ we obtain a linear homogeneous equation

$$x''' + ax' + bx = 0. \tag{3}$$

It is well known ([6], [7], [8]) that (2) admits at least one periodic solution for each $\mu \in [0,1]$, if for $\mu \in (0,1)$ all periodic solutions as well as their derivatives of first and second order are bounded, provided that (3) has no ω-periodic solution except for the trivial solution. This is indeed the case under the condition of positiveness of a and b.

Let $x(t) \equiv x(t+\omega)$ be a solution of (2), then the derivative $y = x'$ satisfies the following equation:

$$\begin{gathered} y'' + ay = q(t), \quad q(t+\omega) \equiv q(t) \\ q(t) = \mu\left[p(t) - f(t,x(t)) - \frac{d}{dt}(\Phi(x(t)) + \Psi(x'(t)) + ax(t))\right] \\ -(1-\mu)bx(t). \end{gathered} \tag{4}$$

Now let $G(t,s)$ be the Green's function of the boundary value problem

$$\begin{aligned} y'' + ay &= q(t), & 0 \leqq t \leqq \omega \\ y(0) &= y(\omega), & y'(0) = y'(\omega), \end{aligned}$$

where

$$G(t,s) = \frac{1}{2\sqrt{a}\sin(\sqrt{a}\frac{\omega}{2})}\begin{cases} \cos\sqrt{a}(\frac{\omega}{2} + t - s), & 0 \leqq t \leqq s \leqq \omega \\ \cos\sqrt{a}(\frac{\omega}{2} - t + s), & 0 \leqq s \leqq t \leqq \omega. \end{cases} \tag{5}$$

We obtain the following representation for $y(t)$ and for its derivatives

$$y(t) = \int_0^\omega G(t,s)\{\mu[p(s) - f(s,x(s))] - (1-\mu)bx(s)\}ds + \int_0^\omega \frac{\partial G(t,s)}{\partial s}[\Phi(x(s)) + \Psi(y(s)) + ax(s)]ds,$$

$$y'(t) = \int_0^\omega \frac{\partial G(t,s)}{\partial s}\{\mu[p(s) - f(s,x(s))] - (1-\mu)bx(s)\}ds + \Phi(x(t)) + \Psi(y(t)) + ax(t) + \int_0^\omega \frac{\partial^2 G}{\partial t \partial s}[\Phi(x(s)) + \Psi(y(s)) + ax(s)]ds.$$

For $0 < \omega < \pi/\sqrt{a}$ we obtain the following bounds for Green's functions and its derivatives

$$|G(t,s)| \leqq \frac{\pi}{2\omega a}, \quad \left|\frac{\partial G}{\partial t}(t,s)\right| \leqq \frac{\pi}{2\omega\sqrt{a}}, \quad \left|\frac{\partial^2 G}{\partial t \partial s}\right| \leqq \frac{\pi}{2\omega}.$$

Denote

$$R = \max_{t\in[0,\omega]} |x(t)|$$
$$\Phi(R) = \max_{|x|\leqq R} |\Phi(x)|$$
$$F(R) = \max_{t\in[0,\omega],|x|\leqq R} |f(t,x)|$$

and let $0 < \mu < 1$, then we derive the following estimates for $y(t)$ and $y'(t)$:

(6) $$|y(x)| \leqq \frac{\pi}{2}\left\{\frac{1}{a}|p| + \frac{1}{a}F(R) + \frac{1}{\sqrt{a}}\Phi(R) + \frac{1}{\sqrt{a}}G + \left(\frac{b}{a} + \sqrt{a}\right)R\right\},$$

(7) $$|y'(t)| \leqq \frac{\pi}{2}\left\{\frac{1}{\sqrt{a}}|p| + \frac{1}{\sqrt{a}}F(R) + \left(\frac{2}{\pi} + 1\right)\Phi(R) + \left(\frac{2}{\pi} + 1\right)G + \left(b + \frac{2a}{\pi} + a\right)R\right\}.$$

Now term by term integration of (2) yields

$$\int_0^\omega [b(1-\mu)x(t) + f(t,x(t))]dt = 0.$$

However, for $\mu \in (0,1)$, $1-\mu > 0$, and we get

$$b(1-\mu)x(t)\operatorname{sgn} x + f(t,x)\operatorname{sgn} x > 0, \quad |x| \geqq \rho.$$

Therefore $|x(t)| \geqq \rho$ for all $t \in [0,\omega]$ does not hold and we have $|x(\tau)| < \rho$ for some $\tau \in (0,\omega)$.

Applying the mean value theorem to an arbitrary interval $[\tau, t] \subset [\tau, \tau+\omega]$, we find

$$|x(t) - x(\tau)| = (t-\tau)|x'(s)|; \quad \tau < s < t$$

(8)
$$|x(t)| < \rho + \frac{\pi\omega}{2}\left[\frac{1}{a}|p| + \frac{1}{a}F(R) + \frac{1}{\sqrt{a}}\Phi(R) + \frac{1}{\sqrt{a}}G + \left(\frac{b}{a} + \sqrt{a}\right)R\right].$$

The above estimates are valid for all t, hence

$$\max|x(t)| = R < \rho + \frac{\pi\omega}{2}\left[\frac{1}{a}|p| + \frac{1}{a}F(R) + \frac{1}{\sqrt{a}}\Phi(R) + \frac{1}{\sqrt{a}}G + \left(\frac{b}{a} + \sqrt{a}\right)R\right]$$
$$t \in [0,\omega].$$

Choosing

(9)
$$\omega < \min\left\{\frac{2}{\pi\left(\dfrac{b}{a} + \sqrt{a}\right)}, \frac{\pi}{\sqrt{a}}\right\}$$

we obtain

(10)
$$1 < \frac{1}{1 - \dfrac{\pi\omega}{2}\left(\dfrac{b}{a} + \sqrt{a}\right)}\left\{\left(\rho + \frac{\pi\omega}{2a}|p| + \frac{\pi\omega}{2\sqrt{a}}G\right)\frac{1}{R} + \frac{\pi\omega}{2a}\frac{F(R)}{R} + \frac{\pi\omega}{2\sqrt{a}}\frac{\Phi(R)}{R}\right\}.$$

From Assumptions (iii) and (iv)

$$\frac{F(R)}{R} \to 0, \qquad \frac{\Phi(R)}{R} \to 0, \quad \text{as} \quad R \to \infty,$$

therefore we conclude from the inequality (10)

$$R = \max_{t\in[0,\omega]}|x(t)| \leqq R_0,$$

$$F(R) = \max_{t\in[0,\omega],|x|\leqq R}|f(t,x)| \leqq F_0 = \max_{t\in[0,\omega],|x|\leqq R_0}|f(t,x)|,$$

$$\Phi(R) = \max_{|x|\leqq R}|\Phi(x)| \leqq \Phi_0 = \max_{x\leqq R_0}|\Phi(x)|.$$

Using the above results and the estimates (6) and (7) we obtain the a priori bounds

$$|x(t)| \leqq R_0,$$
$$|x'(t)| \leqq \frac{\pi}{2}\left\{\frac{1}{a}|p| + \frac{1}{\sqrt{a}}G + \frac{1}{a}F_0 + \frac{1}{\sqrt{a}}\Phi_0 + \left(\frac{b}{a} + \sqrt{a}\right)R_0\right\},$$
$$|x''(t)| \leqq \frac{\pi}{2}\left\{\frac{1}{\sqrt{a}}|p| + \left(\frac{2}{\pi} + 1\right)(G + \Phi_0) + \frac{1}{\sqrt{a}}F_0 + \left(b + \frac{2a}{\pi} + a\right)R_0\right\},$$

which ensure the existence of an ω-periodic solution of the equation (2).

REMARK. In the case

(v') $$f(t,x)\operatorname{sgn}(x) \leqq 0, \quad |x| \geqq \rho$$

we introduce the new independent variable $x = -z$ and we obtain a differential equation of type (1). Thus Theorem 1 remains valid if the assumption (v) is replaced by (v').

EXAMPLE. We consider the differential equation

(11) $$x''' + c_1(\cos x')x'' + c_2(\sin x)x' + c_3(\sin^2 t)x^{1/3} = c_4 \cos t$$

where c_1, c_2 and c_3 are given constants. Here we denote

$$\psi(y) = c_1 \cos y, \quad \phi(x) = c_2 \sin x, \quad f(t,x) = c_3(\sin^2 t)x^{1/3}.$$

Hence the conditions (i)–(v) are satisfied Theorem 1 ensures the existence of at least one 2π-periodic solution of (11).

REFERENCES

[1] BIHARI, I., Oscillations and monotonicity theorems concerning non-linear differential equations of the second order, *Acta. Math. Acad. Sci. Hungar.* **9** (1958), 83–104. *MR* **20** #1824

[2] BIHARI, I., On periodic solutions of certain second order ordinary differential equations with periodic coefficients, *Acta. Math. Acad. Sci. Hungar.* **12** (1961), 11–16. *MR* **23** #A3324

[3] HAMEDANI, G. G. and MEHRI, B., Periodic boundary value problems for certain non-linear second order differential equations, *Studia Sci. Math. Hungar.* **9** (1974), 307–312. *MR* **53** #13704

[4] HAMEDANI, G. G. and MEHRI, B., A nonlinear periodic boundary value problem for a system of equations of the second order, *Studia Sci. Math. Hungar.* **10** (1975), 339–342. *MR* **81j**:34029

[5] REISSIG, R., Periodic solutions of a third order nonlinear differential equation, *Ann. Mat. Pura Appl.* (4) **92** (1972), 193–198. *MR* **47** #5375

[6] REISSIG, R., On the existence of periodic solutions of a certain non-autonomous differential equation, *Ann. Mat. Pura Appl.* (4) **85** (1970), 235–240. *MR* **42** #2105

[7] VILLARI, G., Contributi allo studio dell'esistenza di soluzioni periodiche per i sistemi di equazioni differenziali ordinarie, *Ann. Mat. Pura Appl.* (4) **69** (1965), 171–190. *MR* **32** #7864

[8] VILLARI, G., Soluzioni periodiche di una classe di equazioni differenziali del terz'ordine quasi lineari, *Ann. Mat. Pura Appl.* (4) **73** (1966), 103–110. *MR* **34** #6223

(Received August 18, 1993)

DEPARTMENT OF MATHEMATICS
SHARIF UNIVERSITY OF TECHNOLOGY
P.O.BOX 11365-9415
TEHRAN
IRAN

dshadman@rose.ipm.ac.ir

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 351–362

α_4-PROPERTY VERSUS A-PROPERTY IN TOPOLOGICAL SPACES AND GROUPS

T. NOGURA, D. SHAKHMATOV and Y. TANAKA

Abstract

We give a series of examples demonstrating that A-property (due to E. Michael) and α_4-property (due to A. Arhangel'skii) behave independently from each other in general spaces and groups. They are known to coincide for Fréchet spaces, but are different for sequential spaces (T. Nogura). We show that these properties coincide for: (i) sequential spaces each point of which is a G_δ-set, and (ii) hereditarily normal, sequential topological groups.

1. Introduction

All topological spaces and groups considered in this paper are assumed to be Tychonoff.

In what follows $\overline{A}$ always denotes the closure of a set A in a space X. A space X is called:

sequential if for every non-closed set $A \subseteq X$, there is a sequence of points in A converging to some point outside of A,

Fréchet (=Fréchet–Urysohn) if whenever $A \subseteq X$ and $x \in \overline{A}$, there exists a sequence in A converging to x,

strongly Fréchet [26] *(=countably bi-sequential* in the sense of [13]) if, whenever $\{A_n : n \in \omega\}$ is a decreasing sequence of subsets of X and $x \in \bigcap\{\overline{A_n} : n \in \omega\}$, then there exists a sequence $\{x_n : n \in \omega\}$ converging to x with $x_n \in A_n$ for all $n \in \omega$.

Clearly

first countable $\Rightarrow$ strongly Fréchet $\Rightarrow$ Fréchet $\Rightarrow$ sequential.

In this note we provide a comparison of two convergence properties of topological spaces.

The first one, the α_4-property, was introduced by Arhangel'skii [2,3] in 1972 as an important tool for studying the behaviour of the Fréchet–Urysohn property under the product operation and classification of Fréchet–Urysohn spaces. A countable collection $\mathcal{S} = \{S_n : n \in \omega\}$ of convergent sequences in a space X is called a *sheaf (with a vertex x)* if each sequence S_n converges

---

1991 *Mathematics Subject Classification.* Primary 54A20; Secondary 22A05, 54H11.

Key words and phrases. A-space, α_4-space, convergence, convergent sequence, Fréchet–Urysohn space, sequential space, topological group

0081–6906/97/$ 5.00

to the same point $x \in X$. A space X is an α_4-*space* (equivalently, $X \in \langle 4 \rangle$ in the sense of [2,3], or X is an $\langle \alpha_4 \rangle$-space in the sense of [16]), if for every point $x \in X$ and each sheaf $\mathcal{S} = \{S_n : n \in \omega\}$ with the vertex x, there exists a sequence converging to x which meets infinitely many sequences S_n.

The second property, the A-property, was invented by Michael [14] in 1973. He calls a space X an A-*space* if, whenever $\{A_n : n \in \omega\}$ is a decreasing sequence of subsets of X, and $x \in X$ is a point with $x \in \cap\{\overline{A_n \setminus \{x\}} : n \in \omega\}$, then for every $n \in \omega$ one can find a (possibly empty) set $B_n \subseteqq A_n$ such that $\cup\{\overline{B_n} : n \in \omega\}$ is not closed in X.

Arhangel'skii [3, Theorem 5.23] showed that strongly Fréchet spaces are precisely Fréchet α_4-spaces. Both countably compact and countably bi-k-spaces (in the sense of [13]) are A-spaces. We refer the interested reader to [2,3,6,16,22,25] for properties of α_4-spaces, and to [15] for properties of A-spaces.

Our starting point is the following fact (in which (i) follows from [15, Proposition 8.1] and [3, Theorem 5.23], and (ii) was established in [27, Theorem 1.1]):

FACT 1.1. (i) *A Fréchet space is an α_4-space if and only if it is an A-space.*

(ii) *A sequential α_4-space is an A-space.*

Nogura showed that the implication in (ii) is not reversible:

THEOREM 1.2 [16, Corollary 3.11]. *There exists a sequential compact space (hence an A-space) which is not an α_4-space.*

In this note we study an interrelationship between α_4-spaces and A-spaces in the absence of sequentiality, and also in topological groups. We show that in general α_4-property and A-property are independent from each other even for topological groups (Sections 3-5), but in some special cases the implication in 1.1(ii) is reversible (Section 6). We also formulate some open questions.

2. S_ω versus S_2 for topological groups

In this section we establish a specific, albeit somewhat technical, property of convergence in topological groups which will be used only in the proof of Theorem 6.2. It is therefore possible for the reader to pass directly to Section 3 without losing a continuity of exposition.

The *sequential fan* S_ω is the quotient space obtained from a topological sum of a countable family of convergent sequences by identifying all their limit points to a single point [4].

The *Arens space* S_2 is defined as follows [1]. Let $S_2 = (\omega \times \omega) \cup \omega \cup \{\infty\}$, where each point of $\omega \times \omega$ is isolated, a kth basic open neighbourhood of

$n \in \omega$ consists of all sets in the form $\{n\} \cup \{(m,n) : m \geqq k\}$, and U is a neighbourhood of ∞ if and only if $\infty \in U$ and U is a neighbourhood of all but finitely many $n \in \omega$.

We refer the interested reader to [20] for background results on spaces containing copies of S_ω and S_2, and to [19] for applications of these spaces to metrizability of topological groups.

LEMMA 2.1. *A topological group contains a (closed) homeomorphic copy of S_ω if and only if it contains a (closed) homeomorphic copy of S_2.*

PROOF. Let G be a topological group with the identity element e.

First we establish the "if" part. Let $F = \{e\} \cup \{x_m : m \in \omega\} \cup \{y_{mn} : m,n \in \omega\}$ be a (closed) subset of G naturally homeomorphic to S_2, with $x_m \to e$ and $y_{mn} \to x_m$ for all $m \in \omega$.

Using continuity of algebraic operations in G, one can easily check the following property:

$(*)$ For every function $f \in \omega^\omega$, a point $p \in G$ is a cluster point of the set $\{x_m^{-1} y_{mk} : m \in \omega, k < f(m)\}$ if and only if p is a cluster point of the set $\{y_{mk} : m \in \omega, k < f(m)\}$.

By continuity of group operations, each $L_m = \{e\} \cup \{x_m^{-1} y_{mn} : n \in \omega\}$ is a sequence converging to e. According to $(*)$, every set $\{n \in \omega : L_m \cap L_n$ is infinite$\}$ is finite, so we can find some injection $g \in \omega^\omega$ and a pairwise disjoint family $\{A_{g(m)} : m \in \omega\}$ with each $A_{g(m)}$ an infinite subset of $L_{g(m)}$. Define $A = \cup\{A_{g(m)} : m \in \omega\}$.

Let $\mathcal{F}$ be the set of all functions Φ from ω into the set $G^{<\omega}$ of all finite subsets of G such that $\Phi(m) \subseteqq A_{g(m)}$ for all m. For $\Phi \in \mathcal{F}$ define $E_\Phi = \cup\{\Phi(m) : m \in \omega\}$. We claim that

$$e \notin \overline{\cup\{\overline{E_\Phi} \cap A \backslash E_\Phi : \Phi \in \mathcal{F}\}}. \tag{1}$$

Indeed, from $(*)$ and the choice of $\mathcal{F}$ it immediately follows that

$$e \notin \overline{E_\Phi} \text{ for all } \Phi \in \mathcal{F}. \tag{2}$$

Thus, if (1) fails, then one can find $k \in \omega$ and two sequences $\{\Phi_n : n \in \omega\} \subseteqq \mathcal{F}, \{z_n : n \in \omega\} \subseteqq A_{g(k)}$ such that $z_n \in \overline{E_{\Phi_n}} \backslash E_{\Phi_n}$ for all n, and $z_n \neq z_{n'}$, for $n \neq n'$. Define $\Psi \in \mathcal{F}$ by $\Psi(m) = \cup\{\Psi_l(m) : l \leqq m\}$. Now it can be easily seen that $e \in \overline{E_\Psi}$, in contradiction with (2). Thus (1) holds, which allows us to conclude that

$$T = \{e\} \cup A \backslash \overline{\cup\{\overline{E_\Phi} : \Phi \in \mathcal{F}\}}$$

is a (closed) copy of S_ω.

For the "only if" part, let $H = \{e\} \cup \{y_{mn} : m,n \in \omega\} \subseteqq G$ be a (closed) copy of S_ω, where $y_{mn} \to e$ for $m \in \omega$. Every $R_m = \{y_{0m} y_{mn}^{-1} : n \in \omega\}$ is a sequence converging to e. Arguing in a way similar to the "if" part we may choose $h \in \omega^\omega$, a pairwise disjoint family $\{B_{h(m)} : m \in \omega\}$ with every $B_{h(m)}$

an infinite subset of R_m, and an appropriate open neighbourhood V of e such that

$$S=(\cup\{B_{h(m)}:m\in\omega\}\cup\{y_{0h(m)}:m\in\omega\}\cup\{e\})\cap V$$

would be a (closed) homeomorphic copy of S_2. □

REMARK 2.2. Clearly Lemma 2.1 is specific for topological groups. Indeed, S_ω does not contain a homeomorphic copy of S_2, and vice versa, S_2 does not contain a homeomorphic copy of S_ω.

3. A compact group which is not an α_4-space

We use $\mathfrak{c}$ to denote the cardinality of the continuum.

EXAMPLE 3.1. *$D^{\mathfrak{c}}$ is a compact topological group which is not an α_4-space.* (Observe that $D^{\mathfrak{c}}$ is an A-space, being compact.)

CONSTRUCTION. Let $T=\omega^\omega$, the set of all maps from ω to ω, and $D=\{0,1\}$. Since $|T|=\mathfrak{c}$, $D^{\mathfrak{c}}$ and D^T are isomorphic, so we will work in D^T. For $f\in D^T$ we let $S(f)=\{t\in T:f(t)=1\}$, and we use $\mathbf{0}$ to denote the point of D^T which has all its coordinates equal to 0. One can easily verify the following

FACT. *A sequence $\{f_n:n\in\omega\}\subseteq D^T$ converges to $\mathbf{0}$ if and only if $\{S(f_n):n\in\omega\}$ is a point-finite family in T; that is, $\{n\in\omega:t\in S(f_n)\}$ is finite for each $t\in T$.*

For $m,n\in\omega$, let $F_{mn}=\{t\in T:t(m)=n\}$, and define $f_{mn}\in D^T$ by $S(f_{mn})=F_{mn}$. Since $F_{mi}\cap F_{mj}=\emptyset$ if $i\neq j$, each sequence $L_m=\{f_{mn}:n\in\omega\}\subseteq D^T$ converges to $\mathbf{0}$ by the above fact. Thus $\{L_m:m\in\omega\}$ is a sheaf with the vertex $\mathbf{0}$. We claim that this sheaf violates the α_4-property in D^T. Indeed, let $\varphi,\psi\in\omega^\omega$ be arbitrary maps such that φ is an injection, and define $g_i=f_{\varphi(i)\psi(i)}\in L_{\varphi(i)}$ for each $i\in\omega$. Choose any point $t^*\in T$ such that $t^*(\varphi(i))=\psi(i)$ for $i\in\omega$ (which is possible by injectivity of φ), and observe that $t^*\in\cap\{S(g_i):i\in\omega\}$. Therefore the family $\{S(g_i):i\in\omega\}$ is not point-finite, and so the sequence $\{g_i:i\in\omega\}$ does not converge to $\mathbf{0}$ (Fact). This implies that D^T is not an α_4-space. □

REMARK 3.2. Nyikos [22, Theorem 1.8] constructed a countable space X of weight $\mathfrak{b}$ which is not an α_4-space. (Here $\mathfrak{b}$ denotes the smallest cardinality of a $<^*$-unbounded subset of ω^ω, where for $f,g\in\omega^\omega$ we write $f<^*g$ iff there exists n such that $f(k)<g(k)$ for all $k\geq n$.) Being a zero-dimensional space of weight $\mathfrak{b}$, X can be embedded as a subspace into $D^{\mathfrak{b}}$. A subspace of an α_4-space must be an α_4-space, so we conlude that *$D^{\mathfrak{b}}$ is not an α_4-space.*[1]

[1] Combining this with the first part of [22, Theorem 1.8] one obtains that D^τ is an α_4-space iff $\tau\leq\mathfrak{b}$. See also [19] for other relevant convergence properties of D^τ.

Since $\mathfrak{b} \leqq \mathfrak{c}$, this improves Example 3.1. We decided to give an elementary proof of 3.1 only for the reader's convenience.

4. A countable A-space need not be an α_4-space

EXAMPLE 4.1. A countable A-space (with a single non-isolated point) which is not an α_4 -space.

To construct such an example we need some preliminaries. Let $\beta\omega$ be the Stone-Čech compactification of ω. For $A \subseteqq \omega$, we set $A^* = \mathrm{Cl}_{\beta\omega}A \backslash A$, where $\mathrm{Cl}_{\beta\omega}A$ denotes the closure of A in $\beta\omega$; in particular, $\omega^* = \beta\omega \backslash \omega$. Symbols $\mathrm{Cl}_{\omega^*} Z$ and $\mathrm{Int}_{\omega^*} Z$ denote the closure and the interior of a set $Z \subseteqq \omega^*$ in ω^*. For a family $\mathcal{S}$ of infinite subsets of ω we define $\mathcal{S}^* = \{S^* : S \in \mathcal{S}\}$.

Let $\mathcal{P}$ be a maximal almost disjoint family of infinite subsets of ω; that is:

(i) $P \cap P'$ is finite for different $P, P' \in \mathcal{P}$, and

(ii) if $S \subseteqq \omega$ is infinite, then $S \cap P$ is infinite for some $P \in \mathcal{P}$.

Choose arbitrarily an infinite, countable $\mathcal{Q} \subseteqq \mathcal{P}$, and let $\mathcal{R} = \mathcal{P} \backslash \mathcal{Q}$. Since each A^* is a clopen subset of ω^*, $F = \omega^* \backslash \cup \mathcal{R}^*$ is a closed subset of ω^* . Let $X = \omega \cup \{*\}$ be the quotient space obtained from the subspace $\omega \cup F$ of $\beta\omega$ by identifying the set F to a single point $* \in X$. Then every point of ω is isolated in X, and $\{A \cup \{F\} : F \subseteqq A^*\}$ is a neighbourhood base of $*$ in X. We claim that X is the required space. But before we proceed with verification of that, we need two facts. The first fact, taken from [10] (see also [17]), holds for every closed subset F of ω^*:

FACT 1. *Let $E \subseteqq \omega$. Then:*

(i) $* \in \overline{E}$ *if and only if* $E^* \cap F \neq \emptyset$, *and*

(ii) E *is a sequence converging to* $*$ *if and only if* $E^* \subseteqq F$.

Our second fact uses the specific choice of F and does not hold for an arbitrary closed subset of ω^*.

FACT 2. *Let U be a clopen (= simultaneously closed and open) subset of ω^*. If the set $\mathcal{P}' = \{P \in \mathcal{P} : U \cap P^* \neq \emptyset\}$ is infinite, then:*

(i) $|\mathcal{P}'| \geqq \omega_1$, *and*

(ii) $U \cap F \neq \emptyset$.

PROOF. (i) Suppose that $\mathcal{P}'$ is infinite and countable. Then $Z = U \backslash \cup (\mathcal{P}')^*$ is a non-empty zero-set in ω^*, and so $\mathrm{Int}_{\omega^*} Z \neq \emptyset$ [24, Theorem 3.3]. By the maximality of the family $\mathcal{P}$, there exists $P \in \mathcal{P}$ such that $\emptyset \neq P^* \cap \mathrm{Int}_{\omega^*} Z \subseteqq P^* \cap U$, which implies that $P \in \mathcal{P}'$, and so $P^* \cap \mathrm{Int}_{\omega^*} Z \subseteqq P^* \cap Z = \emptyset$ by the definition of Z, a contradiction.

(ii) Use Fact 2(i) to choose pairwise distinct sets $P_n \in \mathcal{R} \cap \mathcal{P}'$ and note that

$$0 \neq \mathrm{Cl}_{\omega^*}(\cup\{P_n^* \cap U : n \in \omega\}) \backslash \cup \{P_n^* : n \in \omega\} \subseteqq U \cap F,$$

because $\mathrm{Cl}_{\omega^*}(\cup\{P_n^* \cap U : n \in \omega\})$ is a compact set not covered by any finite union $\cup\{P_k^* : k \leqq m\}$. □

Claim 1. X is not α_4.

PROOF. Indeed, let $\mathcal{Q} = \{Q_n : n \in \omega\}$ be any enumeration of $\mathcal{Q}$. Since $\mathcal{P}$ is an almost disjoint family of infinite subsets of ω, $\mathcal{P}^* = \mathcal{Q}^* \cup \mathcal{R}^*$ consists of pairwise disjoint subsets of ω^* [7, p. 98], and so $\cup\mathcal{Q}^* \subseteqq F = \omega^* \setminus \cup\mathcal{R}^*$. Then Fact 1(ii) implies that $\mathcal{Q}$ is a sheaf in X with the vertex $*$. Suppose now that $E \subseteqq \omega$ is a sequence converging to $*$ which meets infinitely many elements of $\mathcal{Q}$. Since $\mathcal{Q}$ is almost disjoint, taking a subsequence of E if necessary, we may assume that each intersection $E \cap Q_n$ is finite. Now one can easily see that E does not converge to $*$ by Fact 1(ii). □

Claim 2. X is an A-space.

PROOF. Let $\{A_n : n \in \omega\}$ be a decreasing sequence in X with $* \in \overline{\cap\{A_n \setminus \{*\} : n \in \omega\}}$. We may assume that $* \notin \cup\{A_n : n \in \omega\}$. Then, for each $n \in \omega$, we have $A_n^* \cap F \neq \emptyset$ by Fact 1(i), and so one can choose $Q_n \in \mathcal{Q}$ with $A_n^* \cap Q_n^* \neq \emptyset$.

Case 1. Some A_n^* is covered by finitely many elements of $\mathcal{P}^*$, say $A_n^* \subseteqq \cup\mathcal{S}^*$ for some finite $\mathcal{S} \subseteqq \mathcal{P}$. Since the family $\mathcal{P}^*$ is disjoint, and $A_m^* \subseteqq A_n^*$ for $m \geqq n$, it follows that $\{Q_m : m \geqq n\} \subseteqq \mathcal{S}$. Since $\mathcal{S}$ is finite, there exist $Q \in \mathcal{S} \cap \mathcal{Q}$ and $k \geqq m$ such that $A_l^* \cap Q^* \neq \emptyset$ for $l \geqq k$. Now choose $x_l \in A_l \cap Q$ for each $l \geqq k$. Then $\{x_l : l \geqq k\}$ converges to $*$ by Fact 1(ii).

Case 2. No A_n^* is covered by finitely many elements of $\mathcal{P}^*$. Since each A_n^* is a clopen subset of ω^*, in this case we have $|\{P \in \mathcal{P} : A_n^* \cap P^* \neq \emptyset\}| \geqq \omega_1$ for every $n \in \omega$ (Fact 2(i)). Therefore one can choose pairwise distinct $P_n^* \in \mathcal{R}$ with $P_n^* \cap A_n^* \neq \emptyset$. Now observe that each set $E_n = P_n \cap A_n$ is non-empty, and $E_n^* \cap F = \emptyset$. Thus every E_n is closed in X by Fact 1(i). To show that the set $\cup\{\overline{E_n} : n \in \omega\} = \cup\{E_n : n \in \omega\}$ is not closed in X, it suffices to check that $* \in \overline{\cup\{E_n : n \in \omega\}}$. In its turn, to get this we only need to prove, by Fact 1(i), that $(\cup\{E_n : n \in \omega\})^* \cap F \neq \emptyset$. Since $\{E_n^* : n \in \omega\}$ is a disjoint collection of nonempty, clopen subsets of $\mathrm{Cl}_{\omega^*} \cup \{E_n^* : n \in \omega\}$, we have $\mathrm{Cl}_{\omega^*} \cup \{E_n^* : n \in \omega\} \setminus \cup \{E_n^* : n \in \omega\} \neq \emptyset$ by compactness of $\mathrm{Cl}_{\omega^*} \cup \{E_n^* : n \in \omega\}$. Pick arbitrarily $p \in \mathrm{Cl}_{\omega^*} \cup \{E_n^* : n \in \omega\} \setminus \cup \{E_n^* : n \in \omega\}$. Since $\mathcal{P}^*$ consists of pairwise disjoint clopen subsets of ω^* and $\{E_n : n \in \omega\} \subseteqq \mathcal{P}$, it follows that $p \notin \cup\mathcal{P}^*$, and so $p \in \omega^* \setminus \cup\mathcal{P}^* \subseteqq F$. Therefore $p \in F \cap \mathrm{Cl}_{\omega^*} \cup \{E_n^* : n \in \omega\} \subseteqq (\cup\{E_n^* : n \in \omega\})^* \cap F \neq \emptyset$. □

Let $C_p(I)$ denote the space of all real-valued continuous functions defined on the unit interval $I = [0,1]$ with the topology of pointwise convergence, i.e. the topology inherited by $C_p(X)$ from $\mathbb{R}^X$. Note that $C_p(I)$ is a topological group (even a topological ring).

EXAMPLE 4.2. *The function space $C_p(I)$ is an A-space (with a countable network) which is not an α_4-space.*

The fact that $C_p(I)$ is an A-space was apparently noticed first by Roman Pol (see [15, Remark added in proof]). Since the argument is very short (and not included in [15]), we present it here for the sake of completeness. Suppose that $A_n \subseteq C_p(I)\setminus\{\mathbf{0}\}$ for every $n\in\omega$, where $\mathbf{0}$ is the function identically equal to zero on I. Assume also that $\mathbf{0}\in\bigcap\{\overline{A_n}:n\in\omega\}$. For $n\in\omega$ and $f\in C_p(I)$ define $U_f^n=\{x\in I: f(x)<1/n\}$. Observe that each $\{U_f^n: f\in A_n\}$ is an open cover of I. By compactness of I, there exists a finite set $B_n\subseteq A_n$ with $I=\bigcup\{U_f^n: f\in B_n\}$. Now it can be easily checked that $\mathbf{0}\in\overline{\bigcup\{B_n:n\in\omega\}}$, which means that $\bigcup\{\overline{B_n}:n\in\omega\}=\bigcup\{B_n:n\in\omega\}$ is not closed in $C_p(I)$.

Gerlits and Nagy [6, Corollary to Theorem 8] showed that $C_p(I)$ is not an α_4-space. Finally, it is well-known that $C_p(I)$ has a countable network. To get such a network, fix countable bases $\mathcal{U}$ and $\mathcal{V}$ for I and $\mathbb{R}$, respectively, and note that $\mathcal{N}=\{F(U,V):U\in\mathcal{U},V\in\mathcal{V}\}$ is as required, where $F(U,V)=\{f\in C_p(I): f(U)\subseteq V\}$.[2] □

5. A countable α_4 -group which is not an A-space

LEMMA 5.1. *There exists a countable space X with the following properties:*

(i) *X has only one non-isolated point,*

(ii) *X has no non-trivial convergent sequences (and therefore, X is obviously an α_4-space), and*

(iii) *X is not an A-space.*

PROOF. Fix an arbitrary point $p\in\beta\omega\setminus\omega$, and define $Y=\omega\cup\{p\}$ and $Z=Y\times N$, where Y is equipped with the subspace topology inherited from $\beta\omega$, and N is the set of natural numbers with the discrete topology. Define $F=\{(p,n):n\in\omega\}$, and let $f:Z\to X=Z/F$ be the quotient map collapsing F to a point $q\in X$. We claim that X is as required. Only the verification of the fact that X is not an A-space is necessary, because other properties are immediate. So let $A_n=f(\bigcup\{Y\times\{i\}:i\geq n\})$ for $n\in\omega$. Then $\{A_n:n\in\omega\}$ is a decreasing sequence in X with $q\in\bigcap\{\overline{A_n\setminus\{q\}}:n\in\omega\}$, but for any choice of $B_n\subseteq A_n, n\in\omega$, the set $\bigcup\{\overline{B_n}:n\in\omega\}$ is closed in X. □

EXAMPLE 5.2. A countable Abelian group G without non-trivial convergent sequences which is not an A-space. (Therefore, G is an α_4-space, but not an A-space.)

We shall give two examples below. The first example uses a result from the theory of free topological groups, while the second one does not require any knowledge of such a theory.

EXAMPLE 1. Let X be the countable space constructed in Lemma 5.1, and let $G=F(X)$ be the (Graev) free topological group over X. (For the

[2] The construction of $\mathcal{N}$ is essentially due to Michael [12].

theory of free topological groups see [8] or [11].) Since X is a paracompact space in which every compact subset is finite, it follows from [5, Theorem 1.5] that the (countable) topological group $F(X)$ contains no non-trivial convergent sequences. Since $F(X)$ contains a closed copy of X, which is not an A-space, $F(X)$ is not an A-space either.[3] □

To construct the second example we need to remind ourselves some general construction. Let $\mathcal{F}$ be a free filter on ω with the finite intersection property, i.e.

(a) $\cap\mathcal{F}=\emptyset$, and

(b) if $F\in\mathcal{F}$ and $H\in\mathcal{F}$, then $F\cap H\in\mathcal{F}$.

Let $X_{\mathcal{F}}=\omega\cup\{p\}$ be a space such that all points of ω are isolated and $\{\{p\}\cup F : F\in\mathcal{F}\}$ is a neighbourhood base of the point p. Then $X_{\mathcal{F}}$ is completely regular. Let $\mathcal{G}$ be the family of all finite subsets of ω with the group operation $(A,B)\to A+B=(A\backslash B)\cup(B\backslash A)$ which makes $\mathcal{G}$ an Abelian group. The zero element of $\mathcal{F}$ is the empty set and $A+A=0$ for every $A\in\mathcal{G}$. For $F\in\mathcal{F}$ the set $F^*=\{A\in\mathcal{G} : A\subseteqq F\}$ is a subgroup of $\mathcal{G}$. We can take the family $\mathcal{B}_{\mathcal{F}}=\{F^* : F\in\mathcal{F}\}$ of subgroups of $\mathcal{G}$ as neighbourhoods of zero of some group topology on $\mathcal{G}$; see [9]. We use the symbol $\mathcal{G}_{\mathcal{F}}$ for denoting the group $\mathcal{G}$ with this group topology. From (a) it follows that $\cap\mathcal{B}_{\mathcal{F}}=\{0\}$, so $\mathcal{G}_{\mathcal{F}}$ is Hausdorff [9].

Observe that the map $i: X_{\mathcal{F}}\to\mathcal{G}_{\mathcal{F}}$ defined by $i(n)=\{n\}$ and $i(p)=0$ is a homeomorphic embedding.

LEMMA 5.3. *$\mathcal{G}_{\mathcal{F}}$ contains a non-trivial convergent sequence if and only if $X_{\mathcal{F}}$ does.*

PROOF. Since $i\colon X_{\mathcal{F}}\to\mathcal{G}_{\mathcal{F}}$ is a homeomorphic embedding, the "if" part holds. To check the "only if" part, suppose that $\mathcal{S}=\{A_n : n\in\omega\}$ is a non-trivial convergent sequence in $\mathcal{G}_{\mathcal{F}}$. Without loss of generality one may assume that $\mathcal{S}$ converges to 0, and that $A_n\neq\emptyset$ for all $n\in\omega$.

Claim. There is a sequence $\{n_k : k\in\omega\}$ such that the family $\{A_{n_k} : k\in\omega\}$ is pairwise disjoint.

Indeed, arguing by induction, suppose that we have already chosen a sequence $n_0,\dots,n_k$ such that $A_{n_0},\dots,A_{n_k}$ are pairwise disjoint. Since $B_k=\cup\{A_{n_i} : i\leqq k\}$ is finite, and $\mathcal{F}$ is a free filter with the finite intersection property, $F\cap B_k=\emptyset$ for some $F\in\mathcal{F}$. Since $F^*\in\mathcal{B}_{\mathcal{F}}$ and $\mathcal{S}$ is a sequence converging to 0, there exists $l\in\omega$ such that $A_j\in F^*$ (equivalently, $A_j\subseteqq F$) for all $j\geqq l$. In particular, $A_l\subseteqq F\subseteqq\omega\backslash B_k$, and so we can set $n_{k+1}=l$. □

Now if one picks $m_i\in A_{n_i}$ for each $i\in\omega$, then $S=\{\{m_i\} : i\in\omega\}\subseteqq X_{\mathcal{F}}$ would be a non-trivial sequence converging to the point p. □

---

[3] The use of the free topological group $F(X)$ in our construction gives us the resulting (algebraically) *free* group G. To get an Abelian group as promised in Example 5.2, we should replace the free topological group $F(X)$ by the free *Abelian* topological group $A(X)$ of X in our construction. The same argument works in this case.

EXAMPLE 2. Now let X be the countable space with a single non-isolated point, say p, which was constructed in Lemma 5.1, and let $\mathcal{F}$ be the family of all open neighbourhood of p with the point p deleted. Then $\mathcal{F}$ satisfies (a) and (b), so we can consider $X_{\mathcal{F}}$ and $\mathcal{G}_{\mathcal{F}}$. Since $i(X_{\mathcal{F}})$ is a closed subset of $\mathcal{G}_{\mathcal{F}}$ homeomorphic to $X = X_{\mathcal{F}}$, and $X_{\mathcal{F}}$ is not an A-space, $\mathcal{G}_{\mathcal{F}}$ is not an A-space either. Since $X_{\mathcal{F}}$ contains no non-trivial convergent sequences, the same is true for $\mathcal{G}_{\mathcal{F}}$ (Lemma 5.3). □

REMARK 5.4. It should be noted that Lemma 5.3 does not have an analogue for X and the free topological group $F(X)$ of X instead of $X_{\mathcal{F}}$ and $G_{\mathcal{F}}$. Indeed, Tkachuk [28] established that the free topological group $F(Y)$ of the Alexandroff double Y of $\beta\omega$ does contain a non-trivial convergent sequence, while Y obviously does not have such a sequence.

REMARK 5.5. The reader definitely noticed that, in Lemma 5.1 and Example 5.2, we constructed α_4-spaces which are not A-spaces in a somewhat brutal (and to a certain extent trivial) way, by simply killing all non-trivial convergent sequences. Fact 1.1(ii) shows, however, that this might be a necessary approach, since such examples cannot be sequential.

REMARK 5.6. Obviously both X from Lemma 5.1 and G from Example 5.2 are even α_1-spaces.

6. Special cases when sequential A-spaces are α_4

THEOREM 6.1. *Let X be a regular sequential space such that each point of X is a G_δ-set. Then X is an α_4-space if and only if X is an A-space.*

PROOF. In view of Fact 1.1(ii) we need only to check the "if" part of our theorem. So let X be an A-space with all points G_δ. We are going to derive a contradiction by assuming that X is not an α_4-space. This assumption allows us to fix a point $x \in X$ and a countable sheaf $\{A_n : n \in \omega\}$ with the vertex x such that any sequence converging to x meets only finitely many A_n. We will also assume, without loss of generality, that $\{A_n : n \in \omega\}$ is pairwise disjoint. Pick a collection $\{U_n : n \in \omega\}$ of open subsets of X such that $\{x\} = \bigcap\{U_n : n \in \omega\}$ and $\overline{U_{n+1}} \subseteq U_n$ for $n \in \omega$. Now if $B_n = A_n \cap U_n$ for $n \in \omega$, then $S = \bigcup\{B_n : n \in \omega\} \cup \{x\}$ is closed in X, and no sequence of $\{x_n : n \in \omega\}$ with $x_n \in B_n$ converges to any point of S. Since X is sequential, we conclude that S is homeomorphic to S_ω (see Section 2 for the definition of S_ω). Thus X contains a closed copy of S_ω. Since the last space is not an A-space, this gives us a contradiction. □

For topological groups we can get especially strong result:

THEOREM 6.2. *Suppose that G is a sequential topological group such that either*

(a) *$e \in G$ is a G_δ-set, or*

(b) *G is hereditarily normal.*

Then the following conditions are equivalent:

(i) *G is an α_4-space,*
(ii) *G is an A-space, and*
(iii) *G is strongly Fréchet.*

PROOF. The implication (iii) → (ii) was proved in [3, Theorem 5.23], and the implication (i) → (ii) can be found in Fact 1.1(ii). So it remains only to check the implication (ii) → (iii). Assume that G is an A-space. Then from [27, Theorem 1.1] it follows that G contains no closed copy of S_ω. Applying Lemma 2.1 one concludes that G does not contain a closed copy of S_2. Finally, G is strongly Fréchet by [27, Theorem 3.1]. □

REMARK 6.3. Example 4.1 and Lemma 5.1 show that sequentiality of X is essential in Theorem 6.1 (note that points of a countable space are G_δ), while Theorem 1.2 demonstrates that at least some additional condition like "points are G_δ" is also necessary in Theorem 6.1. Furthermore, Example 4.2 shows that the sequentiality of G cannot be omitted in Theorem 6.2 (observe that every space with a countable network is hereditarily normal and has all points G_δ).

In conclusion we will formulate some open questions.

QUESTION 6.4. Is a sequential A-group an α_4-space?

QUESTION 6.5. Is every sequential A-group Fréchet?

Nyikos [21] showed that (i) *sequential α_4-groups are Fréchet, and* (ii) *Fréchet groups are α_4-spaces*, so Questions 6.4 and 6.5 are in fact equivalent.[4] In view of Fact 1.1(ii), a positive answer to Question 6.5 would be a strengthening of Nyikos' theorem (i). Nogura's example from Theorem 1.2 shows that Questions 6.4 and 6.5 are specific for topological groups, and their analogue for general (even compact) spaces has a negative answer. Theorem 6.2 provides a partial positive answer to both of the above questions in case when G is hereditarily normal, and when $e \in G$ is a G_δ-set.

Since countably compact spaces are A-spaces, the following particular version of Question 6.5, due to Shakhmatov [23, Problem L11], could be especially interesting:

QUESTION 6.6. Are countably compact, sequential groups Fréchet?

Quite surprisingly, a counterexample to Question 6.6 (if any) seems to be unknown. Finally, Examples 4.1 and 5.2 justify the following

QUESTION 6.7. Is there a *countable* A-group which is not α_4?

---

[4] Nyikos' property (**) from [21] is equivalent to α_4-property, so (ii) follows from [21, Theorem 4]. To get (i), one needs to combine [21, Theorem 1] with the remark in the last paragraph on p. 797 of [21] saying that (*) can be replaced by (**) in [21, Theorem 1].

REFERENCES

[1] ARENS, R., Note on convergence in topology, *Math. Mag.* **23** (1950), 229–234. *MR* **12** #271h

[2] ARHANGEL'SKII, A. V., Frequency spectrum of a topological space and classification of spaces, *Dokl. Akad. Nauk SSSR* **206** (1972), 265–268 (in Russian). *MR* **52** #15376; *Soviet Math. Dokl.* **13** (1972), 265–268.

[3] ARHANGEL'SKII, A. V., The frequency spectrum of a topological space and the product operation, *Trudy Moskov. Mat. Obshch.* **40** (1979), 171–206 (in Russian). *MR* **81i**:54017; *Trans. Moscow Math. Soc.*, Issue **2** (1981), 163–200.

[4] ARHANGEL'SKII, A. V. and FRANKLIN, S. P., Ordinal invariants for topological spaces, *Michigan Math. J.* **15** (1968), 313–320. *MR* **39** #2112

[5] ARHANGEL'SKII, A. V., OKUNEV, O. G. and PESTOV, V. G., Free topological groups over metrizable spaces, *Topology Appl.* **33** (1989), 63–76. *MR* **90h**:22002

[6] GERLITS, J. and NAGY, ZS., On Fréchet spaces, *Rend. Circ. Mat. Palermo* (2) Suppl. **18** (1988), 51–71. *MR* **89m**:54039

[7] GILLMAN, L. and JERISON, M., *Rings of continuous functions*, The University series in higher mathematics, Van Nostrand, Princeton, NJ, 1960. *MR* **22** #6994. Reprint in Graduate Texts in Mathematics, No. 43, Springer-Verlag, New York–Heidelberg, 1976. *MR* **53** #11352

[8] GRAEV, M. I., Free topological groups, *Izv. Akad. Nauk SSSR Ser. Mat.* **12** (1948), 279–324 (in Russian). *MR* **10**–11; *Amer. Math. Soc. Translation*, no. 35 (1951), 61 pp. *MR* **12**–391; Reprinted *Amer. Math. Soc. Transl.* (1) **8** (1962), 305–364. *MR* **38** #1983

[9] HEWITT, E. and ROSS, K. A., *Abstract harmonic analysis*, I, Springer-Verlag, Berlin–Heidelberg–New York, 1965.

[10] MALYHIN, V. I., Countable spaces having no bicompactifications of countable tightness, *Dokl. Akad. Nauk SSSR* **206** (1972), 1293–1296 (in Russian). *MR* **47** #9514; *Soviet. Math. Dokl.* **13** (1972), 1407–1411.

[11] MARKOV, A. A., On free topological groups, *Amer. Math. Soc. Transl.* (1) **8** (1962), 195–272.

[12] MICHAEL, E., On a theorem of Rudin and Klee, *Proc. Amer. Math. Soc.* **12** (1961), 921. *MR* **24** #A2942

[13] MICHAEL, E., A quintuple quotient quest, *General Topology Appl.* **2** (1972), 91–138. *MR* **46** #8156

[14] MICHAEL, E., Countably bi-quotient maps and A-spaces, *Topology Conference* (Virginia Polytech. Inst. and State Univ., Blacksburg, VA, 1973), Lecture Notes in Mathematics, Vol. 375, Springer-Verlag, Berlin, 1974, 183–189. *MR* **50** #11129

[15] MICHAEL, E., OLSON, R. C. and SIWIEC, F., A-spaces and countably bi-quotient maps, *Dissertationes Math.* (*Rozprawy Mat.*) **133** (1976), 5–43. *MR* **54** #6067

[16] NOGURA, T., The product of $\langle\alpha_i\rangle$-spaces, *Topology Appl.* **21** (1985), 251–259. *MR* **87d**:54006

[17] NOGURA, T., Product of sequential convergence properties, *Czechoslovak Math. J.* **39** (1989), 262–279. *MR* **90f**:54039

[18] NOGURA, T., SHAKHMATOV, D. and TANAKA, Y., Metrizability of topological groups having weak topologies with respect to good covers, *Topology Appl.* **54** (1993), 203–212. *MR* **95b**:54046

[19] NOGURA, T., SHAKHMATOV, D. and TANAKA, Y., Examples of a specific behaviour of α_i-properties and the Fréchet–Urysohn property in topological groups (to appear).

[20] NOGURA, T. and TANAKA, Y., Spaces which contain a copy of S_ω or S_2, and their applications, *Topology Appl.* **30** (1988), 51–62. *MR* **89i**:54038

[21] NYIKOS, P. J., Metrizability and the Fréchet–Urysohn property in topological groups, *Proc. Amer. Math. Soc.* **83** (1981), 793–801. *MR* **82k**:54049

[22] Nyikos, P. J., Subsets of ${}^{\omega}\omega$ and the Fréchet–Urysohn and α_i-properties, *Topology Appl.* **48** (1992), 91–116. *MR* **93k**:54011

[23] Problem Section, *Topology Proc.* **14** (1989), 375–382.

[24] Rudin, W., Homogeneity problems in the theory of Čech compactifications, *Duke Math. J.* **23** (1956), 409–419. *MR* **18**–324d

[25] Shakhmatov, D. B., α_i-properties in Fréchet–Urysohn topological groups, *Topology Proc.* **15** (1990), 143–183. *MR* **94d**:54010

[26] Siwiec, F., Sequence-covering and countably bi-quotient mappings, *General Topology and Appl.* **1** (1971), 143–154. *MR* **44** #5933

[27] Tanaka, Y., Metrizability of certain quotient spaces, *Fund. Math.* **119** (1983), 157–168. *MR* **86c**:54028

[28] Tkachuk, V. V., A method of construction of examples of M-equivalent spaces, *Uspekhi Mat. Nauk* **38** (1983), no. 6 (234), 127–128 (in Russian). *MR* **86i**:54008; *Russian Math. Surveys* **38** (1983), 135–136.

(Received November 15, 1993)

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
EHIME UNIVERSITY
MATSUYAMA 790
JAPAN

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
EHIME UNIVERSITY
MATSUYAMA 790
JAPAN

and

FACULTY OF MECHANICS AND MATHEMATICS
MOSCOW STATE UNIVERSITY
MOSCOW 119899
RUSSIA

DEPARTMENT OF MATHEMATICS
TOKYO GAKUGEI UNIVERSITY
KOGANEI-SHI
TOKYO 184
JAPAN

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 363–368

ON LASKERIAN LATTICES AND Q-LATTICES

H. M. NAKKAR and E. A. AL-KHOUJA

Let R be a commutative ring with identity. Recall [2] that a ring R is called a Q-ring if every ideal of R is a product of primary ideals. Also recall that R is called a Laskerian ring if every ideal of R is a finite intersection of primary ideals.

We define a multiplicative lattice L to be a Q-lattice (Laskerian lattice) if every element $x \neq I$ of L is a finite product (intersection) of primary elements (see [8]).

The purpose of this paper is to characterize the relation between Q-lattices and Laskerian lattices. We show that if the multiplicative lattice L is a Q-lattice, then it is Laskerian (Theorem 1). But the converse of this theorem need not be true. Our main result is that the multiplicative lattice L is a Q-lattice if and only if L is Laskerian and every nonmaximal prime element of L is multiplication (Theorem 2). This theorem generalizes the result of ([2], Theorem 10) in commutative rings to the multiplicative lattice and improve the result of [7] that every element in a Noetherian lattice is a product of primary elements if and only if every nonmaximal prime element is multiplication.

Let L be a multiplicative lattice. Recall that L is called a K-lattice if it is a CG-lattice (every element of L is a join of compact elements) and if x and y are compact elements of L, then $x \cdot y$ is a compact element of L. Also recall that L is called an R-lattice if it is a PG-lattice (every element of L is a join of principal elements) and every principal element of L is compact.

Let L be a K-lattice in which the greatest element I is compact. An element $p \in L$ is said to be prime if $p \neq I$ and if $ab \leqq p$ implies $a \leqq p$ or $b \leqq p$, for all $a, b \in L$. An element $q \in L$ is said to be primary if $q \neq I$ and if for all compact elements $a, b \in L, a \cdot b \leqq q$ implies $a \leqq q$ or $b^k \leqq q$ for some positive integer k. By the radical of an element q, we shall mean the join of all elements x having a power contained in q. We shall use the notation $\sqrt{q}$ to denote the radical of q. If q is primary, then $\sqrt{q}$ is the minimal prime containing q, and then we shall say that q is p-primary.

An element b is said to have a primary decomposition if there exist primary elements $q_1, \ldots, q_m$ such that $b = q_1 \wedge \cdots \wedge q_m$.

1991 *Mathematics Subject Classification.* Primary 06F10; Secondary 06A06.
Key words and phrases. Multiplicative lattice, lattice modules, primary element.

For any element b of a multiplicative lattice $L, V(b)$ will denote the set of all prime elements of L containing b.

Let L be a CG-lattice and let b be an element of L, then the prime element $p \in L$ belongs to $\mathrm{Ass}(b)$ ($\underset{f}{\mathrm{Ass}}(b)$) if and only if there is a compact element $h \in L$ such that $p = (b:h)$ (p is minimal in $V(b:h)$).

A non-empty subset S of L is called multiplicatively closed if it is closed under multiplication and every element of S is compact in L. Let L be a K-lattice. For every element x of L, we define $S(x) = \bigvee_{s \in S} (x:s)$. Clearly $b \sim d \Leftrightarrow S(b) = S(d)$ $(b, d \in L)$ is an equivalence relation. Let $[b]$ be the equivalence class of b and let $S^{-1}L = \{[b] : b \in L\}$. The quotient lattice $S^{-1}L$ is a multiplicative lattice. It is again K-lattice. If p is a prime element of L and $S_p = \{s : s \in L, s \nleq p$ and s is compact $\}$, then we write L_p instead of $S_p^{-1}L$.

Recall that the multiplicative lattice L has Noetherian spectrum if L satisfies the ascending chain condition for radical elements.

Let p be a prime element of L. The least upper bound of the length of chains $p > p_1 > p_2 \cdots > p_k$ where the p_i $(i = 1, \ldots, k)$ are prime elements of L is called the rank of p or the height of p ($\mathrm{ht}\, p$).

The dimension of L is the supremum of the length of all chains of distinct prime elements of L and it is denoted by $\dim L$.

Multiplication elements will be used in several places in this paper. An element b is called a multiplication element if for every element $d \leqq b$, there exists an element c with $d = cb$. In general, we adopt the lattice terminology of [3], [4], [5] and [6].

LEMMA 1. *Let L be a K-lattice in which the greatest element I is compact. Suppose that p is a prime element that is multiplication element. If p^n is p-primary for some positive integer n and if $(p^{n+1} : p^n) = p$, then p^{n+1} is p-primary.*

PROOF. Let x, y be two compact elements of L such that $x \cdot y \leqq p^{n+1}$ and $x \nleq p$. It follows that $x \cdot y \leqq p^n$. Since p^n is p-primary, then $y \leqq p^n$. By assumption p is a multiplication element, so is p^n. Hence, there is an element c of L such that $y = cp^n$ and then $x \cdot y = xcp^n \leqq p^{n+1}$. It implies that $xc \leqq (p^{n+1} : p^n) = p$ and hence $c \leqq p$. Thus $y \leqq p^{n+1}$.

LEMMA 2. *Let L be an R-lattice. Suppose that p is a minimal prime element that is a multiplication element. If $\{p^i\}_{i=1}^m$ is the set of distinct p-primary elements of L, then there exists a principal element a of L such that $a \nleq p$ with $a^t \cdot p^m \leqq p^{m+t}$ for each positive integer t.*

PROOF. Now $p < (p^{m+1} : p^m)$, because if we assume $p = (p^{m+1} : p^m)$, then Lemma 1 shows that p^{m+1} is p-primary. Hence $p^m = p^{m+1}$. It means that $p = (p^{m+1} : p^m) = I$, a contradiction.

Let a be a principal element such that $a \nleq p$ and $a \leq (p^{m+1} : p^m)$. It implies that $ap^m \leq p^{m+1}$ and hence $a^t p^m \leq p^{m+t}$ for each positive integer t.

PROPOSITION 1. *Let L be an R-lattice and let d be an element of L. Suppose that p is a multiplication prime element minimal over d. If $d = q_1 \dots q_t$ where q_i is p_i-primary $(i = 1, \dots, t)$, then there exists a principal element $a \nleq p$ such that $(d : a)$ is p-primary.*

PROOF. Since p is prime and $d \leq p$, we may assume that $q_j \leq p$ for some j $(1 \leq j \leq t)$, where q_j is p_j-primary, so $p_j = p$. Since every p-primary element of L is a power of its radical p (Lemma 4 in [7]). Then in the decomposition for d we replace every p-primary element by a power of p. Therefore by rearranging, we can obtain the following decomposition for $d, d = p^n q_1 \dots q_l$ where q_i is p_i-primary $(i = 1, \dots, l)$ and $p_1, \dots, p_l$ are primes of L, each of them different from p. Now, we show that there is a principal element $b \nleq p$ and a positive integer k such that $(d : b) \geq p^k q_1 \dots q_l$ with p^k is p-primary and $p_1, \dots, p_l$ are primes of L with $p_i \nleq p$ for each i $(1 \leq i \leq l)$. For, if rank $p = 0$ and $m < n$ where m is the least positive integer such that $p, p^2, \dots, p^m$ are the distinct p-primary elements of L (Theorem 1 in [7]), then Lemma 2 shows that there is a principal element $b \nleq p$ such that $b \cdot p^m \leq p^n$. Hence, $(d : b) \geq p^m \cdot q_1 \dots q_l$. But if rank $p = 0$ and $m \geq n$, we may take $b = I$. Finally, if rank $p > 0$, then p^n is p-primary for each positive integer n (Theorem 1 in [7]). So we may take $b = I$.

Since $\bigwedge_{i=1}^{l} p_i \nleq p$, then there exists a principal element $x \leq \bigwedge_{i=1}^{l} p_i$ with $x \nleq p$. Thus, there exists a positive integer s such that $x^s \leq q_1 \dots q_l$. This implies that $x^s \cdot p^k \leq q_1 \dots q_l p^k \leq (d : b)$. Therefore $p^k \leq (d : bx^s) \leq p$. We consider two cases. Assume first that $(d : bx^s) \leq p^k$, then $p^k = (d : bx^s)$. If we denote a by bx^s, then a is principal with $a \nleq p$ and $(d : a)$ is p-primary. Now assume that $(d : bx^s) \nleq p^k$, then there exists a positive integer $j, j \leq k$ such that $(d : bx^s) \leq p^{j-1}$ and $(d : bx^s) \nleq p^j$. Since p^{j-1} is multiplication, there is an element c of L such that $(d : bx^s) = cp^{j-1}$ and $c \nleq p$. So, there exists a principal element $y \leq c$ with $y \nleq p$. Hence $yp^{j-1} \leq (d : bx^s)$. Thus $p^{j-1} \leq (d : bx^s y) \leq p$. On the other hand $(d : bx^s y)y \leq (d : bx^s) \leq p^{j-1}$ with $y \nleq p$. Since p^{j-1} is p-primary, then $(d : bx^s y) \leq p^{j-1}$. Therefore $(d : bx^s y) = p^{j-1}$. We denote a by $bx^s y$, then a is also principal with $a \nleq p$. Thus $(d : a)$ is p-primary.

COROLLARY 1. *Let L be an R-lattice and let d be an element of L. Suppose that p is a multiplication prime element minimal over d. If L is a Q-lattice, then there exists a principal element $a \nleq p$ such that $(d : a)$ is p-primary.*

PROOF. Let $d \leq p$ be an element of L. Then $d = q_1 \dots q_n$ where $q_1, \dots, q_n$ are primary. Proposition 1 shows that there exists a principal element $a \nleq p$ such that $(d : a)$ is p-primary.

PROPOSITION 2. *Let L be an R-lattice and let d be an element of L which is contained in a finite number of minimal primes. Then, for every prime element p minimal over d and maximal, there exists a principal element a of L such that $a \nleqq p$ and $(d:a)$ is p-primary.*

PROOF. Let $p_1, p_2, \ldots, p_n, p$ be the only primes minimal over d. Since $\bigwedge_{i=1}^{n} p_i \nleqq p$, then there exists a principal element r of $L, r \leqq \bigwedge_{i=1}^{n} p_i$ with $r \nleqq p$. Since p is maximal, it follows that $p \vee r = I$. By assumption, L is PG-lattice, so there exist principal elements $x_1, \ldots, x_k$ with $x_i \leqq p$ $(i = 1, \ldots, k)$, such that $x_i \vee \cdots \vee x_k \vee r = I$. But $r(x_i \vee \cdots \vee x_k) \leqq \sqrt{d}$ (Theorem (2-4) in [1]). Then, there is a positive integer i such that $r^i(x_1 \vee \cdots \vee x_k)^i \leqq d$. We say $a = r^i$, then a is principal with $a \nleqq p$, and we have $(x_1 \vee \cdots \vee x_k)^i \leqq (d:a) \leqq p$. Now, each $p_i \ngeqq (d:a)$ $(i = 1, \ldots, n)$, because in the converse case, we obtain $I = x_1 \vee \cdots \vee x_k \vee r \leqq p_i$. Thus $I = p_i$, a contradiction.

Since $(d:a) \neq I$ and p is minimal prime over $(d:a)$, then there is a compact element t of L such that $t \nleqq p$ and $(d:a) \leqq ((d:a):t) \leqq p$. Thus $p \in \operatorname{Ass}_f(d:a)$. On the other hand, let $q \in \operatorname{Ass}_f(d:a)$. So there exists a compact element $h \in L$ such that q is minimal in $V((d:a):h)$ (Remark (2-3) in [6]).

Since $d \leqq (d:a) \leqq (d:ah) = ((d:a):h) \leqq q$, then q contains one of the elements $p_1, \ldots, p_n, p$. But $q \ngeqq p_i$ for each i $(i = 1, \ldots, n)$ because in the converse case, it follows that $q = I$. Then $q \geqq p$ and hence $p = q$, since p is maximal. Thus $\operatorname{Ass}_f((d:a)) = \{p\}$. Proposition (4-2) in [6] shows that $(d:a)$ is p-primary.

THEOREM 1. *Let L be a modular R-lattice. If L is a Q-lattice then L is Laskerian.*

PROOF. Since L is Q-lattice, then by Theorem 4 in [6], L has Noetherian spectrum. On the other hand, let $d \neq I$ be an element of L and let p be a prime element minimal over d. We consider two cases. Assume first that p is a maximal element of L, then by Proposition 2, there is a principal element a of L with $a \nleqq p$ and for which the element $(d:a)$ is p-primary. Now assume that p is not maximal, then p is multiplication (Theorem 4 in [8]). Hence, by Proposition 1, there is a principal element a of L with $a \nleqq p$ and for which the element $(d:a)$ is p-primary. Thus L is a Laskerian lattice (Theorem 2 in [8]).

REMARK. Let L be a multiplicative lattice. If L is Laskerian, then L need not to be a Q-lattice. For example, the polynomial ring $R = K[X_1, \ldots, X_n]$ in a finite number $n (n \geqq 3)$ of indeterminates over a field K is a Noetherian ring. If we take $L(R)$ the lattice of ideals of R, then $L(R)$ is a Noetherian lattice in which every element has a primary decomposition. Therefore $L(R)$ is a Laskerian lattice and $\dim L(R) \geqq 3$. But this lattice is not a Q-lattice, because the Q-lattices have dimension at most two.

THEOREM 2. *Let L be a modular R-lattice. Then L is a Q-lattice if and only if L is Laskerian and every nonmaximal prime element of L is a multiplication element.*

PROOF. ($\Rightarrow$) Suppose that L is a Q-lattice. By Theorem 1, L is a Laskerian lattice. By Theorem 4 in [8], it follows that every nonmaximal prime element of L is a multiplication element. ($\Leftarrow$) Suppose that L is a Laskerian lattice in which every nonmaximal prime element is a multiplication element. Then by Theorem 2 in [7], every element of L is a product of primary elements, so L is a Q-lattice.

PROPOSITION 3. *Let L be a K-lattice in which the greatest element I is compact and a an element of L having a primary decomposition. Let b be any element of L and let S be a multiplicatively closed subset of L. Then $[(a:b)]_s=[a]_s:[b]_s$ in $S^{-1}L$.*

PROOF. Let $a=q_1\wedge\cdots\wedge q_n$ be a primary decomposition of a. Since

$$[(q_1\wedge\cdots\wedge q_n):b]_s=[(q_1:b)\wedge\cdots\wedge(q_n:b)]_s=$$
$$=[(q_1:b)]_s\wedge\cdots\wedge[(q_n:b)]_s.$$

So, we can assume that a is p-primary. Of course, we always have that $[(a:b)]_s\leqq[a]_s:[b]_s$. If $s\leqq a$ for some $s\in S$, then $s\leqq(a:b)$ and so, $[(a:b)]_s=[I]_s$. It follows that $[a]_s:[b]_s=[I]_s$. Hence $[(a:b)]_s=[a]_s:[b]_s$. So we can assume that $s\nleqq a$ for every $s\in S$. Let h be a compact element of L with $[h]\leqq[a]_s:[b]_s$. Then $[h]_s[b]_s\leqq[a]_s$. Therefore, $h\cdot b\leqq S(a)=a$, and hence $h\leqq(a:b)$. It implies that $[h]_s\leqq[(a:b)]_s$. Thus $[a]_s:[b]_s\leqq[(a:b)]_s$.

THEOREM 3. *Let L be a K-lattice in which the greatest element I is compact and every principal element not equal to I has a primary decomposition. For an element a of L which is a join of principal elements the following statements are equivalent:*

(1) *a is a multiplication element.*
(2) *a is locally principal.*
(3) *a is finitely generated and locally principal.*
(4) *a is principal element.*

PROOF. (1)$\Rightarrow$(2). Let $a=\bigvee_\alpha x_\alpha$ where x_α is principal of L, and let p be a prime element of L. By passing to L_p, it follows that $[a]_p=\bigvee_\alpha[x_\alpha]_p$. Since a is a multiplication element, so is $[a]_p$ in L_p. This implies that $[a]_p=[x_{\alpha_i}]_p$ for some index i, by Proposition (1-1) and Theorem (1-2) in [1].

(2)$\Rightarrow$(3). Let $a=\bigvee_\alpha a_\alpha$ where a_α is principal, then for any maximal element p of L we have: $[a]_p=[\bigvee_\alpha a_\alpha]_p=\bigvee_\alpha[a_\alpha]_p$. By Theorem (2-3) in [4] there exists an index β such that $[a]_p=[a_\beta]_p$. On the other hand, let $\Theta(a)=$

$\bigvee_\alpha (a_\alpha : a)$ and let m be a maximal element of L. Then by Proposition 3, we get that $[\Theta(a)]_m = \bigvee_\alpha [(a_\alpha : a)]_m = \bigvee_\alpha ([a_\alpha]_m : [a]_m)$ since each principal element of L has primary decomposition. But $[a]_m = [a_\beta]_m$ for some $a_\beta \leqq a$ since a is locally principal. Thus $[\Theta(a)]_m = [I]_m$ for every maximal element m of L, so $\Theta(a) = I$ (Proposition (5-3) in [5]). Since I is compact we get that $I = (a_{\alpha_1} : a) \vee \cdots \vee (a_{\alpha_n} : a) \leqq (a_{\alpha_1} \vee \cdots \vee a_{\alpha_n} : a)$, where a_{α_i} is principal with $a_{\alpha_i} \leqq a$. This implies that $a = a_{\alpha_1} \vee \cdots \vee a_{\alpha_n}$.

(3)⇒(4) and (4)⇒(1) are always true.

THEOREM 4. *Let L be a modular R-lattice. L is a Q-lattice if and only if L is Laskerian and every nonmaximal prime element of L is locally principal.*

PROOF. It follows immediately from Theorem 2 and Theorem 3.

REFERENCES

[1] ANDERSON, D. D., Abstract commutative ideal theory without chain condition, *Algebra Universalis* **6** (1976), 131–145. *MR* **54** #7332

[2] ANDERSON, D. D. and MAHANEY, L. A., Commutative rings in which every ideal is a product of primary ideals, *J. Algebra* **106** (1987), 528–535. *MR* **87m**:13021

[3] DILWORTH, R. P., Abstract commutative ideal theory, *Pacific J. Math.* **12** (1962), 481–498. *MR* **26** #1333

[4] JOHNSON, J. A., Totally quasilocal multiplicative lattices, *Tamkang J. Math.* **13** (1982), 91–104. *MR* **84d**:06023

[5] NAKKAR, H. M., Localization in multiplicative lattices, *Mat. Issled.* **9** (1974), no. 2 (32), 88–108 (in Russian). *MR* **50** #2013

[6] NAKKAR, H. M. and ANDERSON, D. D., Associated and weakly associated prime elements and primary decomposition in lattice modules, *Algebra Universalis* **25** (1988), 196–209. *MR* **89g**:06034

[7] NAKKAR, H. M. and AL-KHOUJA, E. A., Noetherian lattice in which every element is a product of primary elements, *Order* **12** (1995), 413–420.

[8] NAKKAR, H. M. and AL-KHOUJA, E. A., A characterization of local Q-lattices, *Damascus Univ. J.* **9** (1993), 13–27.

(Received March 27, 1994)

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
UNIVERSITY OF ALEPPO
ALEPPO
SYRIA

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 369–391

ON THE POINTWISE APPROXIMATION BY BOREL AND EULER MEANS

B. ROSZAK

Abstract

The rates of approximation of Lebesgue-integrable functions by the Borel and Euler means of its Fourier series are estimated by the characteristics created by the relation defining the Lebesgue-type points. Some corollaries for Lipschitz functions are also derived.

Introduction

Let $L^p_{2\pi}$, with $1 \leqq p < \infty$, be the class of all 2π-periodic complex-valued functions that are Lebesgue-integrable with p-th power in the interval $[-\pi, \pi]$, i.e. $f \in L^p_{2\pi}$ if $\int_{-\pi}^{\pi} |f(x)|^p dx < \infty$. Given any function $f \in L^p_{2\pi}$, let $S_n[f]$ be the n-th partial sum of its Fourier series. Introduce the Borel and Euler means of this series:

$$B_r[f](x) = e^{-r} \sum_{k=0}^{\infty} \frac{r^k}{k!} S_k[f](x), \quad \text{for every } r > 0, \tag{1}$$

$$E_n[f](x) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} S_k[f](x), \quad \text{for } n = 0, 1, 2, \ldots \tag{2}$$

(see e.g. [1], p. 79 and p. 70, respectively).

As a measure of deviation of those means from $f(x)$ in a fixed point $x \in \mathbf{R}$ we will use the quantities:

$$w_x[f](\delta) = \frac{1}{\delta} \int_0^{\delta} |\phi_x(t)| dt,$$

$$\overline{w}_{x,a}[f](\delta) = \frac{1}{\delta} \int_0^{\delta} |\phi_x(t+a) - \phi_x(t)| dt,$$

1991 *Mathematics Subject Classification.* Primary 42A10; Secondary 42A24, 40G05, 40G10.

Key words and phrases. Borel means, Euler means.

where $\phi_x(t) = \frac{1}{2}(f(x+t) + f(x-t) - 2f(x))$ and $\delta > 0$, $a > 0$.

The function $f \in L^p_{2\pi}$ is said to satisfy a Lipschitz condition of order α, with $0 < \alpha \leqq 1$ (notation $f \in \mathrm{Lip}\,(\alpha, p)$), if

$$\omega[f](\delta)_p \leqq M\delta^\alpha$$

for some constant $M > 0$ and all $\delta > 0$, where

$$\omega[f](\delta)_p = \sup_{|t|\leqq\delta} \|\phi_\cdot(t)\|_p = \sup_{|t|\leqq\delta}\left(\int_{-\pi}^{\pi} |\phi_x(t)|^p dx\right)^{1/p}$$

is the modulus of continuity of f.

In the case when $p = \infty$, we write $f \in L^\infty_{2\pi}$ if $\|f\|_\infty = \operatorname*{ess\,sup}_{-\pi\leqq x\leqq\pi} |f(x)| < \infty$ and the above definitions remain unchanged. Of course, if $C_{2\pi}$ denotes the class of all 2π-periodic continuous complex-valued functions, then $C_{2\pi} \subset L^\infty_{2\pi}$ and

$$\|f\| = \max_{-\pi\leqq x\leqq\pi} |f(x)| = \operatorname*{ess\,sup}_{-\pi\leqq x\leqq\pi} |f(x)| = \|f\|_\infty \quad \text{for } f \in C_{2\pi}.$$

Simple calculations show that the Borel and Euler means defined by (1) and (2) can be represented in the forms

$$B_r[f](x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_r(t)dt,$$

where

$$K_r(t) = e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r\sin t + \frac{1}{2}t)}{\sin\frac{1}{2}t} \qquad (r > 0),$$

and

$$E_n[f](x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_n(t)dt,$$

where

$$K_n(t) = \cos^n\left(\frac{1}{2}t\right)\frac{\sin\frac{1}{2}(n+1)t}{\sin\frac{1}{2}t}, \qquad (n = 0, 1, 2, \dots)$$

respectively, (see e.g. [1], p. 364 and [2]).

Hence

$$B_r[f](x) - f(x) = \frac{1}{4\pi}\int_{-\pi}^{\pi} (f(x+t) + f(x-t) - 2f(x))K_r(t)dt$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_x(t)K_r(t)dt = \frac{1}{\pi}\int_0^{\pi} \phi_x(t)K_r(t)dt,$$

and in the same way

$$E_n[f](x)-f(x)=\frac{1}{\pi}\int\limits_0^{\pi}\phi_x(t)K_n(t)dt.$$

Statement of results

We present here four approximation theorems.

THEOREM 1. *Let* $\beta\in(\frac{1}{3},\frac{1}{2})$ *and put*

$$a_r=\frac{\pi}{r+\frac{1}{2}}\quad\text{and}\quad b_r=\frac{\pi}{(r+\frac{1}{2})^{\beta}}\quad\text{for}\quad r\geqq\frac{1}{2}.$$

Then there exists a constant $M>0$ *such that*

$$|B_r[f](x)-f(x)|\leqq$$

$$M\left(w_x(2a_r)+b_r^2\int\limits_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt+\overline{w}_{x,a_r}(b_r)+\int\limits_{a_r}^{b_r}\frac{\overline{w}_{x,a_r}(t)}{t}dt+\frac{1}{a_r}\int\limits_{a_r}^{b_r}|\phi_x(t)|t^2dt\right)$$

for any $f\in L^1_{2\pi}$, $r\geqq\frac{1}{2}$ *and* $x\in\mathbf{R}$.

THEOREM 2. *Let* $\beta\in(\frac{1}{3},\frac{1}{2})$ *and put*

$$a_n=\frac{2\pi}{n+1}\quad\text{and}\quad b_n=\frac{2\pi}{(n+1)^{\beta}}\quad\text{for}\quad n=0,1,2,\ldots.$$

Then there exists a constant $M>0$ *such that*

$$|E_n[f](x)-f(x)|\leqq$$

$$M\left(w_x(2a_n)+b_n^2\int\limits_{b_n}^{2\pi}\frac{w_x(t)}{t^2}dt+\overline{w}_{x,a_n}(b_n)+\int\limits_{a_n}^{b_n}\frac{\overline{w}_{x,a_n}(t)}{t}dt\right)$$

for any $f\in L^1_{2\pi}$, $n=0,1,2,\ldots$ *and* $x\in\mathbf{R}$.

THEOREM 1′. *There exists a constant* $M>0$ *such that*

$$\|B_r[f]-f\|_p\leqq$$

$$M\left(\omega(2a_r)_p+b_r^2\int\limits_{b_r}^{2\pi}\frac{\omega(t)_p}{t^2}dt+\left\|\int\limits_{a_r}^{b_r}\frac{|\phi_.(t)-\phi_.(t+a_r)|}{t}dt\right\|_p+\frac{1}{a_r}\int\limits_{a_r}^{b_r}\|\phi_.(t)\|_pt^2dt\right)$$

for any $f \in L^p_{2\pi}$, *with* $1 \leqq p \leqq \infty$, $r \geqq \frac{1}{2}$ *and* $x \in \mathbf{R}$, *where* a_r *and* b_r *are the same as in Theorem 1.*

THEOREM 2′. *There exists a constant* $M > 0$ *such that*

$$\|E_n[f] - f\|_p \leqq M\left(\omega(2a_r)_p + b_n^2 \int_{b_n}^{2\pi} \frac{\omega(t)_p}{t^2} dt + \left\| \int_{a_n}^{b_n} \frac{|\phi_.(t) - \phi_.(t + a_n)|}{t} dt \right\|_p \right)$$

for any $f \in L^p_{2\pi}$, *with* $1 \leqq p \leqq \infty$, $n = 0, 1, 2, \ldots$ *and* $x \in \mathbf{R}$, *where* a_n *and* b_n *are the same as in Theorem 2.*

Applying Theorem 1′ we can readily derive the following

COROLLARY 1. *Let* $0 < \alpha < 1$ *and* $1 \leqq p \leqq \infty$. *Put* $a_r = \frac{\pi}{r + \frac{1}{2}}$ *for* $r \geqq \frac{1}{2}$. *If* $f \in \operatorname{Lip}(\alpha, p)$ *is such that*

$$\left\| \int_{a_r}^{\pi} \frac{|\phi_.(t) - \phi_.(t + a_r)|}{t} dt \right\| \leqq N r^{-\alpha}$$

for some $N > 0$ *and all* $r \geqq \frac{1}{2}$, *then there exists* $M > 0$ *that*

$$\|B_r[f] - f\|_p \leqq M r^{-\alpha}$$

for all $r \geqq \frac{1}{2}$.

Analogously, from Theorem 2′ we can obtain

COROLLARY 2. *Let* $0 < \alpha < 1$ *and* $1 \leqq p \leqq \infty$. *Put* $a_n = \frac{2\pi}{n+1}$ *for* $n = 1, 2, \ldots$. *If* $f \in \operatorname{Lip}(\alpha, p)$ *is such that*

$$\left\| \int_{a_n}^{\pi} \frac{|\phi_.(t) - \phi_.(t + a_n)|}{t} dt \right\| \leqq N n^{-\alpha}$$

for some $N > 0$ *and* $n = 1, 2, \ldots$, *then there exists* $M > 0$ *that*

$$\|E_n[f] - f\|_p \leqq M n^{-\alpha}$$

for $n = 1, 2, \ldots$.

REMARK 1. The similar result as in the above corollary was obtained by Chui and Holland in [2] for $f \in \operatorname{Lip} \alpha$, with $0 < \alpha < 1$, i.e. for $f \in C_{2\pi}$ and satisfying a Lipschitz condition of order α.

REMARK 2. We mention here the following simple consequence of our theorems that was also obtained in [3]: If $\omega(t)$ is the modulus of continuity of $f \in C_{2\pi}$, then there is a constant $M > 0$ such that

$$\|B_r[f] - f\| \leqq M \int_1^{2(r+\frac{1}{2})} \frac{\omega(\frac{1}{t})}{t} dt \quad \text{for } r \geqq 1, \text{ and}$$

$$\|E_n[f] - f\| \leqq M \sum_{k=1}^{n+1} \frac{\omega(\frac{1}{k})}{k} \quad \text{for } n = 1, 2, \dots .$$

Lemmas

We will need further a few auxiliary results.

LEMMA 1. *The following relations hold for the quantities* $w_x[f]$ *and* $\overline{w}_{x,a_r}[f]$*:*

(3) $$w_x(\delta) \leqq 2w_x(2\delta), \qquad \text{for } \delta > 0,$$

(4) $$\overline{w}_{x,\alpha}(\delta) \leqq 4w_x(2\delta), \qquad \text{for } \delta \geqq \alpha \geqq 0.$$

PROOF OF LEMMA 1. Indeed, for (3) we have

$$w_x(\delta) = \frac{1}{\delta} \int_0^{\delta} |\phi_x(t)| dt \leqq 2 \cdot \frac{1}{2\delta} \int_0^{2\delta} |\phi_x(t)| dt = 2w_x(2\delta),$$

and (4) follows from

$$\overline{w}_{x,\alpha}(\delta) = \frac{1}{\delta} \int_0^{\delta} |\phi_x(t+\alpha) - \phi_x(t)| dt \leqq$$

$$\leqq \frac{1}{\delta} \int_0^{\delta} |\phi_x(t+\alpha)| dt + \frac{1}{\delta} \int_0^{\delta} |\phi_x(t)| dt = \frac{1}{\delta} \int_{\alpha}^{\delta+\alpha} |\phi_x(t)| dt + \frac{1}{\delta} \int_0^{\delta} |\phi_x(t)| dt \leqq$$

$$\leqq 2 \cdot \frac{1}{2\delta} \int_0^{2\delta} |\phi_x(t)| dt + 2 \cdot \frac{1}{2\delta} \int_0^{2\delta} |\phi_x(t)| dt = 4w_x(2\delta).$$ □

LEMMA 2. *If* $a_r = \frac{\pi}{r+\frac{1}{2}}$ *and* $b_r = \frac{\pi}{(r+\frac{1}{2})^{\beta}}$, *with* $\frac{1}{3} < \beta < \frac{1}{2}$, *then, for* $r \geqq 12$, *the following relations hold:*

(5) $$w_x(2b_r) \leqq 5b_r \int_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt,$$

$$(6)\qquad \int_{a_r}^{b_r} \frac{|\phi_x(t)-\phi_x(t+a_r)|}{t}dt = \overline{w}_{x,a_r}(b_r) - \overline{w}_{x,a_r}(a_r) + \int_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t}dt.$$

LEMMA 3. *If* $a_n = \frac{2\pi}{n+1}$ *and* $b_n = \frac{2\pi}{(n+1)^\beta}$, *with* $\frac{1}{3} < \beta < \frac{1}{2}$, *then, for* $n \geqq 12$, *the following relations hold:*

$$(7)\qquad w_x(2b_n) \leqq 20b_n \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2}dt,$$

$$(8)\qquad \int_{a_n}^{b_n} \frac{|\phi_x(t)-\phi_x(t+a_n)|}{t}dt = \overline{w}_{x,a_n}(b_n) - \overline{w}_{x,a_n}(a_n) + \int_{a_n}^{b_n} \frac{\overline{w}_{x,a_n}(t)}{t}dt.$$

PROOF OF LEMMA 2. Initially it is

$$b_r^2 \int_{2b_r}^{2\pi} \frac{dt}{t^3} = b_r^2\Big(-\frac{1}{2t^2}\Big)\Big|_{2b_r}^{2\pi} = \frac{1}{8} - \frac{1}{8(r+\frac{1}{2})^{2\beta}} > \frac{1}{8} - \frac{1}{8r^{2\beta}} >$$

$$> \frac{1}{8} - \frac{1}{8r^{\frac{2}{3}}} > \frac{1}{8} - \frac{1}{8\cdot 5} = \frac{1}{10}$$

since $r^{\frac{2}{3}} \geqq 5$ when $r \geqq 12$. Next,

$$w_x(2b_r) = 10\cdot\frac{1}{10}w_x(2b_r) < 10b_r^2\int_{2b_r}^{2\pi} w_x(2b_r)\frac{dt}{t^3} =$$

$$= 10b_r^2 \int_{2b_r}^{2\pi}\Big(\frac{1}{2b_r}\int_0^{2b_r}|\phi_x(u)|du\Big)\frac{dt}{t^3} = 5b_r\int_{2b_r}^{2\pi}\Big(\frac{1}{t}\int_0^{2b_r}|\phi_x(u)|du\Big)\frac{dt}{t^2} \leqq$$

$$\leqq 5b_r\int_{2b_r}^{2\pi}\Big(\frac{1}{t}\int_0^{t}|\phi_x(u)|du\Big)\frac{dt}{t^2} = 5b_r\int_{2b_r}^{2\pi}\frac{w_x(t)}{t^2}dt \leqq 5b_r\int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt,$$

thus the inequality (5) holds.

Further, integration by parts yields

$$\int_{a_r}^{b_r}\frac{|\phi_x(t)-\phi_x(t+a_r)|}{t}dt = \int_{a_r}^{b_r}\frac{1}{t}\frac{d}{dt}\Big(\int_0^t |\phi_x(u)-\phi_x(u+a_r)|du\Big)dt =$$

$$=\frac{1}{t}\int\limits_0^t |\phi_x(u)-\phi_x(u+a_r)|du\Bigg|_{a_r}^{b_r} - \int\limits_{a_r}^{b_r}\left[\left(-\frac{1}{t^2}\right)\int\limits_0^t |\phi_x(u)-\phi_x(u+a_r)|du\right]dt=$$

$$=\overline{w}_{x,a_r}(b_r)-\overline{w}_{x,a_r}(a_r)+\int\limits_{a_r}^{b_r}\frac{\overline{w}_{x,a_r}(t)}{t}dt.$$

That ends the proof of the lemma. □

PROOF OF LEMMA 3 is very similar to the preceding one. Initially we make the following estimate valid for $n\geqq 12$:

$$b_n^2\int\limits_{2b_n}^{2\pi}\frac{dt}{t^3}=b_n^2\left(-\frac{1}{2t^2}\right)\Bigg|_{2b_n}^{2\pi}=\frac{1}{8}-\frac{4\pi^2}{8\pi^2(n+1)^{2\beta}}>\frac{1}{8}-\frac{1}{2n^{2\beta}}>$$

$$>\frac{1}{8}-\frac{1}{2n^{\frac{2}{3}}}>\frac{1}{8}-\frac{1}{2\cdot 5}=\frac{1}{40}.$$

Taking b_n instead of b_r and $\frac{1}{40}$ instead of $\frac{1}{10}$ and proceeding similarly as in the proof of the inequality (5) we can obtain the (7) one. Further, similarly as in (6), integration by parts yields (8).

That completes the proof of Lemma 3. □

To estimate the order of magnitude of $w_x(\pi)$ we use the next lemma:

LEMMA 4. *The following inequalities hold for the function $w_x[f]$:*

$$w_x(\pi)\ln\sqrt{2}\leqq\int\limits_{\pi}^{2\pi}\frac{w_x(t)}{t}dt\leqq w_x(\pi)\ln 4.$$

PROOF OF LEMMA 4. If $\pi\leqq t\leqq 2\pi$ then

(9) $$w_x(\pi)\leqq 2w_x(t)\quad\text{and}\quad w_x(t)\leqq 2w_x(\pi).$$

Indeed,

$$w_x(t)=\frac{1}{t}\int\limits_0^{\pi}|\phi_x(u)|du+\frac{1}{t}\int\limits_{\pi}^{t}|\phi_x(u)|du\geqq\frac{1}{t}\int\limits_0^{\pi}|\phi_x(u)|du\geqq$$

$$\geqq\frac{1}{2\pi}\int\limits_0^{\pi}|\phi_x(u)|du=\frac{1}{2}w_x(\pi),$$

and, since ϕ_x is even and 2π-periodic function for every $x\in\mathbf{R}$, we can write

$$w_x(t)=\frac{1}{t}\int_0^t|\phi_x(u)|du\leqq\frac{1}{t}\int_0^{2\pi}|\phi_x(u)|du=\frac{1}{t}\left(\int_0^{\pi}|\phi_x(u)|du+\int_{\pi}^{2\pi}|\phi_x(u)|du\right)=$$

$$=\frac{1}{t}\left(\int_0^{\pi}|\phi_x(u)|du+\int_{-\pi}^{0}|\phi_x(u)|du\right)=\frac{2}{t}\int_0^{\pi}|\phi_x(u)|du\leqq\frac{2}{\pi}\int_0^{\pi}|\phi_x(u)|du=2w_x(\pi).$$

Now using the inequalities (9) we find that

$$\frac{1}{2}w_x(\pi)\ln 2=\int_{\pi}^{2\pi}\frac{1}{2}w_x(\pi)\frac{1}{t}dt\leqq\int_{\pi}^{2\pi}\frac{w_x(t)}{t}dt\leqq 2w_x(\pi)\int_{\pi}^{2\pi}\frac{1}{t}dt=2w_x(\pi)\ln 2,$$

that completes the proof of lemma. □

Proofs of main results

For convenience, throughout the proofs, M and N stand for positive constants not necessarily the same at each occurrence.

PROOF OF THEOREM 1. First assume that $r\geqq 12$. In this case let us write

$$|B_r[f](x)-f(x)|=\left|\frac{1}{\pi}\left(\int_0^{a_r}+\int_{a_r}^{b_r}+\int_{b_r}^{\pi}\right)\phi_x(t)K_r(t)dt\right|\leqq|\varrho_1|+|\varrho_2|+|\varrho_3|,$$

say, where as previously $a_r=\frac{\pi}{r+\frac{1}{2}}$ and $b_r=\frac{\pi}{(r+\frac{1}{2})^{\beta}}$, with $\beta\in(\frac{1}{3},\frac{1}{2})$.

Now

$$|\varrho_1|=\left|\frac{1}{\pi}\int_0^{a_r}\phi_x(t)e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r\sin t+\frac{1}{2}t)}{\sin\frac{1}{2}t}dt\right|\leqq$$

$$\leqq\frac{1}{\pi}\int_0^{a_r}|\phi_x(t)|\frac{r\sin t+\frac{1}{2}t}{\frac{1}{\pi}t}dt\leqq\frac{1}{\pi}\cdot\pi\int_0^{a_r}|\phi_x(t)|\frac{(r+\frac{1}{2})t}{t}dt=$$

$$=\pi\cdot\frac{1}{a_r}\int_0^{a_r}|\phi_x(t)|dt=\pi w_x(a_r).$$

In this way, by inequality (3),

$$|\varrho_1|\leqq\pi w_x(a_r)\leqq 2\pi w_x(2a_r)=Mw_x(2a_r).$$

Let us pass now to the integral ϱ_3. Then,

$$\pi|\varrho_3| = \left|\int_{b_r}^{\pi} \phi_x(t) e^{-2r\sin^2\frac{1}{2}t} \frac{\sin(r\sin t + \frac{1}{2}t)}{\sin\frac{1}{2}t} dt\right| \leqq$$

$$\leqq \int_{b_r}^{\pi} |\phi_x(t)| e^{-2r\sin^2\frac{1}{2}t} \frac{1}{\frac{1}{\pi}t} dt = \pi \int_{b_r}^{\pi} \frac{|\phi_x(t)|}{t} e^{-2r\sin^2\frac{1}{2}t} dt.$$

It is clear that, for $r \geqq \frac{1}{2}$ and $b_r < t < \pi$, we have

$$e^{-2r\sin^2\frac{1}{2}t} \leqq e^{-2r\sin^2\frac{1}{2}b_r} \leqq e^{-2r(\frac{1}{\pi}b_r)^2} = e^{-2r\frac{1}{(r+\frac{1}{2})^{2\beta}}} \leqq$$

$$\leqq \exp\left[-\left(r+\frac{1}{2}\right)\frac{1}{(r+\frac{1}{2})^{2\beta}}\right] = \exp\left[-\left(r+\frac{1}{2}\right)^{1-2\beta}\right].$$

It could be shown that for every $\beta \in (0, \frac{1}{2})$ there exists a constant $M > 0$ (that depends on β) such that $\exp\left[-(r+\frac{1}{2})^{1-2\beta}\right] \leqq M\frac{\pi^2}{(r+\frac{1}{2})^{2\beta}}$. Using this fact and partial integration we find that

$$|\varrho_3| \leqq \exp\left[-\left(r+\frac{1}{2}\right)^{1-2\beta}\right] \int_{b_r}^{\pi} \frac{|\phi_x(t)|}{t} dt \leqq M b_r^2 \int_{b_r}^{\pi} \frac{|\phi_x(t)|}{t} dt =$$

$$= M b_r^2 \int_{b_r}^{\pi} \frac{1}{t}\left(\frac{d}{dt}\int_0^t |\phi_x(u)| du\right) dt =$$

$$= M b_r^2 \left(\frac{1}{t}\int_0^t |\phi_0^t| \phi_x(u)| du\Bigg|_{b_r}^{\pi} - \int_{b_r}^{\pi}\left(-\frac{1}{t^2}\right)\left(\int_0^t |\phi_x(u)| du\right) dt\right) =$$

$$= M b_r^2 \left(w_x(\pi) - w_x(b_r) + \int_{b_r}^{\pi} \frac{w_x(t)}{t} dt\right).$$

Whence, by Lemma 1.4 we may write

$$|\varrho_3| \leqq M b_r^2 \left(\frac{1}{\ln\sqrt{2}} \int_{\pi}^{2\pi} \frac{w_x(t)}{t} dt + \frac{1}{\ln\sqrt{2}} \int_{b_r}^{\pi} \frac{w_x(t)}{t} dt\right) = \frac{M}{\ln\sqrt{2}} b_r^2 \int_{b_r}^{2\pi} \frac{w_x(t)}{t} dt \leqq$$

$$\leqq \frac{2\pi M}{\ln\sqrt{2}} b_r^2 \int_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt = M b_r^2 \int_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt.$$

The study of ϱ_2 is more complicated and requires some calculations. The main idea of the proof is to manage the task in the following way:

$$2\pi\varrho_2 = 2\int_{a_r}^{b_r} \phi_x(t)e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r\sin t+\frac{1}{2}t)}{\sin\frac{1}{2}t}dt =$$

$$= 2\int_{a_r}^{b_r} \phi_x(t)e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r\sin t+\frac{1}{2}t)-\sin(rt+\frac{1}{2}t)}{\sin\frac{1}{2}t}dt +$$

$$+2\int_{a_r}^{b_r} \phi_x(t)e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r+\frac{1}{2})t}{\sin\frac{1}{2}t}dt =$$

$$= I_0 + \int_{a_r}^{b_r} \phi_x(t)e^{-2r\sin^2\frac{1}{2}t}\frac{\sin(r+\frac{1}{2})t}{\sin\frac{1}{2}t}dt -$$

$$- \int_0^{b_r-a_r} \phi_x(t+a_r)e^{-2r\sin^2\frac{1}{2}(t+a_r)}\frac{\sin(r+\frac{1}{2})t}{\sin\frac{1}{2}(t+a_r)}dt =$$

$$= I_0 + \int_{a_r}^{b_r} \frac{\phi_x(t)-\phi_x(t+a_r)}{\sin\frac{1}{2}t}e^{-2r\sin^2\frac{1}{2}t}\sin(r+\frac{1}{2})t\,dt +$$

$$+ \int_{a_r}^{b_r} \frac{\phi_x(t+a_r)}{\sin\frac{1}{2}t}(e^{-2r\sin^2\frac{1}{2}t}-e^{-2r\sin^2\frac{1}{2}(t+a_r)})\sin\Big(r+\frac{1}{2}\Big)t\,dt +$$

$$+ \int_{a_r}^{b_r} \phi_x(t+a_r)e^{-2\sin^2\frac{1}{2}(t+a_r)}\bigg(\frac{1}{\sin\frac{1}{2}t}-\frac{1}{\sin\frac{1}{2}(t+a_r)}\bigg)\sin\Big(r+\frac{1}{2}\Big)t\,dt -$$

$$- \int_0^{a_r} \frac{\phi_x(t+a_r)}{\sin\frac{1}{2}(t+a_r)}e^{-2r\sin^2\frac{1}{2}(t+a_r)}\sin\Big(r+\frac{1}{2}\Big)t\,dt +$$

$$+ \int_{b_r-a_r}^{b_r} \frac{\phi_x(t+a_r)}{\sin\frac{1}{2}(t+a_r)}e^{-2r\sin^2\frac{1}{2}(t+a_r)}\sin\Big(r+\frac{1}{2}\Big)t\,dt =$$

$$= I_0 + I_1 + I_2 + I_3 + I_4 + I_5, \quad \text{say.}$$

Now, for I_0 we obtain

$$|I_0| = \left| 2\int_{a_r}^{b_r} \phi_x(t) e^{-2r\sin^2 \frac{1}{2}t} \frac{\sin(r\sin t + \frac{1}{2}t) - \sin(rt + \frac{1}{2}t)}{\sin \frac{1}{2}t} dt \right| =$$

$$= \left| 2\int_{a_r}^{b_r} \phi_x(t) e^{-2r\sin^2 \frac{1}{2}t} \frac{2\cos \frac{r\sin t + (r+1)t}{2} \sin \frac{r(\sin t - t)}{2}}{\sin \frac{1}{2}t} dt \right|$$

$$\leqq 2\int_{a_r}^{b_r} |\phi_x(t)| \frac{r|\sin t - t|}{\frac{1}{\pi}t} dt =$$

$$= 2\pi \int_{a_r}^{b_r} |\phi_x(t)| \frac{r}{t} \left| -\frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \frac{t^9}{9!} - \dots \right| dt =$$

$$= 2\pi \int_{a_r}^{b_r} |\phi_x(t)| r t^2 \left| -\frac{1}{3!} + \frac{t^2}{5!} - \frac{t^4}{7!} + \frac{t^6}{9!} - \dots \right| dt \leqq$$

$$\leqq 2\pi e^{\pi}\left(r + \frac{1}{2}\right) \int_{a_r}^{b_r} |\phi_x(t)| t^2 dt = M \frac{1}{a_r} \int_{a_r}^{b_r} |\phi_x(t)| t^2 dt.$$

Next, we will estimate the integral I_1 as follows

$$|I_1| = \left| \int_{a_r}^{b_r} \frac{\phi_x(t) - \phi_x(t + a_r)}{\sin \frac{1}{2}t} e^{-2r\sin^2 \frac{1}{2}t} \sin\left(r + \frac{1}{2}\right) t \, dt \right| \leqq$$

$$\leqq \pi \int_{a_r}^{b_r} \frac{|\phi_x(t) - \phi_x(t + a_r)|}{t} dt.$$

Equality (6) implies that

$$|I_1| \leqq \pi \left(\overline{w}_{x,a_r}(b_r) - \overline{w}_{x,a_r}(a_r) + \int_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t} dt \right) \leqq$$

$$\leqq \pi \overline{w}_{x,a_r}(b_r) + \pi \int_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t} dt,$$

since $\overline{w}_{x,a_r}(a_r) \geqq 0$.

Now, let us consider the integral

$$|I_2| = \left| \int_{a_r}^{b_r} \frac{\phi_x(t+a_r)}{\sin \frac{1}{2}t} \left(e^{-2r \sin^2 \frac{1}{2}t} - e^{-2r \sin^2 \frac{1}{2}(t+a_r)} \right) \sin\left(r + \frac{1}{2}\right) t \, dt \right|.$$

By the mean-value theorem there exists a $\xi_t \in (t, t+a_r)$ such that

$$e^{-2r \sin^2 \frac{1}{2}t} - e^{-2r \sin^2 \frac{1}{2}(t+a_r)} = r a_r e^{-2r \sin^2 \frac{1}{2}\xi_t} \sin \xi_t.$$

Here it is $a_r < t < \xi_t < t + a_r < 2t$, so we get

$$|I_2| \leqq \int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)|}{\frac{1}{\pi}t} r a_r 2t \, dt \leqq \pi \int_{a_r}^{b_r} |\phi_x(t+a_r)| \left(r + \frac{1}{2}\right) \frac{\pi}{r + \frac{1}{2}} dt =$$

$$= \pi^2 \int_{2a_r}^{b_r + a_r} |\phi_x(t)| dt \leqq \pi^2 \int_0^{2b_r} |\phi_x(t)| dt =$$

$$= 2\pi^2 b_r w_x(2b_r) \leqq 10\pi^2 b_r^2 \int_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt,$$

by inequality (5).

Let us pass to the I_3-integral. We have as follows:

$$|I_3| = \left| \int_{a_r}^{b_r} \phi_x(t+a_r) e^{-2r \sin^2 \frac{1}{2}(t+a_r)} \left(\frac{1}{\sin \frac{1}{2}t} - \frac{1}{\sin \frac{1}{2}(t+a_r)} \right) \sin\left(r + \frac{1}{2}\right) t \, dt \right| =$$

$$= \left| \int_{a_r}^{b_r} \phi_x(t+a_r) e^{-2r \sin^2 \frac{1}{2}(t+a_r)} \frac{\sin \frac{1}{2}(t+a_r) - \sin \frac{1}{2}t}{\sin \frac{1}{2}t \sin \frac{1}{2}(t+a_r)} \sin\left(r + \frac{1}{2}\right) t \, dt \right| =$$

$$= \left| \int_{a_r}^{b_r} \phi_x(t+a_r) e^{-2r \sin^2 \frac{1}{2}(t+a_r)} \frac{2 \cos \frac{1}{4}(t+a_r+t) \sin \frac{1}{4}(t+a_r-t)}{\sin \frac{1}{2}t \sin \frac{1}{2}(t+a_r)} \times \right.$$

$$\left. \times \sin\left(r + \frac{1}{2}\right) t \, dt \right| \leqq$$

$$\leqq \int_{a_r}^{b_r} |\phi_x(t+a_r)| \frac{2 \cdot \frac{1}{4} a_r}{\frac{1}{\pi} t \frac{1}{\pi}(t+a_r)} dt = \frac{\pi^2}{2} \int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)|}{t(t+a_r)} dt =$$

$$= \frac{\pi^2}{2} \int_{a_r}^{b_r} \left(\frac{|\phi_x(t+a_r)|}{t} - \frac{|\phi_x(t+a_r)|}{t+a_r} \right) dt =$$

$$= \frac{\pi^2}{2} \left(\int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)|}{t} dt - \int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)|}{t+a_r} dt \right) =$$

$$= \frac{\pi^2}{2} \left(\int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)|}{t} dt - \int_{2a_r}^{b_r+a_r} \frac{|\phi_x(t)|}{t} dt \right) =$$

$$= \frac{\pi^2}{2} \left(\int_{a_r}^{2a_r} \frac{|\phi_x(t+a_r)|}{t} dt + \int_{2a_r}^{b_r} \left(\frac{|\phi_x(t+a_r)|}{t} - \frac{|\phi_x(t)|}{t} \right) dt - \int_{b_r}^{b_r+a_r} \frac{|\phi_x(t)|}{t} dt \right) =$$

$$= \frac{\pi^2}{2} \left(\int_{a_r}^{2a_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt + \int_{a_r}^{2a_r} \frac{|\phi_x(t)|}{t} dt + \right.$$

$$\left. + \int_{2a_r}^{b_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt - \int_{b_r}^{b_r+a_r} \frac{|\phi_x(t)|}{t} dt \right) \leqq$$

$$\leqq \frac{\pi^2}{2} \left(\int_{a_r}^{2a_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt + \frac{1}{a_r} \int_{a_r}^{2a_r} |\phi_x(t)| dt + \right.$$

$$\left. + \int_{2a_r}^{b_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt \right) \leqq$$

$$\leqq \frac{\pi^2}{2} \left(\int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt + 2 \cdot \frac{1}{2a_r} \int_{0}^{2a_r} |\phi_x(t)| dt \right) =$$

$$= \frac{\pi^2}{2} \int_{a_r}^{b_r} \frac{|\phi_x(t+a_r)| - |\phi_x(t)|}{t} dt + \pi^2 w_x(2a_r) \leqq$$

$$\leqq \frac{\pi^2}{2} \int_{a_r}^{b_r} \frac{|\phi_x(t+a_r) - \phi_x(t)|}{t} dt + \pi^2 w_x(2a_r).$$

Now, taking into account the relation (6) we arrive at

$$|I_3| \leqq M\left(w_x(2a_r) + \overline{w}_{x,a_r}(b_r) + \int\limits_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t} dt \right).$$

Consequently, to estimate the next term, we note that

$$|I_4| = \left| - \int\limits_0^{a_r} \frac{\phi_x(t+a_r)}{\sin\frac{1}{2}(t+a_r)} e^{-2r\sin^2\frac{1}{2}(t+a_r)} \sin\left(r+\frac{1}{2}\right)t\,dt \right| =$$

$$= \left| \int\limits_{a_r}^{2a_r} \frac{\phi_x(t)}{\sin\frac{1}{2}t} e^{-2r\sin^2\frac{1}{2}t} \sin\left(r+\frac{1}{2}\right)t\,dt \right| \leqq \int\limits_{a_r}^{2a_r} \frac{|\phi_x(t)|}{\frac{1}{\pi}t}\left(r+\frac{1}{2}\right)t\,dt =$$

$$= \pi^2\frac{r+\frac{1}{2}}{\pi} \int\limits_{a_r}^{2a_r} |\phi_x(t)|dt \leqq 2\pi^2\frac{1}{2a_r} \int\limits_0^{2a_r} |\phi_x(t)|dt = 2\pi^2 w_x(2a_r).$$

For I_5 we have

$$|I_5| = \left| \int\limits_{b_r-a_r}^{b_r} \frac{\phi_x(t+a_r)}{\sin\frac{1}{2}(t+a_r)} e^{-2r\sin^2\frac{1}{2}(t+a_r)} \sin\left(r+\frac{1}{2}\right)t\,dt \right| =$$

$$= \left| - \int\limits_{b_r}^{b_r+a_r} \frac{\phi_x(t)}{\sin\frac{1}{2}t} e^{-2r\sin^2\frac{1}{2}t} \sin\left(r+\frac{1}{2}\right)t\,dt \right| \leqq$$

$$\leqq \int\limits_{b_r}^{b_r+a_r} \frac{|\phi_x(t)|}{\frac{1}{\pi}t} e^{-2r\sin^2\frac{1}{2}t}dt \leqq \pi \int\limits_{b_r}^{\pi} \frac{|\phi_x(t)|}{t} e^{-2r\sin^2\frac{1}{2}t}\,dt.$$

Now, following the same procedure as for ϱ_3 we achieve the estimate

$$|I_5| \leqq M b_5^2 \int\limits_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt.$$

Collecting the estimations for I_i we may write

$$|\varrho_2| \leqq M\left(w_x(2a_r) + b_r^2 \int\limits_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt + \overline{w}_{x,a_r}(b_r) + \right.$$

$$+\int_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t}dt + \frac{1}{a_r}\int_{a_r}^{b_r} |\phi_x(t)|t^2 dt\Bigg).$$

In this way we arrive at the conclusion that for $r \geqq 12$ our theorem holds.

It remains now to consider the case when $\frac{1}{2} \leqq r < 12$. In this case we can estimate the kernel. Namely,

$$|K_r(t)| = \left| e^{-2r\sin^2\frac{1}{2}t} \frac{\sin(r\sin t + \frac{1}{2}t)}{\sin\frac{1}{2}t} \right| \leqq \frac{|r\sin t + \frac{1}{2}t|}{\frac{1}{\pi}t} \leqq \frac{\pi r t + \frac{1}{2}\pi t}{t} \leqq 13\pi,$$

for $0 < 1 \leqq \pi$ and $\frac{1}{2} \leqq r < 12$.

Since we have $\frac{\pi}{5} < b_r \leqq \pi$, hence, by Lemma 1.4, we find that

$$|B_r[f](x) - f(x)| \leqq \frac{1}{\pi}\int_0^\pi |\phi_x(t)||K_r(t)|dt \leqq 13\pi\frac{1}{\pi}\int_0^\pi |\phi_x(t)|dt =$$

$$= 13\pi w_x(\pi) \leqq 13\pi \frac{2\pi}{\ln\sqrt{2}}\int_\pi^{2\pi} \frac{w_x(t)}{t^2}dt \leqq \frac{26\pi^2}{\ln\sqrt{2}}\int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt =$$

$$= \frac{26\pi^2}{\ln\sqrt{2}}\frac{1}{b_r^2}\cdot b_r^2 \int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt \leqq \frac{26\cdot 5^2}{\ln\sqrt{2}} b_r^2 \int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt =$$

$$= M b_r^2 \int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt.$$

In this way the proof of Theorem 1 is complete. □

Proof of Theorem 2. This proof is similar to that of Theorem 1. It also splits into two main parts. First we assume that $n \geqq 12$. Let us write

$$|E_n[f](x) - f(x)| = \left|\frac{1}{\pi}\left(\int_0^{a_n} + \int_{a_n}^{b_n} + \int_{b_n}^{\pi}\right)\phi_x(t)K_n(t)dt\right| \leqq$$

$$\leqq |\phi_1| + |\phi_2| + |\phi_3|, \quad \text{say},$$

where, in this case we have, $a_n = \frac{2\pi}{n+1}$ and $b_n = \frac{2\pi}{(n+1)^\beta}$, with $\beta \in (\frac{1}{3}, \frac{1}{2})$.

For ϱ_1 we obtain

$$|\varrho_1| = \left|\frac{1}{\pi}\int\limits_0^{a_n} \phi_x(t)\cos^n\Big(\frac{1}{2}t\Big)\frac{\sin\frac{1}{2}(n+1)t}{\sin\frac{1}{2}t}dt\right| \leqq \frac{1}{\pi}\cdot\pi\int\limits_0^{a_n}|\phi_x(t)|\frac{\frac{1}{2}(n+1)t}{t}dt =$$

$$= \frac{\pi}{a_n}\int\limits_0^{a_n}|\phi_x(t)|dt = \pi w_x(a_n) \leqq 2\pi w_x(2a_n) = Mw_x(2a_n),$$

by inequality (3).

Let us pass now to the integral ϱ_3 and write

$$\pi|\varrho_3| = \left|\int\limits_{b_n}^{\pi}\phi_x(t)\cos^n\frac{1}{2}t\frac{\sin\frac{1}{2}(n+1)t}{\sin\frac{1}{2}t}dt\right| \leqq \pi\int\limits_{b_n}^{\pi}\frac{|\phi_x(t)|}{t}\cos^n\frac{1}{2}tdt.$$

For $n \geqq 7$ and $b_n \leqq t \leqq \pi$ we have

$$\cos^n\frac{1}{2}t \leqq \Big(1 - \frac{1}{16}t^2\Big)^n \leqq \Big(1 - \frac{1}{16}b_n^2\Big)^n = \Big(1 - \frac{\pi^2}{4}\cdot\frac{1}{(n+1)^{2\beta}}\Big)^{(n+1)-1} \leqq$$

$$\leqq \frac{8}{3}\left[\Big(1 - \frac{\pi^2}{4}\cdot\frac{1}{(n+1)^{\beta}}\Big)^{\frac{4}{\pi^2}(n+1)^{2\beta}}\right]^{\frac{\pi^2}{4}(n+1)^{1-2\beta}} \leqq \frac{8}{3}\exp\Big(-\frac{\pi^2}{4}(n+1)^{1-2\beta}\Big),$$

since the expression between square brackets increases and tends to e^{-1} as $n \to +\infty$. Hence, recalling the similar fact as in the proof of the preceding theorem, we conclude that there exists a constant $M > 0$ such that for $n \geqq 12$ and $b_n \leqq t \leqq \pi$

$$\cos^n\frac{1}{2}t \leqq M\frac{4\pi^2}{(n+1)^{2\beta}}.$$

We can now proceed analogously as before and conclude that

$$|\varrho_3| \leqq Mb_n^2\int\limits_{b_n}^{\pi}\frac{|\phi_x(t)|}{t}dt \leqq Mb_n^2\left(w_x(\pi) + \int\limits_{b_n}^{\pi}\frac{w_x(t)}{t}dt\right) \leqq$$

$$\leqq Mb_n^2\left(\frac{1}{\ln\sqrt{2}}\int\limits_{\pi}^{2\pi}\frac{w_x(t)}{t}dt + \frac{1}{\ln\sqrt{2}}\int\limits_{b_n}^{\pi}\frac{w_x(t)}{t}dt\right) \leqq Mb_n^2\int\limits_{b_n}^{2\pi}\frac{w_x(t)}{t^2}dt.$$

The term ϱ_2 can be handled in the same way as in the proof of Theorem 1. Similarly, let us write

$$2\pi\varrho_2 = 2\int\limits_{a_n}^{b_n}\phi_x(t)\cos^n\frac{1}{2}t\frac{\sin\frac{1}{2}(n+1)t}{\sin\frac{1}{2}t}dt =$$

$$= \int_{a_n}^{b_n} \phi_x(t) \cos^n \frac{1}{2}t \frac{\sin \frac{1}{2}(n+1)t}{\sin \frac{1}{2}t} dt - \int_0^{b_n - a_n} \phi_x(t+a_n) \cos^n \frac{1}{2}(t+a_n) \frac{\sin \frac{1}{2}(n+1)t}{\sin \frac{1}{2}(t+a_n)} dt =$$

$$= \int_{a_n}^{b_n} \frac{\phi_x(t) - \phi_x(t+a_n)}{\sin \frac{1}{2}t} \cos^n \frac{1}{2}t \sin \frac{1}{2}(n+1)t\, dt +$$

$$+ \int_{a_n}^{b_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}t} \left(\cos^n \frac{1}{2}t - \cos^n \frac{1}{2}(t+a_n)\right) \sin \frac{1}{2}(n+1)t\, dt +$$

$$+ \int_{a_n}^{b_n} \phi_x(t+a_n) \cos^n \frac{1}{2}(t+a_n) \left(\frac{1}{\sin \frac{1}{2}t} - \frac{1}{\sin \frac{1}{2}(t+a_n)} \right) \sin \frac{1}{2}(n+1)t\, dt -$$

$$- \int_0^{a_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}(t+a_n)} \cos^n \frac{1}{2}(t+a_n) \sin \frac{1}{2}(n+1)t\, dt +$$

$$+ \int_{b_n - a_n}^{b_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}(t+a_n)} \cos^n \frac{1}{2}(t+a_n) \sin \frac{1}{2}(n+1)t\, dt =$$

$$= I_1 + I_2 + I_3 + I_4 + I_5, \quad \text{say.}$$

To estimate the integral I_1, we note that

$$|I_1| = \left| \int_{a_n}^{b_n} \frac{\phi_x(t) - \phi_x(t+a_n)}{\sin \frac{1}{2}t} \cos^n \frac{1}{2}t \sin \frac{1}{2}(n+1)t\, dt \right| \leqq$$

$$\leqq \pi \int_{a_n}^{b_n} \frac{|\phi_x(t) - \phi_x(t+a_n)|}{t} dt \leqq$$

$$\leqq \pi \overline{w}_{x,a_n}(b_n) + \pi \int_{a_n}^{b_n} \frac{\overline{w}_{x,a_n}(t)}{t} dt,$$

by relation (8).

Now, let us consider the integral

$$|I_2| = \left| \int_{a_n}^{b_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}t} \left(\cos^n \frac{1}{2}t - \cos^n \frac{1}{2}(t+a_n)\right) \sin \frac{1}{2}(n+1)t\, dt \right|.$$

Here again, we apply the mean-value theorem which assures that there exists a $\xi_t \in (t, t+a_r)$ such that

$$\cos^n \frac{1}{2}t - \cos^n \frac{1}{2}(t+a_n) = \frac{1}{2} n a_n \cos^{n-1} \frac{1}{2}\xi_t \sin \frac{1}{2}\xi_t.$$

Here $a_r < t < \xi_t < t + a_r < 2t$, so we get, by (8),

$$|I_2| \leqq \int\limits_{a_n}^{b_n} \frac{|\phi_x(t+a_n)|}{\frac{1}{\pi}t} \frac{1}{2}(n+1)a_r \frac{1}{2} \cdot 2t\,dt \leqq \pi^2 \int\limits_{a_n}^{b_n} |\phi_x(t+a_n)|dt =$$

$$= \pi^2 \int\limits_{2a_n}^{b_n+a_n} |\phi_x(t)|dt \leqq \pi^2 \int\limits_{0}^{2b_n} |\phi_x(t)|dt =$$

$$= \pi^2 2b_n w_x(2b_n) \leqq 2 \cdot 20\pi^2 b_n^2 \int\limits_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt.$$

Let us pass to the integral I_3. Similarly, as earlier, we conclude that

$$|I_3| = \left| \int\limits_{a_n}^{b_n} \phi_x(t+a_n) \cos^n \frac{1}{2}(t+a_n) \left(\frac{1}{\sin \frac{1}{2}t} - \frac{1}{\sin \frac{1}{2}(t+a_n)} \right) \sin \frac{1}{2}(n+1)t\,dt \right| \leqq$$

$$\leqq \frac{\pi^2}{2} \int\limits_{a_n}^{b_n} a_n \frac{|\phi_x(t+a_n)|}{t(t+a_n)} dt \leqq \frac{\pi^2}{2} \left(\overline{w}_{x,a_n}(b_n) + \int\limits_{a_n}^{b_n} \frac{\overline{w}_{x,a_n}(t)}{t} dt \right) + \pi^2 w_x(2a_n).$$

For the integral I_4 we can get

$$|I_4| = \left| - \int\limits_{0}^{a_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}(t+a_n)} \cos^n \frac{1}{2}(t+a_n) \sin \frac{1}{2}(n+1)t\,dt \right| \leqq \frac{\pi^2}{2} w_x(2a_n).$$

As in the preceding proof, I_5 can be estimated similarly as it was done for ϱ_3

$$|I_5| = \left| \int\limits_{b_n-a_n}^{b_n} \frac{\phi_x(t+a_n)}{\sin \frac{1}{2}(t+a_n)} \cos^n \frac{1}{2}(t+a_n) \sin \frac{1}{2}(n+1)t\,dt \right| =$$

$$= \left| - \int\limits_{b_n}^{b_n+a_n} \frac{\phi_x(t)}{\sin \frac{1}{2}(t)} \cos^n \frac{1}{2}t \sin \frac{1}{2}(n+1)t\,dt \right| \leqq$$

$$\leqq \pi \int_{b_n}^{\pi} \frac{|\phi_x(t)|}{t} \cos^n \frac{1}{2} t\, dt \leqq M b_n^2 \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt.$$

Collecting the above estimations we obtain, for $n = 12, 13, \dots$

$$|E_n[f](x) - f(x)| \leqq$$

$$\leqq M \left(w_x(2a_n) + b_n^2 \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt + \overline{w}_{x,a_n}(b_n) + \int_{a_n}^{b_n} \frac{\overline{w}_{x,a_n}(t)}{t} dt \right).$$

In the case when $0 \leqq n < 12$, we can also estimate the kernel:

$$|K_n(t)| = \left| \cos^n \left(\frac{1}{2} t \right) \frac{\sin \frac{1}{2}(n+1)t}{\sin \frac{1}{2} t} \right| \leqq \frac{\frac{1}{2}(n+1)t}{\frac{1}{\pi} t} \leqq 7\pi, \quad \text{for} \quad 0 < t \leqq \pi.$$

Using the relations $\frac{1}{2}\pi < b_n < 2^{\frac{2}{3}}\pi$, for $1 \leqq n < 12$, and $w_x(\pi) \leqq 2w_x(t)$, for $\pi \leqq t \leqq 2\pi$ (inequality (9)) we obtain

$$|E_n[f](x) - f(x)| \leqq \frac{1}{\pi} \int_0^{\pi} |\phi_x(t)||K_n(t)| dt \leqq 7\pi \frac{1}{\pi} \int_0^{\pi} |\phi_x(t)| dt = 7\pi w_x(\pi) \leqq$$

$$\leqq 7\pi \frac{1}{\ln 2^{\frac{1}{3}}} \int_{2^{\frac{2}{3}}\pi}^{2\pi} w_x(\pi) \frac{dt}{t} \leqq \frac{14\pi}{\ln 2^{\frac{1}{3}}} \int_{2^{\frac{2}{3}}\pi}^{2\pi} \frac{w_x(t)}{t} dt \leqq$$

$$\leqq \frac{14\pi \cdot 2\pi}{\ln 2^{\frac{1}{3}}} \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt =$$

$$= \frac{2 \cdot 14\pi^2}{\ln 2^{\frac{1}{3}}} \frac{1}{b_n^2} \cdot b_n^2 \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt \leqq$$

$$\leqq \frac{3 \cdot 28\pi^2}{\ln 2} \cdot \frac{2}{\pi} b_n^2 \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt = M b_n^2 \int_{b_n}^{2\pi} \frac{w_x(t)}{t^2} dt.$$

For $n = 0$, $a_n = b_n = 2\pi$, and so

$$|E_n[f](x) - f(x)| \leqq 7\pi w_x(\pi) \leqq 28\pi w_x(4\pi) = M w_x(2a_n).$$

In this way the proof is complete. □

PROOF OF THEOREM 1′. We have $\beta \in (\frac{1}{3}, \frac{1}{2})$ therefore by Theorem 1 there is an $M > 0$ such that

$$\|B_r[f] - f\|_p \leqq$$
$$\leqq M \left\| w_x(2a_r) + b_r^2 \int\limits_{b_r}^{2\pi} \frac{w_x(t)}{t^2} dt + \overline{w}_{x,a_r}(b_r) + \int\limits_{a_r}^{b_r} \frac{\overline{w}_{x,a_r}(t)}{t} dt + \frac{1}{a_r} \int\limits_{a_r}^{b_r} |\phi_x(t)| t^2 dt \right\|_p .$$

Applying the generalized Minkowski inequality, when $1 \leqq p < \infty$, we can estimate the quantity

$$\|w_\cdot(\delta)\|_p = \|w_x(\delta)\|_p = \left(\int\limits_{-\pi}^{\pi} (w_x(\delta))^p dx \right)^{1/p} = \left(\int\limits_{-\pi}^{\pi} \frac{1}{\delta^p} \left(\int\limits_0^\delta |\phi_x(t)| dt \right)^p dx \right)^{1/p} =$$
$$= \frac{1}{\delta} \left(\int\limits_{-\pi}^{\pi} \left(\int\limits_0^\delta |\phi_x(t)| dt \right)^p dx \right)^{1/p} \leqq \frac{1}{\delta} \left(\int\limits_0^\delta \left(\int\limits_{-\pi}^{\pi} |\phi_x(t)|^p dx \right)^{1/p} dt \right) \leqq$$
$$\leqq \frac{1}{\delta} \left(\int\limits_0^\delta \sup_{|t| \leqq \delta} \left(\int\limits_{-\pi}^{\pi} |\phi_x(t)|^p dx \right)^{1/p} dt \right) =$$
$$= \frac{1}{\delta} \left(\int\limits_0^\delta \omega(\delta)_p dt \right) = \omega[f](\delta)_p \quad \text{for any } \delta > 0.$$

When $p = \infty$ we have

$$\|w_\cdot(\delta)\|_\infty = \left\| \frac{1}{\delta} \int\limits_0^\delta |\phi_x(t)| dt \right\|_\infty \leqq \frac{1}{\delta} \int\limits_0^\delta \|\phi_\cdot(t)\|_\infty dt \leqq$$
$$\leqq \frac{1}{\delta} \int\limits_0^\delta \omega(\delta)_\infty dt = \omega(\delta)_\infty, \quad \text{again } \delta > 0.$$

Using (6) and (4) we have

$$\left\| \overline{w}_{\cdot,a_r}(b_r) + \int\limits_{a_r}^{b_r} \frac{\overline{w}_{\cdot,a_r}(t)}{t} dt \right\|_p = \left\| \overline{w}_{\cdot,a_r}(a_r) + \int\limits_{a_r}^{b_r} \frac{|\phi_x(t) - \phi_x(t + a_r)|}{t} dt \right\|_p \leqq$$
$$\leqq \|\overline{w}_{\cdot,a_r}(a_r)\|_p + \left\| \int\limits_{a_r}^{b_r} \frac{|\phi_\cdot(t) - \phi_\cdot(t + a_r)|}{t} dt \right\|_p \leqq$$

$$\leqq 4\|w_{\cdot}(2a_r)\|_p + \left\| \int_{a_r}^{b_r} \frac{|\phi_{\cdot}(t) - \phi_{\cdot}(t+a_r)|}{t} dt \right\|_p .$$

This completes the proof. □

In the same way, applying in turn Theorem 2 we can prove Theorem 2′.

Proof of Corollary 1. Put $\beta = \frac{\alpha+1}{\alpha+3}$. We have then $\beta \in (\frac{1}{3}, \frac{1}{2})$ for $\alpha \in (0,1)$. By the assumption $f \in \operatorname{Lip}(\alpha, p)$, so $\omega[f](\delta)_p \leqq N\delta^\alpha$ for some $N > 0$ and any $\delta > 0$. Hence

$$\omega(2a_r)_p \leqq N(2a_r)^\alpha = N\frac{1}{(r+\frac{1}{2})^\alpha} \leqq Nr^{-\alpha},$$

$$b_r^2 \int_{b_r}^{2\pi} \frac{\omega(t)_p}{t^2} dt \leqq b_r^2 \int_{b_r}^{2\pi} \frac{Nt^\alpha}{t^2} dt = Nb_r^2 \int_{b_r}^{2\pi} t^{\alpha-2} dt \leqq$$

$$\leqq Nb_r^2 \frac{1}{1-\alpha} b_r^{\alpha-1} = Nb_r^{\alpha+1} =$$

$$= N\frac{1}{(r+\frac{1}{2})^{\beta(\alpha+1)}} \leqq N\frac{1}{r^\alpha},$$

$$\text{since} \quad \alpha \leqq \frac{(\alpha+1)^2}{\alpha+3} \quad \text{for} \quad 0 < \alpha < 1.$$

Further we obtain for $r \geqq \frac{1}{2}$

$$\left\| \int_{a_r}^{b_r} \frac{|\phi_{\cdot}(t) - \phi_{\cdot}(t+a_r)|}{t} dt \right\|_p \leqq \left\| \int_{a_r}^{\pi} \frac{|\phi_{\cdot}(t) - \phi_{\cdot}(t+a_r)|}{t} dt \right\|_p \leqq Nr^{-\alpha}.$$

For the last component we have

$$\frac{1}{a_r} \int_{a_r}^{b_r} \|\phi_{\cdot}(t)\|_p t^2 dt \leqq \frac{1}{a_r} \int_{a_r}^{b_r} \omega(t)_p t^2 dt \leqq \frac{1}{a_r} \int_{a_r}^{b_r} Nt^{\alpha+2} dt \leqq$$

$$\leqq \frac{N}{a_r} b_r^{\alpha+3} = N\left(r + \frac{1}{2}\right)^{1-(\alpha+3)\beta} \leqq Nr^{-\alpha}$$

since $1 - (\alpha+3)\beta = -\alpha$ (recall that $\beta = \frac{\alpha+1}{\alpha+3}$). Thus the proof is complete. □

Proof of Remark 2. Initially, for $r \geqq 1$,

$$a_r^2 \int_{2a_r}^{2\pi} \frac{dt}{t^3} = a_r^2 \left(-\frac{1}{2t^2}\right)\Big|_{2a_r}^{2\pi} = \frac{1}{8} - \frac{1}{8(r+\frac{1}{2})} > \frac{1}{8} - \frac{1}{8 \cdot \frac{3}{2}} = \frac{1}{24}.$$

Further, using the previously stated lemmas we can obtain, successively,

$$\frac{1}{a_r}\int_{a_r}^{b_r}|\phi_x(t)|t^2dt \leqq \frac{b_r^2}{a_r}\int_{a_r}^{b_r}|\phi_x(t)|dt \leqq \frac{b_r^3}{a_r}w_x(b_r) =$$

$$= \pi^2\Big(r+\frac{1}{2}\Big)^{1-3\beta} w_x(b_r) \leqq \pi^2 w_x(b_r) \leqq 2\pi^2 w_x(2b_r) \leqq$$

$$\leqq 2\cdot 5\pi^2 b_r \int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt \leqq 10\pi^2\int_{b_r}^{2\pi}\frac{w_x(t)}{t}dt \leqq 20\pi^2\int_{b_r}^{2\pi}\frac{w_x(2t)}{t}dt,$$

$$w_x(2a_r) \leqq 12a_r\int_{a_r}^{2\pi}\frac{w_x(t)}{t^2}dt \leqq 12\int_{a_r}^{2\pi}\frac{w_x(t)}{t}dt \leqq 24\int_{a_r}^{2\pi}\frac{w_x(2t)}{t}dt,$$

$$b_r^2\int_{b_r}^{2\pi}\frac{w_x(t)}{t^2}dt \leqq b_r\int_{b_r}^{2\pi}\frac{w_x(t)}{t}dt \leqq 2\pi\int_{b_r}^{2\pi}\frac{w_x(2t)}{t}dt,$$

$$\overline{w}_{x,a_r}(b_r) \leqq 4w_x(2b_t) \leqq 4\cdot 5\cdot 2\int_{b_r}^{2\pi}\frac{w_x(2t)}{t}dt,$$

$$\int_{a_r}^{b_r}\frac{\overline{w}_{x,a_r}(t)}{t}dt \leqq 4\int_{a_r}^{b_r}\frac{w_x(2t)}{t}dt.$$

Thus

$$|B_r[f](x)-f(x)| \leqq M\int_{a_r}^{2\pi}\frac{w_x(2t)}{t}dt = M\int_1^{2(r+\frac{1}{2})}\frac{w_x(\frac{4\pi}{t})}{t}dt,$$

and

$$\|B_r[f]-f\| \leqq M\int_1^{2(r+\frac{1}{2})}\frac{\|w_{\cdot}(\frac{4\pi}{t})\|}{t}dt \leqq M\int_1^{2(r+\frac{1}{2})}\frac{\omega(4\pi\cdot\frac{1}{t})}{t}dt \leqq$$

$$\leqq M(4\pi+1)\int_1^{2(r+\frac{1}{2})}\frac{\omega_x(\frac{1}{t})}{t}dt,$$

by the relation $\omega(\lambda\delta) \leqq (\lambda+1)\omega(\delta)$, valid for any $\lambda \geqq 0$ and $\delta \geqq 0$.

Similarly we can obtain that

$$\|E_n[f]-f\| \leqq M \int_1^{n+1} \frac{\omega_x\left(\frac{1}{t}\right)}{t} dt \qquad \text{for } n=1,2,\ldots,$$

and since $\omega(t)$ is non-decreasing function we conclude that

$$\|E_n[f]-f\| \leqq M \int_1^{n+1} \frac{\omega_x\left(\frac{1}{t}\right)}{t} dt = M \sum_{k=1}^{n} \int_k^{k+1} \frac{\omega_x\left(\frac{1}{t}\right)}{t} dt \leqq$$

$$\leqq M \sum_{k=1}^{n} \int_k^{k+1} \frac{\omega_x\left(\frac{1}{k}\right)}{k} dt = M \sum_{k=1}^{n} \frac{\omega_x\left(\frac{1}{k}\right)}{k}.$$

□

ACKNOWLEDGEMENT. The author wishes to express his thanks to Professor Włodzimierz Łenski for suggesting the problem and for many stimulating conversations.

REFERENCES

[1] HARDY, G. H., *Divergent series*, Clarendon Press, Oxford, 1949. *MR* **11**, 25a

[2] CHUI, C. K. and HOLLAND, A. S. B., On the order of approximation by Euler and Taylor means, *J. Approx. Theory* **39** (1983), 24–38. *MR* **84i**:41026

[3] HOLLAND, A. S. B. and SAHNEY, B. N., On degree of approximation by Euler (E,q)-means, *Studia Sci. Math. Hungar.* **11** (1976), 431–435. *MR* **81h**:42004

(Received July 19, 1994)

INSTITUTE OF MATHEMATICS
T. KOTARBIŃSKI PEDAGOGICAL UNIVERSITY
PLAC SŁOWIAŃSKI 9
PL–65-069 ZIELONA GÓRA
POLAND

roszak@omega.im.wsp.zgora.pl

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 393–418

DEVIATION PRINCIPLE FOR SET INDEXED PROCESSES WITH INDEPENDENT INCREMENTS

P. BARBE and M. BRONIATOWSKI

Abstract

Let $\{X_\lambda(A) : A \in \mathcal{A}\}_{\lambda>0}$ be a family of processes indexed by a collection of sets $\mathcal{A}$. Assuming that these processes have independent increments, and that for any fixed $A \in \mathcal{A}, \{X_\lambda(A)\}_{\lambda>0}$ obey a large deviation principle (LDP) as $\lambda \to \infty$, we prove that the processes $\{A \in \mathcal{A} \to X_\lambda(A)\}_{\lambda>0}$ obey a functional LDP. As an application, we obtain a LDP for partial extreme processes based on i.i.d. r.v.'s and the sequential empirical measure. Then, we show how to apply this result to obtain a LDP for weighted V-statistics.

1. Introduction

Consider a family of stochastic processes $A \in \mathcal{A} \to X_\lambda(A) \in \mathcal{X}, \lambda \in (0,\infty)$ indexed by some sets $A \in \mathcal{A}$. Suppose that there exist two functions

$$(\lambda, A) \in (0,\infty) \times \mathcal{A} \mapsto h(\lambda, A) \in [0,\infty]$$

and

$$(\Omega \subset \mathcal{X}, A \in \mathcal{A}) \mapsto I(\Omega, A) \in [0,\infty]$$

such that for any fixed $A \in \mathcal{A}$,

$$\mathbf{P}\{X_\lambda(A) \in \Omega\} \cong \exp(-h(\lambda, A) I(\Omega, A)) \quad \text{as} \quad \lambda \to \infty \tag{1.1}$$

(see Section 2 for a more precise statement). If (1.1) holds, we say that the family of set indexed random processes $\{X_\lambda(.)\}_{\lambda>0}$ obeys a marginal large deviation principle. Of course, (1.1) concerns only the marginal behaviour of the processes $X_\lambda(.)$, i.e. for any fixed $A \in \mathcal{A}$. Now, a natural question arises: if (1.1) holds, what can we say about the processes $X_\lambda(.)$ as $\lambda \to \infty$?

These processes may be viewed as random functions in some space $\mathcal{F}$ of functions from $\mathcal{A}$ into $\mathcal{X}$. Hence, one can ask whether there exist functions

$$\lambda \mapsto h(\lambda) \in [0,\infty]$$

and

$$\Omega \in \mathcal{F} \mapsto J(\Omega) \in [0,\infty]$$

1991 *Mathematics Subject Classification.* Primary 60F10.

Key words and phrases. Moderate, large, super-large deviation principle, processes with independent increments, sequential empirical process and measure, partial extreme process, weighted V-statistics.

0081–6906/97/\$ 5.00

such that

$$\mathbf{P}\{X_\lambda(.)\in\Omega\}\cong\exp(-h(\lambda)J(\Omega)) \quad \text{as} \quad \lambda\to\infty. \tag{1.2}$$

A special case of this problem was solved by Varadhan [27], McBride [23], Lynch and Sethuraman [22] and Broniatowski and Mason [12] when $\lambda\in\mathbb{N}$, $\mathcal{A}=\{(0,t]:t>0\}$ and $X_\lambda(A)=\Sigma_{i\in\lambda A}X_i$ is a sum of independent and identically distributed (i.i.d.) random variables (r.v.'s). It turns out that the approach of McBride [23] or Lynch and Sethuraman [22] may be substantially extended to obtain (1.2), assuming that (1.1) holds and that for every λ the process X_λ has independent increments (see Section 2 for the precise assumptions).

The main result is given in Section 2. In Section 3 we deduce from our main theorem a functional large deviation result for partial extreme processes. In Section 4 we give a large deviation principle for a sequential empirical probability measure which extends previous results by Groeneboom, Oosterhoff and Ruymgaart [19] and Bahadur and Zabell [2] among others. This result is of very broad applicability and some applications will be presented in a series of forthcoming papers. All the proofs of our results are postponed to the last section.

2. The main result

Before stating our main theorem, we first need to precise the meaning of (1.1) and to make some assumptions.

Consider a class of sets $\mathcal{A}$, and a Hausdorff topological space $\mathcal{X}$. We denote by $\mathcal{F}$ the set of all $\mathcal{X}$-valued functions defined on $\mathcal{A}$. We shall consider a family $X_\lambda(.)$ of random elements in $\mathcal{F}$ with independent increments in the sense that

(2.1) If $A,B\in\mathcal{A}$ and $A\cap B=\emptyset$, then $X_\lambda(A)$ and $X_\lambda(B)$ are independent.

One can easily exhibit such a process as follows: first define an associative operation T on $\mathcal{X}$. Consider an ordered discrete set $\mathcal{D}$ and let $\mathcal{A}$ be the set of all subsets of $\mathcal{D}$. Then let $\{X_d:d\in\mathcal{D}\}$ be a set of i.i.d. r.v.'s and define for any $A\in\mathcal{A}$,

$$X(A):=\underset{d\in A}{T}X_d.$$

To make (1.1) and (1.2) more precise, we introduce the following definition.

Definition 2.1. A function $I(.)$ from a topological space Θ in $[0,\infty]$ is a good rate function if

$$0\leqq I(.)\leqq\infty \tag{2.2}$$

(2.3) $$I(.) \text{ is lower semicontinuous (lsc).}$$

A consequence of (2.3) is that the level sets $\Gamma_c := \{x \in \Theta : I(x) \leqq c\}, c \in [0, \infty]$, are closed in Θ. But we do not assume that they are compact which is of importance for the applications (see Remark 2.4).

DEFINITION 2.2. The processes $X_\lambda(.) \in \mathcal{F}$ satisfy a marginal deviation principle (MDP) if for any $A \in \mathcal{A}$, there exists a good rate function $I(x, A)$ from $\mathcal{X}$ into $[0, \infty]$ and a real valued function $h(\lambda, A)$ such that for any closed set $F \subset \mathcal{X}$,

(2.4) $$\limsup_{\lambda \to \infty} h(\lambda, A)^{-1} \log \mathbf{P}\{X_\lambda(A) \in F\} \leqq -I(F, A)$$

while for any open set $G \subset \mathcal{X}$,

(2.5) $$\liminf_{\lambda \to \infty} h(\lambda, A)^{-1} \log \mathbf{P}\{X_\lambda(A) \in G\} \geqq -I(G, A),$$

where for any $\Omega \subset \mathcal{X}$,

(2.6) $$I(\Omega, A) := \inf_{x \in \Omega} I(x, A).$$

REMARK 2.1. In (2.4) and (2.5), the functions $h(\lambda, A)$ and $I(., A)$ are not uniquely defined. For example, it is obvious that we can divide $h(\lambda, A)$ by some constant $\mu(A)$ and multiply $I(., A)$ by $\mu(A)$ without changing the meaning of (2.4) or (2.5); similarly, we can replace $h(\lambda, A)^{-1}$ by λ^{-1}.

REMARK 2.2. Definition 2.2 is nothing else than what is often called a large deviations principle ([27], [17], [16]). However, the terminology is not well established, since if $X_\lambda(.)$ is a partial sum process, our definition includes small or moderate, large or super-large deviations (see e.g. [4]).

Following Lynch and Sethuraman [22], we introduce the definition:

DEFINITION 2.3. The processes $X_\lambda(.)$ are marginally deviation tight (MDT) if for any $A \in \mathcal{A}$ and any $M > 0$ there exists a compact set $K_{A,M} \subset \mathcal{X}$ such that

(2.7) $$\limsup_{\lambda \to \infty} h(\lambda, A)^{-1} \log \mathbf{P}\{X_\lambda(A) \notin K_{A,M}\} \leqq -M.$$

An important consequence of (2.7) is that provided we deal with a fixed marginal $X_\lambda(A)$ of our processes, we can reduce $\mathcal{X}$ to its compact subsets when we evaluate a probability of deviation.

Definitions 2.2 and 2.3 are concerned with the marginal behaviour of the process in the sense that $A \in \mathcal{A}$ is fixed. Now, to consider the function $A \to X_\lambda(A)$ we assume that

(2.8) $\mathcal{F}$ is equipped with the topology of pointwise convergence or any coarser topology.

Consequently, if a sequence $(g_n)_{n\geqq 1}$ in $\mathcal{F}$ converges to $g\in\mathcal{F}$, then $\lim_{n\to\infty} g_n(A) = g(A)$ for any $A\in\mathcal{A}$.

We can make (1.2) more precise:

DEFINITION 2.4. The process $X_\lambda(.)$ satisfies a functional deviation principle (FDP) if there exists a function $h(\lambda)$ and a good rate function $I(.)\colon \mathcal{F}\mapsto [0,\infty]$ such that for any closed set $F\subset\mathcal{F}$,

$$\limsup_{\lambda\to\infty} h(\lambda)^{-1}\log \mathbf{P}\{X_\lambda\in F\}\leqq -I(F) \tag{2.9}$$

while for any open set $G\subset\mathcal{F}$,

$$\liminf_{\lambda\to\infty} h(\lambda)^{-1}\log \mathbf{P}\{X_\lambda\in G\}\geqq -I(G), \tag{2.10}$$

where for any $\Omega\subset\mathcal{F}$, we denote

$$I(\Omega):=\inf_{x\in\Omega} I(x). \tag{2.11}$$

REMARK 2.3. As in Definition 2.2, the functions $h(.)$ and $I(.)$ are not uniquely defined. It is also clear that if $h(.)$ is monotone, then one can always take $h(\lambda)=\lambda$ by changing the way X_λ is indexed (i.e. replace λ by $h^{-1}(\lambda)$).

Now, the question we asked in Section 1 becomes: if the processes $X_\lambda(.)$ satisfy a MDP, do they obey a FDP?

The answer is yes if we assume (with notation of Definition 2.2) that there exist functions $\mu\colon \mathcal{A}\mapsto[0,\infty]$ and $h\colon [0,\infty]\mapsto[0,\infty]$ such that

$$\lim_{\lambda\to\infty} h(\lambda,A)/h(\lambda)=\mu(A) \tag{2.12}$$

and

(2.13) for any $g\in\mathcal{F}$, the function $A\in\mathcal{A}\to\mu(A)I(g(A),A)\in[0,\infty]$ is subadditive.

By (2.13) we mean that if $A,B,C\in\mathcal{A}$ with $A=B\cup C$ and $B\cap C=\emptyset$ then

$$\mu(A)I(g(A),A)\leqq\mu(B)I(g(B),B)+\mu(C)I(g(C),C).$$

Next, we have to relate the rate function $I(.)$ in (2.9)–(2.10) to the rate function $I(.,A)$ in (2.4)–(2.5). For this, we say that $A=(A_1,\dots,A_k)\in\mathcal{A}^k$ is a partition if all the sets $A_1,\dots,A_k$ are disjoint (but we do not suppose that their union covers $\mathcal{A}$). We denote by $\mathcal{P}$ the set of all partitions. For any $f\in\mathcal{F}$, we denote

$$Jf:=\sup_{A\in\mathcal{P}}\sum_{A_i\in A}\mu(A_i)I(f(A_i),A_i). \tag{2.14}$$

A consequence of (2.13) is that the sum in the r.h.s. of (2.14) as a function of $A \in \mathcal{P}$ increases when we refine the partition A. We shall assume that there is a sequence of partitions $(A_n^*)_{n \geqq 1}$ such that

$$\text{the } \sigma\text{-fields } \sigma(A_n^*) \text{ increase with } n, \text{ and for any } f \in \mathcal{F}, \tag{2.15}$$
$$J(f) = \lim_{n \to \infty} \sum_{A_{i,n}^* \in A_n^*} \mu(A_{i,n}^*) I(f(A_{i,n}^*), A_{i,n}^*).$$

For the upper bounds on closed sets we also need to assume that

$$\lim_{\lambda \to \infty} h(\lambda) = \infty. \tag{2.16}$$

THEOREM 2.1. *Assume that the processes* X_λ *obey a MDP given by* (2.4)–(2.6), *and that* (2.12)–(2.14) *hold. Then*

(i) X_λ *obey a FDP on open sets given by* (2.10) *and* (2.11) *with the function* $h(.)$ *given in* (2.12) *and rate function* $J(.)$ *given in* (2.14).

(ii) *Moreover, if* (2.15), (2.16) *hold and* $X_\lambda(.)$ *is MDT, then* (2.9) *also holds.*

REMARK 2.4. It must be noticed that the pointwise topology on $\mathcal{F}$ is not crucial in our upper bound result. It is only used to show that $J(.)$ is lsc (see our Proposition 5.1.1). Therefore, any topology under which $J(.)$ is lsc is suitable to get Theorem 2.1. In particular, if $X_\lambda(.)$ is a partial sum process in the space $\mathcal{D}[0,1]$, we can use Skorokhod topology and strength Lynch and Sethuraman's result [22] (see their comments p. 621). Also, we do not need $I(., A)$ to be continuous (compare to the proof of Theorem 4.1 in [22]), which is crucial for some applications, as well as not supposing that $\mathcal{X}$ is a metric space (see Remark 5.3.1).

3. Large deviations for partial extreme process

Let $(X_i)_{i \geqq 1}$ be a sequence of i.i.d. real valued r.v.'s with common distribution function F which is assumed to be continuous. Then, let

$$M_n(t) := \max_{1 \leqq i \leqq nt} X_i, \qquad 0 < t \leqq 1, \quad n \geqq 1$$

be the associated partial extreme process. In this section we present a FDP for $M_n(.)$ which is a direct consequence of results of Theorem 2.1.

We view $M_n(.)$ as a random element in the space $\mathcal{D}\uparrow(0,1]$ of all non-decreasing and càdlàg functions from $(0,1]$ to $F^{\leftarrow}(\overline{\mathbb{R}})$, the compactified of supp F equipped with its usual topology. The space $\mathcal{D}\uparrow(0,1]$ is equipped with the topology of pointwise convergence.

THEOREM 3.2. *The family of processes $\{M_n\}_{n\geqq 1}$ obeys a FDP. If Ω is closed in $\mathcal{D}\uparrow(0,1]$, then*

$$\limsup_{n\to\infty} n^{-1}\log \mathbf{P}\{M_n(.)\in\Omega\}\leqq -J(\Omega),$$

and if Ω is open in $\mathcal{D}\uparrow(0,1]$, then

$$\liminf_{n\to\infty} n^{-1}\log \mathbf{P}\{M_n(.)\in\Omega\}\geqq -J(\Omega),$$

where

$$J(\Omega):=\inf_{g\in\Omega} J(g),$$

and

$$J(g):=-\int_0^1 \log F(g(t))\,dt.$$

4. Empirical process based on i.i.d. random field

Let $\mathcal{Z}:=\{Z_i : i\in\mathbb{Z}^{+d}\}$ be a random field of i.i.d. r.v.'s with common probability measure (p.m.) $\mathbf{P}$, each Z_i taking its values in some Hausdorff space S equipped with its Borel σ-field $\mathcal{B}(S)$. Define the sequential empirical measure

$$\mathbb{P}_n:=n^{-d}\sum_{i\in n(0,1]^d}\delta_{(i/n,Z_i)}.$$

The term sequential is introduced to point out that the knowledge of $\mathbb{P}_n$ enables us to reconstruct the whole field $(X_i)_{i\in[0,1]^d}$, while the usual empirical measure only gives the sequence up to a permutation. The result of Section 2 leads to some LDP for $\mathbb{P}_n$.

LDP for the usual empirical p.m. have been obtained among others by Borovkov [9], Donsker and Varadhan [18], Bahadur and Zabell [2], Groeneboom, Oosterhoff and Ruymgaart [19] (GOR in the sequel), Csiszár [13], Deuschel and Stroock [17], de Acosta [1].

To state our theorem, we need some topology on the space $S_d:=(0,1]^d\times S$ which will be equipped with its corresponding Borel σ-field $\mathcal{B}(S_d)$.

It is well known (see e.g. Billingsley [8]) that a measure on S_d is completely determined by its values on the rectangles $A\times B$, $A\in\mathcal{B}(0,1]^d$, $B\in\mathcal{B}(S)$. Let Λ be the set of all p.m.'s on S_d and let $\mathcal{A}$ be the class of all Borel subsets $A\subset(0,1]^d$ such that

$$\lim_{n\to\infty}\#(nA)/n^d=|A|,$$

where $|A|$ denotes the d-dimensional Lebesgue measure of A. In particular, $\mathcal{A}$ contains all the open and closed subsets of $(0,1]^d$. If $(Q_n)_{n\geqq 1}$ is a sequence in Λ, we say that Q_n converges to Q if for any $A\in\mathcal{A}$ and any $B\in\mathcal{B}(S)$,

$$\lim_{n\to\infty} Q_n(A\times B)=Q(A\times B).$$

Hence, we consider on Λ the product topology of convergence on all elements of $\mathcal{A}$ with the τ-topology (see e.g. [19]) on S. This topology is of course much thinner than the weak topology.

If $Q,R\in\Lambda$, we define the Kullback–Leibler information number between Q and R by

$$K(Q,R):=\begin{cases}\int \log(dQ/dR)dQ & \text{if } Q\ll R \text{ and } Q(A\times S)=R(A\times S)\\ & \text{for any } A\in\mathcal{B}(0,1]^d,\\ \infty & \text{otherwise.}\end{cases}$$

THEOREM 4.3. *If* $\mathbf{P}$ *is tight, the sequential empirical p.m.* $\mathbb{P}_n$ *obeys the following LDP. If* Ω *is closed in* Λ, *then*

$$\limsup_{n\to\infty} n^{-d}\log \mathbf{P}\{\mathbb{P}_n\in\Omega\}\leqq -K(\Omega,\mathbb{P}),$$

while if Ω *is open in* Λ,

$$\liminf_{n\to\infty} n^{-d}\log \mathbf{P}\{\mathbb{P}_n\in\Lambda\}\geqq -K(\Omega,\mathbb{P}),$$

where $\mathbb{P}:=\lambda_d\times\mathbf{P}$, *and* λ_d *is the Lebesgue measure over the unit cube* $(0,1]^d$, *and*

$$K(\Omega,\mathbb{P}):=\inf_{Q\in\Omega} K(Q,\mathbb{P}).$$

REMARK 4.5. The assumption that $\mathbf{P}$ is tight has to be taken in the sense that for any $\epsilon>0$, there exists a compact set $C_\epsilon\subset S$ such that $\mathbf{P}(C_\epsilon)>1-\epsilon$ (see [6]). The only use of this assumption is to prove that the usual empirical measure is deviation tight.

After the writing of the technical report of Barbe and Broniatowski [3], we have been aware of the paper of Dembo and Zajic [15], where a large deviation principle is given for the process $t\mapsto n^{-1}\sum_{1\leqq i\leqq nt}\delta_{X_i}$. The reader may notice that the mapping $\mathbb{Q}\in S_1\mapsto(t\mapsto\int \mathbb{I}_{[0,t]}(u)\delta_x d\mathbb{Q}(u,x)$ is continuous (equip the space of measured-valued functions defined on $[0,1]$ with the pointwise convergence and use the τ-topology on the set of nonnegative measures). The result of Dembo and Zajic [15] may be viewed as a consequence of Theorem 4.3 (see also Remark 5.1.1). Our result also shows that it is possible to build up a LDP for a p.m. (the sequential empirical p.m.)

which encompass both the LDP for partial sum process and the LDP for the empirical p.m.

One can easily get a moderate deviation result using the same techniques. More precisely, if $(\alpha_n)_{n\geqq 1}$ denotes a sequence of positive real numbers such that $\lim_{n\to\infty} \alpha_n^2/n=0$, one can introduce the signed measure $\Pi_n := \alpha_n(\mathbb{P}_n - \mathbb{P})$. Next, if μ is a signed measure, we define its chi-square norm by

$$\chi_P^2(\mu) := \begin{cases} \int (d\mu/d\mathbb{P})d\mu & \text{if } \mu \ll \mathbb{P}, \\ \infty & \text{otherwise.} \end{cases}$$

We can extend the topology on Λ to the set of all signed measures $\mathcal{M}$ on $((0,1]^d \times S, \mathcal{B}(0,1]^d \otimes \mathcal{B}(S))$, and, using our result and extending [19], one can show that $\mathbb{P}_n$ obeys a moderate DP given by

$$\limsup_{n\to\infty} \alpha_n^2 n^{-d} \log \mathbf{P}\{\alpha_n \mathbb{P}_n \in \Omega\} \leqq -(1/2)\chi_P^2(\Omega) \tag{4.1}$$

if Ω is closed, while

$$\liminf_{n\to\infty} \alpha_n^2 n^{-d} \log \mathbf{P}\{\alpha_n \mathbb{P}_n \in \Omega\} \geqq -(1/2)\chi_P^2(\Omega) \tag{4.2}$$

if Ω is open, and with

$$\chi_P^2(\Omega) := \inf_{Q\in\Omega} \chi_P^2(Q).$$

However, if (4.1)–(4.2) are very useful to generate conjectures, the topology on $\mathcal{M}$ is too coarse to give interesting applications.

To illustrate shortly the use of Theorem 4.3, let us consider the case where $d=1$, and introduce the weighted V-statistic

$$V_n := n^{-k} \sum_{1\leqq i_1 \leqq n} \cdots \sum_{1\leqq i_k\leqq n} w(i_1/n,\ldots,i_k/n)h(X_{i_1},\ldots,X_{i_k})$$

where $w(.)$ is a weight function and $h(.)$ a kernel which are both symmetric in their arguments. The limiting distribution of V_n or of the corresponding U-statistics has recently been obtained by O'Neil and Redner [25] in the case $k=2$ and Major [24] in the general case. Clearly, we have

$$V_n = V(\mathbb{P}_n),$$

with

$$V(\mathbb{Q}) = \int_{(0,1]^k\times S^k} w(u_1,\ldots,u_k)h(x_1,\ldots,x_k)d\mathbb{Q}(u_1,x_1)\ldots d\mathbb{Q}(u_k,x_k).$$

If $w(.)$ and $h(.)$ are bounded and continuous, V is a continuous mapping from Λ into $\mathbb{R}$. We readily infer a large deviation principle on V_n from

Theorem 4.3, using a contraction principle. The rate functional cannot be given explicitely. However, if we assume that the X_i's are real valued and admit a density $p(x)$ with respect to the Lebesgue measure, one can evaluate the rate functional in minimizing $K(\mathbb{Q}, \mathbb{P})$ under the constraint $V(\mathbb{Q}) \geqq x$. We obtain

$$I_V(x) := \int\limits_{(0,1]\times\mathbb{R}} (\log c(u) + \lambda\theta(u,x))c(u)f(x)\exp(\lambda\theta(u,x))dudx,$$

where the functions $c(.)$ and $\theta(.,.)$ are defined by the system of integral equations:

$$q(u,x) = c(u)f(x)\exp(\lambda\theta(u,x)),$$

$$c(u) = \left(\int\limits_{\mathbb{R}} f(x)\exp(\lambda\theta(u,x))dx\right)^{-1},$$

$$\theta(u,x) = \int\limits_{(0,1]^{k-1}\times\mathbb{R}^{k-1}} w(u_1,...,u_{k-1},u)h(x_1,...,x_{k-1},x)\times$$
$$\times\, q(u_1,x_1)...q(u_{k-1},x_{k-1})du_1...du_{k-1}dx_1...dx_{k-1},$$

and

$$\int\limits_{(0,1]^k\times\mathbb{R}^k} w(u_1,...,u_k)h(x_1,...,x_k)q(u_1,x_1)...q(u_k,x_k)du_1...du_k dx_1...dx_k = x.$$

If $h(.)$ is unbounded, we can use a truncature argument similar to that of Donsker and Varadhan [18] to get a LDP for V_n.

5. Proofs

5.1. Proof of Theorem 2.1

The proof follows the general scheme of Lynch and Sethuraman [22] but some difficulties happen due to the weakness of our assumptions and the lack of compactness of the level sets and the lack of continuity of the rate functionals. The proof will be captured by a series of lemmas and propositions. From now on, we assume that the assumptions of Theorem 2.1.i hold. It is convenient to introduce the notation

$$\psi(A,f) := \sum_{A_i\in A} \mu(A_i)I(f(A_i),A_i)$$

for any $f\in\mathcal{F}$ and $A\in\mathcal{P}$.

PROPOSITION 5.1.1. *The function $f \in \mathcal{F} \to J(f)$ is lsc.*

PROOF. Using (2.13), we have for any $c > 0$,

$$\{f : J(f) \leqq c\} \subset \cap_A \{f : \psi(A, f) \leqq c\}. \tag{5.1.1}$$

Then (2.8) and the lower semi-continuity of ψ ensures that the r.h.s. of (5.1.1) is an intersection of closed sets. So the l.h.s. of (5.1.1) is a closed set. □

A consequence of Proposition 5.1.1 is that the level sets $\{f : J(f) \leqq c\}$ are closed.

PROPOSITION 5.1.2. *For any open set $G \subset \mathcal{F}$,*

$$\liminf_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in G\} \geqq -J(G). \tag{5.1.2}$$

PROOF. Let $\epsilon > 0$ and let $f \in G$ such that $J(G) \leqq J(f) + \epsilon/2$. There exists a partition $A = (A_1, \dots, A_k) \in \mathcal{P}$ such that

$$J(G) \leqq \psi(A, f) + \epsilon. \tag{5.1.3}$$

For $i = 1, 2, \dots, k$ let V_i be an open neighbourhood of $f(A_i)$, and let

$$N_A(V_1, \dots, V_k) = \{h \in \mathcal{F} : \forall\ 1 \leqq i \leqq k, h(A_i) \in V_i\}.$$

Under (2.8) the set $N_A(V_1, \dots, V_k)$ is open in $\mathcal{F}$. The open sets $V_1, \dots, V_k$ may be chosen such that

$$N_A(V_1, \dots, V_k) \subset G.$$

Then the following inequalities hold:

$$\begin{aligned}
\liminf_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in G\} &\geqq \liminf_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in N_A(V_1, \dots, V_k)\} \\
&\geqq \liminf_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{\forall\ 1 \leqq i \leqq k, X_\lambda(A_i) \in V_i\} \\
&= \liminf_{\lambda \to \infty} h(\lambda)^{-1} \sum_{1 \leqq i \leqq k} \log \mathbf{P}\{X_\lambda(A_i) \in V_i\} \\
&\geqq - \sum_{1 \leqq i \leqq k} \mu(A_i) I(V_i, A_i) \\
&\geqq - \sum_{1 \leqq i \leqq k} \mu(A_i) I(f(A_i), A_i) \\
&= -\psi(A, f).
\end{aligned}$$

Finally use (5.1.3) and let ϵ tend to 0 to get (5.1.2). □

Now, we deal with the upper bound on closed sets and we assume that the assumptions of Theorem 2.1.ii hold.

The next lemma asserts a minimax property and is the analogue of Theorem 3.5 in [22], but we do not assume the compactness of level sets of $J(.)$.

LEMMA 5.1.1. *For any closed set* $\Omega \subset \mathcal{F}$,

$$\inf_{f\in\Omega} J(f) = \sup_{A\in\mathcal{P}} \inf_{f\in\Omega} \psi(A,f).$$

PROOF. Let $\eta_1 := \sup_{A\in\mathcal{P}} \inf_{f\in\Omega} \psi(A,f)$ and $\eta_2 := \inf_{f\in\Omega} J(f)$.

The inequality $\eta_1 \leqq \eta_2$ follows trivially from the fact that $\psi(A,f) \leqq J(f)$ for any $A \in \mathcal{P}$, $f \in \mathcal{F}$.

We prove $\eta_2 \leqq \eta_1$. Consider the sequence $(A_n^*)_{n\geqq 1}$ in $\mathcal{P}$ defined in (2.15). Then

$$\inf_{f\in\Omega} \psi(A_n^*, f) \leqq \eta_1.$$

Thus, for any $\epsilon > 0$, the set

$$G_n(\epsilon) := \{f \in \Omega : \psi(A_n^*, f) \leqq \eta_1 + \epsilon\}$$

is nonempty. Since $\sigma(A_n^*)$ is increasing, $(G_n(\epsilon))_{n\geqq 1}$ is a sequence of nonempty nested closed sets (the closeness comes from (2.8) and the lower semicontinuity of $I(.,A)$ for any A). Therefore,

$$G(\epsilon) := \bigcap_{n\geqq 1} G_n(\epsilon) \neq \emptyset.$$

Let $f \in G(\epsilon)$. Then, $J(f) \leqq \eta_1 + \epsilon$ due to (2.14) and (2.15). Since Ω is closed, $f \in \Omega$ and $\eta_2 \leqq \eta_1 + \epsilon$. □

PROPOSITION 5.1.3. *Let* F *be a closed set of* $\mathcal{F}$. *Then*

$$\limsup_{\lambda\to\infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in F\} \leqq -J(F).$$

PROOF. Let $A = (A_1, \ldots, A_k) \in \mathcal{P}$ and define the closed subset of $\mathcal{X}^k$

$$F_A := \{(f(A_i))_{1\leqq i\leqq k} : f \in F\}.$$

Also define

$$V_\epsilon := \Big\{x \in \mathcal{X}^k : \sum_{1\leqq i\leqq k} \mu(A_i) I(x_i, A_i) > \inf_{f\in F} \psi(A,f) - \epsilon\Big\}.$$

Since X_λ is MDT, there exists a compact subset K of $\mathcal{X}^k$ such that

$$\limsup_{\lambda\to\infty} h(\lambda)^{-1} \log \mathbf{P}\{(X_\lambda(A_i))_{1\leqq i\leqq k} \notin K\} \leqq -M < J(F) - 1. \tag{5.1.4}$$

Since $I(.,A_i)$ is lsc, so is $\psi(A,.)$, and V_ϵ is open in $\mathcal{X}^k$. Thus, $V_\epsilon \cap K$ is open in K equipped with the relative topology. Since K is normal (see e.g.

Bourbaki [11] (I.9, no. 3) and [10] (IX.4, no. 1), we can find an open set $U_\epsilon \subset K$ such that

$$F_A \cap K \subset U_\epsilon \subset \mathrm{cl}_K(U_\epsilon) \subset V_\epsilon.$$

For $x = (x_1, \ldots, x_k) \in \mathcal{X}^k$, let

$$I_A(x) := \sum_{1 \leqq i \leqq k} \mu(A_i) I(x, A_i).$$

Define further

$$I_A(F) := \inf_{x \in F} I_A(x).$$

Using the fact that $\mathrm{cl}_K(U_\epsilon)$ is closed in $\mathcal{X}^k$,

$$\begin{aligned} \limsup_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in F\} &\leqq \limsup_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{(X_\lambda(A_i))_{1 \leqq i \leqq k} \in F_A \cap K\} \\ &\leqq \limsup_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{(X_\lambda(A_i))_{1 \leqq i \leqq k} \in \mathrm{cl}_K(U_\epsilon)\} \\ &\leqq -I_A(\mathrm{cl}_K(U_\epsilon)) \end{aligned} \tag{5.1.5}$$

(the second inequality uses (5.1.4) and the last one comes from Exercise 4.2.7 of Dembo and Zeitouni [16]. Next, observe that

$$\inf_{f \in F} \psi(A, f) - \epsilon \leqq I_A(V_\epsilon) \leqq I_A(\mathrm{cl}_K(U_\epsilon)) \tag{5.1.6}$$

so that making ϵ tend to 0, (5.1.4), (5.1.5) and (5.1.6) give for any $A \in \mathcal{P}$,

$$\limsup_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in F\} \leqq -\psi(A, F). \tag{5.1.7}$$

Of course, we can choose A such that

$$\psi(A, F) > J(F) - 1$$

so that (5.1.4) and (5.1.7) yield

$$\limsup_{\lambda \to \infty} h(\lambda)^{-1} \log \mathbf{P}\{X_\lambda \in F\} \leqq -\psi(A, F). \tag{5.1.8}$$

Then, take the infimum over $A \in \mathcal{P}$ in the rhs of (5.1.8) and apply Lemma 5.1.1 to get the result. □

REMARK 5.1.1. A careful look at our proof shows that the independence of the increments of the processes X_λ is used only when we need to apply Exercise 4.2.7 of [16]. Thus, our result still holds under the much weaker assumption (A-1) of [15].

5.2. *Proof of Theorem 3.1*

Let $\mathcal{A} := \{(a, b], 0 \leqq a < b \leqq 1\}$. Observe that if $A \in \mathcal{A}$, then

$$\lim_{n\to\infty} \#(nA)/n = |A|.$$

We define the partial extreme process indexed by $\mathcal{A}$

$$\mu_n(A) := \max_{i \in nA} X_i.$$

One easily shows that $\mu_n(.)$ obey a MDP as in Definition 2.2 with $h(n, A) = n$ and $I(x, A) = -|A| \log F(x)$. Obviously (2.7), (2.12) and (2.16) hold.

Now, we view μ_n as a random element in the set $\mathcal{D}$ of all functions g from $\mathcal{A}$ into $\mathbb{R}$ such that for any disjoint subsets $A, B \in \mathcal{A}$, $g(A \cup B) = g(A) \vee g(B)$ and $g(0, 1] < \infty$. The set $\mathcal{D}$ is equipped with the topology of pointwise convergence and we are going to apply Theorem 2.1.

Clearly, condition (2.12) holds with $\mu(A) = 1$. To check (2.13) take $g \in \mathcal{D}$, $B, C \in \mathcal{A}$ disjoint and let further $A = B \cup C$. Then

$$\begin{aligned} I(g(A), A) &= -|A| \log F(g(A)) \\ &= -|A| \log F(g(B) \vee g(C)) \\ &= -|A| (\log F(g(B)) \vee \log F(g(C))) \\ &\leqq -|B| \log F(g(B)) - |C| \log F(g(C)), \end{aligned}$$

so that (2.13) holds. The rate functional will be given for $g \in \mathcal{D}$ by

$$J(g) = \sup_{A \in \mathcal{P}} - \sum_{A_i \in A} |A_i| \log F(g(A_i)). \tag{5.2.1}$$

To check (2.15), let

$$L_{i,m} := ((i-1)/2^m, i/2^m], \qquad 1 \leqq i \leqq 2^m,$$

and define

$$A_m^* := \{L_{i,m} : 1 \leqq i \leqq 2^m\}.$$

Since $L_{i,m} \in \mathcal{A}$ we just need to prove that

$$\liminf_{m\to\infty} - \sum_{1 \leqq i \leqq 2^m} |L_{i,m}| \log F(g(L_{i,m})) \geqq J(g). \tag{5.2.2}$$

For this, let $\epsilon > 0$ and let $A \in \mathcal{P}$ such that

$$\psi(A, g) = - \sum_{A_i \in A} |A_i| \log F(g(A_i)) \geqq J(g) - \epsilon.$$

Next, write

$$\psi(A_m^*, g) = \psi_{1,m} + \psi_{2,m} + \psi_{3,m}$$

with, for m large enough

$$\psi_{1,m} := -\sum_{A_i \in A} \sum_{L_{j,m} \cap A_i \neq \emptyset} |L_{j,m} \cap A_i| \log F(g(L_{j,m} \cap A_i)),$$

$$\psi_{2,m} := \sum_{A_i \in A} \sum_{\begin{cases} L_{j,m} \cap A_i \neq \emptyset \\ L_{j,m} \cap A_i^c \neq \emptyset \end{cases}} |L_{j,m} \cap A_i| \log F(g(L_{j,m} \cap A_i)),$$

$$\psi_{3,m} := -(1/2) \sum_{A_i \in A} \sum_{\begin{cases} L_{j,m} \cap A_i \neq \emptyset \\ L_{j,m} \cap A_i^c \neq \emptyset \end{cases}} |L_{j,m}| \log F(g(L_{j,m} \cap A_i)),$$

(the factor $1/2$ in $\psi_{3,m}$ comes from the fact that if $L_{j,m}$ intersects A_i but is not included in A_i then it intersects exactly one other A_j if m is such that $2^{-m} < \min_{A_i \in A} |A_i|$). Due to the subadditivity of $I(g(.), .)$, we have

$$\psi_{1,m} \geqq \psi(A, g). \tag{5.2.3}$$

Let $\epsilon > 0$. Since an $L_{j,m}$ intersects at most two $A_i \in A$ and $|L_{j,m}| = 2^{-m}$, we have

$$0 \geqq \psi_{2,m} \geqq 2(\#A)2^{-m} \log F(g(0,1]) \geqq -\epsilon$$

if m is large enough, and clearly $\psi_{3,m} \geqq 0$. Therefore with (5.2.3) we deduce that

$$\psi(A_m^*, g) \geqq \psi(A, g) - 2\epsilon$$

for any m large enough. Thus, (5.2.2) holds and so (2.15).

We can apply Theorem 2.1 to get that $\mu_n(.)$ obeys a FDP with rate $J(.)$ defined in (5.2.1).

Now, we map the partial extreme process $\mu_n(.)$ indexed by $\mathcal{A}$ into partial extreme process $M_n(.)$. For this, let θ be the function from $\mathcal{D}$ into $\mathcal{D}\uparrow(0,1]$ defined for any $g \in \mathcal{D}$ by

$$\theta(g)(x) = g(0, x], \qquad 0 < x \leqq 1.$$

Clearly, θ is continuous from $\mathcal{D}$ into $\mathcal{D}\uparrow(0,1]$ both equipped with the topology of pointwise convergence, and moreover $\theta(\mu_n) = M_n$. Hence, we can apply the contraction principle, and $M_n(.)$ obeys a FDP with rate $J(\theta^{-1}(.))$. It remains to prove that for any $g \in \mathcal{D}\uparrow(0,1]$,

$$J(\theta^{-1}(g)) = -\int_{(0,1]} \log F(g(x))dx. \tag{5.2.4}$$

Define the mapping Θ from $D\uparrow(0,1]$ into $\mathcal{D}$ by

$$\Theta(g)(A) := \sup_{x\in A} g(x).$$

Then, for any $g\in\mathcal{D}\uparrow(0,1]$, the function $\Theta(g)$ belongs to $\theta^{-1}(g)$ so that

$$J(\Theta(g)) \geqq J(\theta^{-1}(g)). \tag{5.2.5}$$

But using Riemann's integral,

$$\begin{aligned} J(\Theta(g)) &= \sup_{A\in\mathcal{P}} \sum_{A_i\in A} |A_i| \inf_{x\in A_i} -\log F(g(x)) \\ &= \int_{(0,1]} -\log F(g(x))dx. \end{aligned}$$

Next, let $h\in\theta^{-1}(g)$, so that $h(0,x]=g(x)$ for any $x\in(0,1]$. If $A\in\mathcal{P}$, we can write $A=\{(a_i,b_i]:1\leqq i\leqq \#A\}$ and get

$$\begin{aligned} \psi(A,h) &= -\sum_{A_i\in A} |A_i| \log F(h(A_i)) \\ &\leqq -\sum_{A_i\in A} |A_i| \log F(h(0,b_i]) \\ &= -\sum_{A_i\in A} (b_i-a_i)\log F(g(b_i)), \end{aligned} \tag{5.2.6}$$

where the first inequality comes from the fact that $h\in\mathcal{D}$. Then, since $g(.)$ is monotone, the sum (5.2.6) tends to $\int_{(0,1]} \log F(g(x))dx$ when $\max_{1\leqq i\leqq n} b_i-a_i$ tends to 0. Thus,

$$\limsup_{m\to\infty} \psi(A_m^*,h) \leqq -\int_{(0,1]} \log F(g(x))dx. \tag{5.2.7}$$

Then, (5.2.5), (5.2.7) and definition and property (2.15) of A_n^* show that (5.2.4) holds, and so Theorem 3.2 is proved. □

5.3. Proof of Theorem 4.1

To use Theorem 2.1, we first define the measured-valued measure on the Borel σ-field of $(0,1]^d$,

$$\mathbb{P}_n\langle A\rangle := n^{-d}\sum_{i\in nA}\delta_{X_i}, \qquad A\in\mathcal{B}(0,1]^d.$$

Clearly, $\{A \to \mathbb{P}_n\langle A\rangle\}_{n\geqq 1}$ is a sequence of processes with independent increments as defined in Section 1. The weak convergence and strong approximation of these processes when the X_i's are real valued and A runs over all the intervals $(0,t]$, $0<t\leqq 1$, have been studied by Bickel and Wichura [5], Kiefer [20], Komlós, Major and Tusnády [21], M. Csörgő and Révész [14].

We shall use a slightly different construction than in the proof of Theorem 3.2 since we shall use a bigger set $\mathcal{A}$ to index $\mathbb{P}\langle\cdot\rangle$. We shall first prove a MDP, then apply Theorem 2.1 to get the functional version, and finally, we shall identify the process $\mathbb{P}_n\langle\cdot\rangle$ to $\mathbb{P}_n$ as defined in Section 4.

Our proof is captured by a series of lemmas and propositions. But let us first introduce some notation. Let M_1^+ be the space of all nonnegative measures on S with total mass less or equal 1 equipped with the τ-topology (see [19]). Clearly, for any $A\in\mathcal{B}(0,1]^d$, $\mathbb{P}_n\langle A\rangle\in M_1^+$. For $\mu,\nu\in M_1^+$, we define the Kullback–Leibler information number between μ and ν by

$$\widetilde{K}(\mu,\nu):=\begin{cases}\int\log(d\mu/d\nu)dQ & \text{if } Q\ll R \text{ and } \mu(S)=\nu(S),\\ \infty & \text{otherwise.}\end{cases}$$

If Ω is a subset of M_1^+ and $\nu\in M_1^+$, we define

$$\widetilde{K}(\Omega,\nu):=\inf_{\mu\in\Omega}\widetilde{K}(\mu,\nu).$$

REMARK 5.3.1. The space M_1^+ equipped with the τ-topology is not metrizable and the function $\widetilde{K}(.,\nu)$ is lsc but not continuous. So the full form of assumptions of Theorem 2.1 is required here.

For any $\epsilon,\lambda>0$ and $\Omega\subset M_1^+$, we introduce

$$\Omega_{\epsilon,\lambda}:=\{t\mu:\mu\in\Omega,(1-\epsilon)/\lambda\leqq t\leqq(1+\varepsilon/\lambda\}.$$

Next, if $B=(B_1,\dots,B_k)$ is a partition on S, with $B_i\in\mathcal{B}(S)$ we define a pseudo-metric on M_1^+ by

$$d_B(\mu,\nu):=\max_{1\leqq i\leqq k}|\mu(B_i)-\nu(B_i)|$$

for any $\mu,\nu\in M_1^+$. Then we set

$$\Omega^{\epsilon,\lambda}:=\{\mu/\lambda:\mu\in\Omega,\exists B\in\mathcal{P}(S),\{\nu:d_B(\mu,\nu)<\epsilon\}\subset\Omega\}.$$

Clearly, for any $\Omega\subset M_1^+$, we have

$$\lambda\Omega^{\epsilon,\lambda}\subset\Omega\subset\lambda\Omega_{\epsilon,\lambda} \tag{5.3.1}$$

and $\Omega^{\epsilon,\lambda}$ is τ-open if Ω is open, while $\Omega_{\epsilon,\lambda}$ is τ-closed if Ω is closed.

LEMMA 5.3.1. *If Ω is τ-closed in M_1^+, then*

$$\lim_{\epsilon\to 0} \widetilde{K}(\Omega_{\epsilon,\lambda}, \mathbf{P}) = \widetilde{K}(\Omega/\lambda, \mathbf{P}) \tag{5.3.2}$$

and if Ω is τ-open in M_1^+, then

$$\lim_{\epsilon\to 0} \widetilde{K}(\Omega^{\epsilon,\lambda}, \mathbf{P}) = \widetilde{K}(\Omega/\lambda, \mathbf{P}). \tag{5.3.3}$$

PROOF. (i) We consider the case where Ω is τ-closed. The inclusion (5.3.1) implies

$$\limsup_{\epsilon \overset{>}{\to} 0} \widetilde{K}(\Omega_{\epsilon,\lambda}, \mathbf{P}) \leqq \widetilde{K}(\Omega/\lambda, \mathbf{P}). \tag{5.3.4}$$

To prove (5.3.2) assume first that

$$\liminf_{\epsilon \overset{>}{\to} 0} \widetilde{K}(\Omega_{\epsilon,\lambda}, \mathbf{P}) = c < \infty.$$

Since $\Omega_{\epsilon,\lambda}$ is τ-closed, Lemma 3.2 of [18] shows that there exists $g_\epsilon \in \Omega_{\epsilon,\lambda}$ such that

$$\widetilde{K}(g_\epsilon, \mathbf{P}) = \widetilde{K}(\Omega_{\epsilon,\lambda}, \mathbf{P}).$$

There exists a sequence, say g_n, such that

$$\lim_{n\to\infty} \widetilde{K}(g_n, \mathbf{P}) = c.$$

Thus, for n large enough, we can assume that g_n belongs to the compact set $\{\mu : \widetilde{K}(\mu, \mathbf{P}) \leqq c + 1\}$ (see Lemma 2.3 in [19] for the compactness), so that there exists a subsequence g_{n_k} converging to some g. Since $\Omega_{\epsilon,\lambda}$ is closed, $\Omega/\lambda = \bigcap_{\epsilon>0} \Omega_{\epsilon,\lambda}$ and $g \in \Omega/\lambda$. Using the lsc of $\widetilde{K}(.,\mathbf{P})$ (see Lemma 2.2 of [19]), we have

$$c = \liminf_{k\to\infty} \widetilde{K}(g_{n_k}, \mathbf{P}) \geqq \widetilde{K}(g, \mathbf{P}) \geqq \widetilde{K}(\Omega/\lambda, \mathbf{P}). \tag{5.3.5}$$

Clearly, (5.3.4) and (5.3.5) yield (5.3.2).

(ii) Now, consider a τ-open Ω in M_1^+. Then (5.3.1) gives

$$\liminf_{\epsilon \overset{>}{\to} 0} \widetilde{K}(\Omega^{\epsilon,\lambda}, \mathbf{P}) \geqq \widetilde{K}(\Omega/\lambda, \mathbf{P}). \tag{5.3.6}$$

Next, let $\mu_n \in \Omega/\lambda$ such that

$$\lim_{n\to\infty} \widetilde{K}(\mu_n, \mathbf{P}) = \widetilde{K}(\Omega/\lambda, \mathbf{P}).$$

Since Ω is τ-open, there exist $B_n \in \mathcal{P}(S)$ and $\delta_n > 0$ such that

$$\{\nu : d_{B_n}(\mu_n, \nu) < \delta_n\} \subset \Omega/\lambda.$$

Thus, $\mu_n \in \Omega^{\delta_n,\lambda}$ and $\widetilde{K}(\Omega^{\delta_n,\lambda}, \mathbf{P}) \leqq \widetilde{K}(\mu_n, \mathbf{P})$ and

$$\limsup_{n\to\infty} \widetilde{K}(\Omega^{\delta_n,\lambda}, \mathbf{P}) \leqq \widetilde{K}(\Omega/\lambda, \mathbf{P}). \tag{5.3.7}$$

Since $\widetilde{K}(\Omega^{\epsilon,\lambda}, \mathbf{P})$ decreases when ϵ decreases to 0, (5.3.7) yields to

$$\limsup_{\epsilon\to 0} \widetilde{K}(\Omega^{\epsilon,\lambda}, \mathbf{P}) \leqq \widetilde{K}(\Omega/\lambda, \mathbf{P}),$$

which with (5.3.6) is (5.3.3). □

We shall consider $\mathbb{P}_n\langle\cdot\rangle$ as a set indexed process, and with the notation of Section 2, we consider $\mathcal{A}$ the set of all Borel sets $A \subset (0,1]^d$ such that

$$\lim_{n\to\infty} n^{-d} \#(nA) = \lambda_d(A) = |A|. \tag{5.3.8}$$

For example, any open or closed subset of $(0,1]^d$ belongs to $\mathcal{A}$.

We can now prove a MDP for $\mathbb{P}_n\langle\cdot\rangle$.

PROPOSITION 5.3.1. *The processes $\{\mathbb{P}_n\langle.\rangle\}_{n\geq 1}$ obey a MDP given by*

$$\limsup_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle A\rangle \in \Omega\} \leqq -|A|\widetilde{K}(\Omega/|A|, \mathbf{P}) \tag{5.3.9}$$

if Ω is τ-closed in M_1^+, while

$$\liminf_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle A\rangle \in \Omega\} \geqq -|A|\widetilde{K}(\Omega/|A|, \mathbf{P}) \tag{5.3.10}$$

if Ω is τ-open.

PROOF. It is convenient to introduce an i.i.d. sequence $\{Y_i\}_{i\geqq 1}$ of S-valued r.v.'s, with common p.m. $\mathbf{P}$. Then let

$$\mathbf{P}_n := n^{-1} \sum_{1\leqq i\leqq n} \delta_{Y_i}$$

be their empirical p.m. Clearly, for any $A \subset (0,1]^d$,

$$\mathbb{P}_n\langle A\rangle \stackrel{d}{=} n^{-d} \#(nA) \mathbf{P}_{\#(nA)}.$$

Then, since (5.3.8) holds, for any $A \in \mathcal{A}$, any $\epsilon > 0$ and any n large enough,

$$(1-\epsilon)/|A| \leqq n^d/\#(nA) \leqq (1+\epsilon)/|A|.$$

Consequently, for any $\epsilon > 0$ and $A \in \mathcal{A}$,

$$\limsup_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle A\rangle \in \Omega\} \leqq \limsup_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbf{P}_{\#(nA)} \in \Omega_{\epsilon,|A|}\}$$
$$\leqq -|A|\widetilde{K}(\Omega_{\epsilon,|A|}, \mathbf{P}),$$

where the last inequality comes from Lemma 3.1 of [19]. Then use Lemma 5.3.1 and let $\epsilon \to 0$ to obtain (5.3.9).

Now, we prove the lower bound (5.3.10). If $\widetilde{K}(\Omega/|A|, \mathbf{P}) = \infty$, the result is trivial, so we can assume that the rhs of (5.3.10) is finite.

Observe that for any $\epsilon > 0$ and provided n is large enough, then

$$\Omega^{\epsilon,|A|} \subset (n^d/\#(nA))\Omega. \tag{5.3.11}$$

Indeed, for $\mu \in \Omega$, let $\mu_n := \mu\#(nA)/(n^d|A|)$. Since $A \in \mathcal{A}$, we have

$$\lim_{n\to\infty} \sup_{\mu\in\Omega} \sup_{B\in\mathcal{P}(S)} d_B(\mu, \mu_n) = 0,$$

which gives (5.3.11).

Then Lemma 3.1 in [18] and (5.3.11) and (5.3.1) imply for any $\epsilon > 0$,

$$\liminf_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle A\rangle \in \Omega\} \geqq -|A|\widetilde{K}(\Omega^{\epsilon,|A|}, \mathbf{P}).$$

Let ϵ tend to 0 and use Lemma 5.3.1 to get (5.3.10). □

The next step is to use Theorem 2.1 to get the functional version of Proposition 5.3.1. For this, let $\mathcal{F}$ be the set of all σ-additive functions from $\mathcal{A}$ to M_1^+ equipped with the topology of pointwise convergence.

PROPOSITION 5.3.2. *The family of processes $\{\mathbb{P}_n\langle . \rangle\}_{n\geqq 1}$ obeys a FDP. If Ω is closed in $\mathcal{F}$, then*

$$\limsup_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle . \rangle \in \Omega\} \leqq -J(\Omega, \mathbf{P})$$

and if Ω is open in $\mathcal{F}$,

$$\liminf_{n\to\infty} n^{-d} \log \mathbf{P}\{\mathbb{P}_n\langle . \rangle \in \Omega\} \geqq -J(\Omega, \mathbf{P}),$$

where

$$J(\Omega, \mathbf{P}) := \inf_{g\in\Omega} J(g, \mathbf{P}),$$

$$J(g, \mathbf{P}) = \sup_{A\in\mathcal{P}} \sum_{A_i\in A} |A_i|\widetilde{K}(g(A_i)/|A_i|, \mathbf{P}) \tag{5.3.12}$$

and $\mathcal{P}$ denotes the set of all partitions of $(0,1]^d$ *in sets belonging to* $\mathcal{A}$.

PROOF. We check that the conditions of Theorem 2.1 hold.

Check of condition (2.7): We first prove that $\mathbf{P}_n$ is large deviation tight in modifying the arguments of Lynch and Sethuraman ([22], Lemma 2.6) in our case where the underlying space is not metrizable. Let $(\delta_k)_{k\geqq 1}$ be a sequence in $(0,1/2)$ converging to 0 as $k\to\infty$, and let η_k such that

$$(\delta_k/2)\log(\delta_k/2\eta_k) > k + e^{-1}. \tag{5.3.13}$$

Since $\mathbf{P}$ is tight, there exists a compact set $C_k \subset S$ such that

$$\mathbf{P}(C_k) > 1 - \eta_k.$$

Define the partitions $A_k := \{A_{k,1} := S - C_k, A_{k,2} := C_k\}$ and let for any $c>0$

$$\Gamma_c := \{Q \in \Lambda(S) : K(Q,\mathbf{P}) \leqq c\},$$

where $\Lambda(S)$ denotes the space of all probability measures on S endowed with the τ-topology, and $K(.,.)$ the usual Kullback–Leibler number of information. For any $R \in \Lambda(S)$, and $A \in \Lambda(S)$ let

$$B(R,A,\delta) := \{Q \in \Lambda(S) : d_A(Q,R) := \sum_{A_i\in A} |R(A_i) - Q(A_i)| < \delta\}.$$

For any R, A and $\delta, B(R,A,\delta)$ is τ-open in $\Lambda(S)$, and

$$\{B(R,A_k,\delta_k) : R \in \Gamma_{2kM}\}$$

is an open covering of Γ_{2kM} which is τ-compact (see Lemma 2.3 of [19]). Thus, we can extract a finite covering

$$\{B(R_i,A_k,\delta_k) : i \in I_k^{(1)}\} \qquad (\#I_k^{(1)} < \infty)$$

and if we define

$$W_k := \bigcup_{i\in I_k^{(1)}} B(R_i,A_k,\delta_k),$$

we get, using that $\Gamma_{2kM} \subset W_k$,

$$\begin{aligned}\limsup_{n\to\infty} n^{-1}\log\mathbf{P}\{\mathbf{P}_n \notin W_k\} &\leqq -K(W_k^c,\mathbf{P}) \leqq -K(\Gamma_{2kM}^c,\mathbf{P})\\ &\leqq -2kM.\end{aligned}$$

Consequently, we can extend $I_k^{(1)}$ into a finite set I_k such that

$$\mathbf{P}\{\mathbf{P}_n \notin V_k\} \leqq \exp(-knM)$$

for any $n \geqq 1$, with

$$V_k := \bigcup_{i \in I_k} B(R_i, A_k, \delta_k).$$

Next, let

$$G_k := \bigcup_{i \in I_k} \mathrm{cl}_\tau B(R_i, A_k, \delta_k).$$

Then, $W_k \subset V_k \subset G_k$, and let further

$$C := \bigcap_{k \geqq 1} G_k.$$

We now prove that

$$\limsup_{n \to \infty} n^{-1} \log \mathbf{P}\{\mathbf{P}_n \notin C\} \leqq -M,$$

and that C is compact. Clearly,

$$\begin{aligned}\mathbf{P}\{\mathbf{P}_n \notin C\} = \mathbf{P}\Big\{\mathbf{P}_n \in \bigcup_{k \geqq 1} \bigcup_{i \in I_k} G_k^c\Big\} &\leqq \sum_{k \geqq 1} \exp(-knM) \\ &= \exp(-nM)/(1-\exp(-nM)).\end{aligned}$$

Thus, $\mathbf{P}_n$ is LDT if we prove that C is τ-compact. Following the arguments of Lemma 2.3 of [19], we just need to prove that C is τ-closed in the space Ξ of all functions from $\mathcal{B}(S)$ into $[0,1]$ equipped with the τ-topology. In fact, all we need to prove is that if $(Q_n)_{n \geqq 1}$ is a sequence in C which converges to some Q in Ξ, then Q is also σ-additive.

If $Q_n \in C$, for any $k, Q_n \in G_k$ which implies that there exists $R \in \Gamma_{2kM}$ such that $Q_n \in \mathrm{cl}_\tau B(R, A_k, \delta_k)$. Since $R \in \Gamma_{2kM}$ and $P(A_{k,1}) < \eta_k$, the proof of Lemma 2.3 of [19] shows that $R(S - C_k) < \delta_k$ since (5.3.13) holds. Therefore,

$$\begin{aligned}Q_n(A_{k,1}) &= Q_n(S - C_k) \\ &\leqq |Q_n(A_{k,1}) - R(A_{k,1})| + R(A_{k,1}) \\ &\leqq 3\delta_k\end{aligned}$$

and (Q_n) is a tight sequence. Since $(Q_n)_{n \geqq 1}$ converges to Q in Ξ, Q is also a probability measure, and this proves that $\mathbf{P}_n$ is LDT.

Since $\mathbb{P}_n\langle A\rangle \overset{d}{=} \mathbf{P}_{\#(nA)}\#(nA)/n^d$ and $\lim_{n\to\infty} \#(nA)/n^d = |A|$, $\mathbb{P}_n\langle . \rangle$ is MDT, and assumption (2.7) holds.

Check of conditions (2.12), (2.13), (2.15) and (2.16): condition (2.12) is obvious. To check (2.13), take $g \in \mathcal{F}$ and $A = B \cup C$ with $B \cap C = \emptyset$ and $A, B, C \in \mathcal{A}$. Then

$$\begin{aligned}&|A|\widetilde{K}(g(A)/|A|, \mathbf{P}) \\ &= (|B|+|C|)\widetilde{K}((|B|/(|B|+|C|))g(B)/|B| + (|C|/(|B|+|C|))g(C)/|C|, \mathbf{P}) \\ &\leqq |B|\widetilde{K}(g(B)/|B|, \mathbf{P}) + |C|\widetilde{K}(g(C)/|C|, \mathbf{P})\end{aligned}$$

the last inequality coming from the fact that the function $x \log x$ is concave and Jensen's inequality.

Assumption (2.15) is checked hereafter, along the proof of Lemma 5.3.2, and (2.16) obviously holds.

Theorem 2.1 gives Proposition 5.3.2. □

To end the proof of Theorem 4.3, we now proceed to the identification of $\mathbb{P}_n\langle . \rangle$ and $\mathbb{P}_n$.

Notice first that $\mathcal{A}$ is a field, and $g\langle A \rangle$, $A \in \mathcal{A}$ is a measure on $\mathcal{A}$. It can be extended on the σ-field $\sigma(\mathcal{A})$. Since $\mathcal{A}$ contains all the open and closed sets of $(0,1]^d$, $\sigma(\mathcal{A}) = \mathcal{B}(0,1]^d$.

Next, we define a mapping ψ from Λ to $\mathcal{F}$ by, for any $Q \in \Lambda$,

$$\text{for any } \; A \in \mathcal{A}, \; B \in \mathcal{B}(S), \; \psi(Q)\langle A \rangle(B) = Q(A \times B).$$

Since the set of all rectangles $\{A \times B : A \in \mathcal{A}, \; B \in \mathcal{B}(S)\}$ is a π-system (see Billingsley [7]) which generates the σ-field $\mathcal{B}(0,1]^d \otimes \mathcal{B}(S)$, ψ defines a one-to-one mapping from Λ into $\psi(\Lambda)$. Clearly $\mathbb{P}_n\langle . \rangle = \psi(\mathbb{P}_n)\langle . \rangle$. Consequently, if $\Omega \subset \Lambda$,

$$\mathbf{P}\{\mathbb{P}_n \in \Omega\} = \mathbf{P}\{\psi(\mathbb{P}_n) \in \psi(\Omega)\}.$$

The mapping ψ is continuous for the topologies we defined on Λ and $\mathcal{F}$, and using the contraction principle, we infer from Proposition 5.3.2 that the sequence $(\mathbb{P}_n)_{n \geq 1}$ obeys a LDP with rate function $J(\psi(Q), \mathbf{P})$ where $J(., \mathbf{P})$ is defined in (5.3.12). Therefore, Theorem 3.2 holds if the following lemma holds.

LEMMA 5.3.2. *For any $Q \in \Lambda$, we have*

$$J(\psi(Q), \mathbf{P}) = K(Q, \mathbb{P})$$

where $J(.,.)$ is defined in (5.3.12) *and $K(.,.)$ in Section 2.*

PROOF. (i) We first show that $J \leq K$: if $A \in \mathcal{A}$, and Q is a measure on $(0,1]^d \times S$, we denote $Q(A \times .)$ the measure on S defined by $Q(A \times .)(B) := Q(A \times B)$ for any $B \in \mathcal{B}(S)$. Then,

$$J(\psi(Q), \mathbf{P}) = \sup_{A \in \mathcal{P}} \sum_{A_i \in A} |A_i| \widetilde{K}(Q(A_i \times .)/|A_i|, \mathbf{P}).$$

Using the definition of $\widetilde{K}$ and relation (2.2) in [18],

$$\begin{aligned} &J(\psi(Q), \mathbf{P}) \\ &= \sup_{A \in \mathcal{P}} \sum_{A_i \in A} \Big(\sup_{B \in \mathcal{P}(S)} \sum_{B_j \in B} \log(Q(A_i \times B_j)/|A_i| \mathbf{P}(B_j)) Q(A_i \times B_j) \\ &\qquad + \infty \mathbb{I}\{Q(A_i \times S) \neq |A_i|\} \Big) \end{aligned}$$

$$\leqq \sup_{C\in\mathcal{P}_1} \sum_{C_i\in C} \log(Q(C_i)/(\lambda\times\mathbf{P})(C_i))Q(C_i)$$
$$+\infty \sup_{A\in\mathcal{A}} \mathbb{I}\{Q(A\times S)\neq |A|\}$$
$$= K(Q,\lambda\times\mathbf{P}) + \infty \sup_{A\in\mathcal{A}} \mathbb{I}\{Q(A\times S)\neq |A|\}$$

where $\mathcal{P}_1$ denotes the set of all finite partitions of $(0,1]^d\times S$ into measurable subsets.

(ii) We prove now that $K(.,.)\leqq J(\psi(.),.)$ and that (2.15) holds: We first assume that $Q\ll\lambda\times\mathbf{P}$ and define $q(x,y)=(dQ/d(\lambda\times\mathbf{P}))(x,y)$. Let $T(x)=x\log x$.

For any $A\in\mathcal{P}$, let

$$S(A) := \sum_{A_i\in A} |A_i|\widetilde{K}(Q(A_i\times .)/|A_i|,\mathbf{P}).$$

One can easily check that

$$S(A) = \sum_{A_i\in A} |A_i| \int_S T\Big(|A_i|\int_{A_i} q(x,y)d\lambda_d(x)\Big)d\mathbf{P}(y).$$

For any $\tau\geqq 1$, let

$$q_\tau = q\mathbb{I}\{q\leqq\tau\} + \tau\mathbb{I}\{q>\tau\}$$

be the density $q(.,.)$ truncated at the level τ. Since $T(.)$ is increasing on $(1,\infty)$, we have

$$S(A) \geqq \sum_{A_i\in A} |A_i| \int_S T(|A_i|\int_{A_i} q_\tau(x,y)d\lambda_d(x))d\mathbf{P}(y). \tag{5.3.14}$$

Now, we define the sequence of partitions

$$A_n^* := \{A_{n,i} : i\in 2^n(0,1]^d\cap\mathbb{N}^d\} \qquad \text{and}$$
$$A_{n,i} := \prod_{1\leqq k\leqq d} ((i_k-1)/2^n, i_k/2^n] \qquad \text{for } i=(i_1,\dots,i_k).$$

For any $x\in(0,1]^d$, let $i_n(x)$ be the index such that $x\in A_{n,i_n(x)}$. Then let

$$g_{n,\tau}(x,y) := \begin{cases} T(|A_{n,i_n(x)}|^{-1} \int_{A_{n,i_n(x)}} q_\tau(u,y)d\lambda_d(u))d\mathbf{P}(y) & \text{if } x\notin\mathbb{Q}^d \\ 0 & \text{if } x\in\mathbb{Q}^d. \end{cases}$$

One easily checks that the rhs of (5.3.14) is

$$\int_{(0,1]^d\times S} g_{n,\tau}(x,y)d\lambda_d(x)d\mathbf{P}(y).$$

We prove that, as $n \to \infty$ and along A_n^*, the family of step functions $g_{n,\tau}$ converges pointwise to some function g_τ which in turn converges to q as τ tends to infinity.

Since for any $x \in (0,1]^d$ the sets $(A_{n,i_n(x)})_{n \geqq 1}$ shrink nicely to x in the sense of Rudin ([26], Ch. 8), we have

$$\lim_{n\to\infty} g_{n,\tau}(x,y) = g_\tau(x,y)$$

where

$$g_\tau(x,y) := \begin{cases} T(q_\tau(x,y)) & \text{if } x \notin \mathbb{Q}^d \\ 0 & \text{otherwise.} \end{cases}$$

Since q_τ is upper bounded by τ, we also have

$$g_{n,\tau}(x,y) \in [0, T(\tau)] \quad \text{for any} \quad x, y.$$

Thus, Lebesgue's dominated convergence theorem yields

$$\lim_{n\to\infty} \int_{(0,1]^d \times S} g_{n,\tau} d(\lambda_d \times \mathbf{P}) = \int_{(0,1]^d \times S} g_\tau d(\lambda_d \times \mathbf{P}). \tag{5.3.15}$$

Clearly, the measure defined on $(0,1]^d \times S$ by

$$\int_{A\times B} q_\tau d(\lambda_d \times \mathbf{P})$$

converges (when $\tau \to \infty$) to Q on any Borel set of $(0,1]^d \times S$. Then, the lower semi-continuity of the Kullback–Leibler information number ensures that

$$\liminf_{\tau\to\infty} \int_{(0,1]^d \times S} g_\tau d(\lambda_d \times \mathbf{P}) \geqq K(Q, \mathbb{P}). \tag{5.3.16}$$

Combine (5.3.14)–(5.3.16) to show that

$$\liminf_{n\to\infty} S(A_n^*) \geqq K(Q, \mathbb{P}). \tag{5.3.17}$$

With part (i) of the proof of Lemma 5.3.2, (5.3.17) shows that the sequence $(A_n^*)_{n\geqq 1}$ verifies condition (2.15) as soon as $Q \ll \lambda_d \times \mathbf{P}$.

Finally, if Q is not absolutely continuous w.r.t. $\lambda_d \times \mathbf{P}$, we can take $Q + (\lambda_d \times \mathbf{P})$ as a dominating measure and use the same technique to show that

$$\lim_{n\to\infty} S(A_n^*) = \infty$$

in this case. □

REFERENCES

[1] ACOSTA, A. D. DE, On large deviations of empirical measures in the τ-topology, *J. Appl. Probab.* **31A** (1994), 41–47. *MR* **95d**:60055

[2] BAHADUR, R. R. and ZABELL, S. L., Large deviations of the sample mean in general vector spaces, *Ann. Probab.* **7** (1979), 587–621. *MR* **80i**:60031

[3] BARBE, PH. and BRONIATOWSKI, M., Deviation principle for set indexed processes with independent increments, Technical Report, IRMA, Strasbourg, 1993.

[4] BARBE, PH. and BRONIATOWSKI, M., Large deviation probability and the local dimension of sets, Technical Report, INRIA, 1995.

[5] BICKEL, P. J. and WICHURA, M. J., Convergence criteria for multiparameter stochastic processes and some applications, *Ann. Math. Statist.* **42** (1971), 1656–1670. *MR* **52** #4363

[6] BILLINGSLEY, P., *Convergence of probability measures*, Wiley, New York, 1968. *MR* **38** #1718

[7] BILLINGSLEY, P., *Probability and measure*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York, 1979. *MR* **80h**:60001

[8] BILLINGSLEY, P., *Probability and measure*, Wiley, New York, 1986. *MR* **87f**:60001

[9] BOROVKOV, A. A., Boundary value problems for random walks and large deviations in function spaces, *Teor. Verojatnost. i Primenen.* **12** (1967), 635–654. *MR* **39** #4906; *Theor. Probab. Appl.* **12** (1967), 575–595.

[10] BOURBAKI, N., *Topologie générale*, Ch. 5–10, CCLS diffusion, Paris, 1974.

[11] BOURBAKI, N., *Topologie générale*, Ch. 1–4, Masson, Paris, 1990.

[12] BRONIATOWSKI, M. and MASON, D. M., Extended large deviations, *J. Theoret. Probab.* **7** (1994), 647–666. *MR* **95h**:60037

[13] CSISZÁR, I., Sanov property, generalized I-projection and a conditional limit theorem, *Ann. Probab.* **12** (1984), 768–793. *MR* **86h**:60065

[14] CSÖRGŐ, M. and RÉVÉSZ, P., A new method to prove Strassen type laws of invariance principle, II, *Z. Wahrsch. Verw. Gebiete* **31** (1975), 261–269. *MR* **51** #11605a

[15] DEMBO, A. and ZAJIC, T., Large deviations: from empirical mean and measure to partial sums process, *Stochastic Process. Appl.* **57** (1995), 191–224. *MR* **96m**:60064

[16] DEMBO, A. and ZEITOUNI, O., *Large deviations techniques and applications*, Jones and Bartlett, Boston, 1993. *MR* **95a**:60034

[17] DEUSCHEL, J. D. and STROOCK, D. W., *Large deviations*, Pure and Applied Mathematics, 137, Academic Press, Boston, 1989. *MR* **90h**:60026

[18] DONSKER, M. D. and VARADHAN, S. R. S., Asymptotic evaluation of certain Markov process expectations for large time, I, *Comm. Pure Appl. Math.* **28** (1975), 1–47. *MR* **52** #6883

[19] GROENEBOOM, P., OOSTERHOFF, J. and RUYMGAART, F., Large deviation theorems for empirical probability measures, *Ann. Probab.* **7** (1979), 553–586. *MR* **80i**:60030

[20] KIEFER, J., Skorohod embedding of multivariate rv's and the sample df, *Z. Wahrsch. Verw. Gebiete* **24** (1972), 1–35. *MR* **49** #6382

[21] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent rv's and the sample df, I, *Z. Wahrsch. Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[22] LYNCH, J. and SETHURAMAN, J., Large deviation for processes with independent increments, *Ann. Probab.* **15** (1987), 610–627. *MR* **88m**:60076

[23] MCBRIDE, J., Functional analogues of iterated logarithm type laws for empirical distribution function whose arguments tend to 0 at an intermediate state, Ph. Thesis, Chicago, Illinois, 1974.

[24] MAJOR, P., Asymptotic distributions for weighted U-statistics, *Ann. Probab.* **22** (1994), 1514–1535. *MR* **96f**:60035

[25] O'NEIL, K. A. and REDNER, R. A., Asymptotic distributions of weighted U-statistics of degree two, *Ann. Probab.* **21** (1993), 1159–1169. *MR* **94h**:62027

[26] RUDIN, W., *Real and complex analysis*, McGraw-Hill Book Company, New York–Toronto–London, 1966, *MR* **35** #1420

[27] VARADHAN, S. R. S., Asymptotic probabilities and differential equations, *Comm. Pure Appl. Math.* **19** (1966), 261–286. *MR* **34** #3083

[28] VARADHAN, S. R. S., *Large deviations and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, 46, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984. *MR* **86h**:60067b

(Received August 11, 1995)

LABORATOIRE DE STATISTIQUE ET PROBABILITÉS
UNIVERSITÉ PAUL SABATIER
F–31062 TOULOUSE
FRANCE

barbe@cict.fr

DÉPARTEMENT DES SCIENCES ÉCONOMIQUES
UNIVERSITÉ DE REIMS ET CNRS
57 AVENUE TAITTINGER
F–51100 REIMS
FRANCE

mbr@ccr.jussieu.fr

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 419–431

POLYNOMIAL WAVELETS AND WAVELET PACKET BASES

T. KILGORE, J. PRESTIN and K. SELIG

Abstract

We discuss wavelet-oriented ideas to construct bases of algebraic polynomials. In particular, the splitting in the frequency domain is extended in order to define wavelet packets.

1. Introduction

We show here how algebraic polynomials on the interval $[-1,1]$ can be treated as wavelets and can be handled by wavelet techniques. Benefits include the potential for computational efficiency and accuracy in applications, for example to approximation problems. Theoretical developments also follow from systematic development and exploitation of orthogonality and from a generalization of the concept of shift invariance, which allow the application of the wavelet techniques on the interval. We will give here some of the basic ideas and techniques used in the wavelet approach to polynomials, which is also related to an important application, the construction of a series of mutually orthogonal polynomials of "optimal degree."

As the idea of wavelets originated in connection with signal analysis, let us look first at the original setting. Signal analysis naturally involves a "time domain" and a "frequency domain". One splits the frequency domain dyadically into *wavelet spaces*, with dilations and translations of a single function (mother wavelet) employed systematically to construct bases for these spaces. *Wavelet packet spaces* are subspaces which in turn further split the wavelet spaces, using smaller frequency ranges. In signal analysis, a function (signal) is "time-localized" if it is relatively large in magnitude at a certain "time" and relatively small otherwise. A "frequency-localized" function on the other hand is more or less of a single frequency. In a manifestation of Heisenberg's uncertainty principle, perfect time localization and perfect frequency localization are mutually incompatible. Thus, one goal in signal analysis is the construction of "time-frequency-localized" bases, involving a balanced consideration of both domains. In a wavelet treatment of polynomials on $[-1,1]$,

1991 *Mathematics Subject Classification.* Primary 41A10; Secondary 33C25, 42C15.
Key words and phrases. Wavelet packets, Chebyshev polynomials, orthogonal bases, generalized Chebyshev shift.

the time domain clearly should correspond to the underlying interval $[-1,1]$, while the frequency domain should correspond more or less to the degree of involved monomials. More precise statements and adaptations of this and of other concepts require more systematic treatment.

Wavelet techniques for polynomials on the interval $[-1,1]$ with respect to the Chebyshev weight have been developed in Kilgore and Prestin [4], in Tasche [12] and in Plonka, Selig, and Tasche [7], where the generalized Chebyshev shift was discussed and applied to the development of wavelets on the interval. An adaptation of the uncertainty principle can be found in Rösler and Voit [10], which in turn could be applied to wavelets on $[-1,1]$ analogous to Narcowich and Ward [6] and to Selig [11].

More recently, in Kilgore, Prestin, and Selig [5], wavelet techniques have been used to show the existence of and to perform the construction of an orthogonal Schauder basis of polynomials of optimal degree for the space $C[-1,1]$, where optimal degree signifies that the nth polynomial in the basis is always of degree less than $n(1+\epsilon)$, for previously given $\epsilon>0$. Here, the use of wavelet packets is precisely what is needed to construct a polynomial basis in which the degree of the polynomials grows within the prescribed limitations; as ϵ decreases, the dimension of the packet spaces decreases, and the number of packet spaces into which a given wavelet space must be split increases. This basis problem has a long history which is discussed in further detail in the paper [5].

Here, we will construct different wavelet bases and wavelet packet bases on the interval. At first, we will define polynomial subspaces by means of bases with the most frequency localization. Then, the idea of time-frequency-localized bases will be realized by building finite linear combinations in order to obtain wavelets and wavelet packets as generalized translates within each subspace.

The wavelet spaces as well as the wavelet packet spaces will be *orthogonal*; their orthogonality is given with respect to the weighted inner product

$$\langle f,g\rangle=\frac{2}{\pi}\int_{-1}^{1} f(x)g(x)\frac{dx}{\sqrt{1-x^2}}\,.$$

Hence, we use the orthogonal Chebyshev polynomials $T_n(x)=\cos n\arccos x$ $(n\in I\!N_0)$ for which

$$(1)\qquad \langle T_n,T_m\rangle=\begin{cases}2 & \text{for } n=m=0,\\ 1 & \text{for } n=m>0,\\ 0 & \text{otherwise.}\end{cases}$$

We will directly and explicitly describe the algebraic polynomials used in our wavelet and wavelet packet bases by giving their Chebyshev expansions. Our examples are related to the trigonometric Dirichlet kernel and the de la Vallée Poussin kernels and corresponding shift-invariant spaces (see e.g. Privalov [9], and Prestin and Selig [8]).

Note that in our construction the Chebyshev polynomials can be replaced by other polynomial systems orthonormal with respect to an arbitrary weight function w which yield corresponding bases for L^2_w. Having similar frequency localization in terms of the involved orthonormal polynomials, the resulting polynomials will differ in their properties of localization on $[-1,1]$ according to the weight w which will be related to different approximation properties of the bases. For the construction of the wavelets orthogonal with respect to an arbitrary weighted inner product we refer to Depczynski and Jetter [1, 2] and Fischer and Prestin [3]. However, results for the wavelet packets and the uncertainty principles are still in progress. Therefore, we restrict ourselves here to the Chebyshev weight and Chebyshev polynomials.

2. Wavelets and wavelet packets on $[-1,1]$

Let $N, M \in I\!N$ be fixed, with $N = 2^{\eta+1}M$ for some $\eta \in I\!N, \eta \geqq 2$. Furthermore, let us introduce, for any $l = 0, \ldots, 2^\eta - 2$, real coefficients

$$a^l_M(k), \quad (k = -M, \ldots, M) \quad \text{and} \quad a^0_{2M}(k), \quad (k = -2M, \ldots, 2M).$$

With any fixed set of such coefficients, we define the following spaces of polynomials

$$\begin{aligned}
V^M_N &:= \text{span } (\{T_k : k = 0, \ldots, N - M\} \\
&\qquad \cup \{a^0_M(k - M)T_{N-M+k} + a^0_M(M - k)T_{N+M-k} : k = 1, \ldots, M\}), \\
W^M_N &:= \text{span } (\{a^0_M(-k)T_{N+k} - a^0_M(k)T_{N-k} : k = 1, \ldots, M - 1\} \\
&\qquad \cup \{T_k : k = N + M, \ldots, 2N - 2M\} \\
&\qquad \cup \{a^0_{2M}(k-2M)T_{2N-2M+k} + a^0_{2M}(2M-k)T_{2N+2M-k} : k = 1, \ldots, 2M\}), \\
W^M_{N,l} &:= \text{span}(\{a^{l-1}_M(-k)T_{N+2M(l-1)+k} - a^{l-1}_M(k)T_{N+2M(l-1)-k} : k = 1, \ldots, M-1\} \\
&\qquad \cup \{a^l_M(k - M)T_{N+M(2l-1)+k} + a^l_M(M - k)T_{N+M(2l+1)-k} : k = 0, \ldots, M\}),
\end{aligned}$$

for $l = 1, \ldots, 2^\eta - 2$, and

$$\begin{aligned}
&W^M_{N,2^\eta-1} := \\
&\text{span}(\{a^{2^\eta-2}_M(-k)T_{2N-4M+k} - a^{2^\eta-2}_M(k)T_{2N-4M-k} : k = 1, \ldots, M - 1\} \\
&\cup \{T_k : k = 2N - 3M, \ldots, 2N - 2M\} \\
&\cup \{a^0_{2M}(k - 2M)T_{2N-2M+k} + a^0_{2M}(2M - k)T_{2N+2M-k} : k = 1, \ldots, 2M\}).
\end{aligned}$$

Given a general scheme for constructing the coefficients a^l_M, it is then possible to double repeatedly the values of M and N together. This successive doubling gives a nested sequence of spaces V^M_N, a corresponding sequence of spaces W_N^M, and inside of each space W_N^M a set of subspaces $W_{N,1}^M, \dots, W_{N,2^\eta-1}^M$.

Three relevant examples for the choice of the coefficients are the following, where for all $l=1,\dots,2^\eta-2$

(a) $$a_M^l(k) = \begin{cases} 1, & -M \leqq k \leqq 0, \\ 0, & 0 < k \leqq M, \end{cases}$$

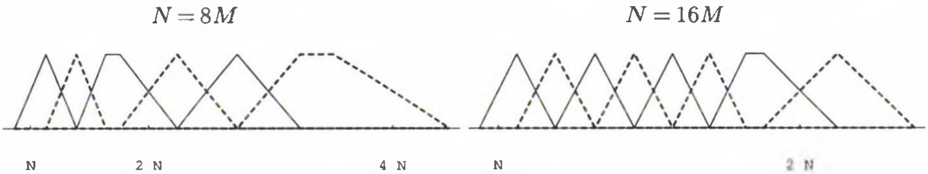
(b) $$a_M^l(k) = \frac{M-k}{2M}, \qquad -M \leqq k \leqq M,$$

(c) $$a_M^l(k) = \frac{M-k}{\sqrt{2M^2+2k^2}}, \qquad -M \leqq k \leqq M,$$

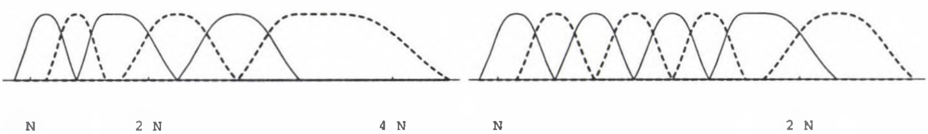
and $a_{2M}^0(k) = a_M^0(k/2)$ for all $-2M \leqq k \leqq 2M$. Arisen from their trigonometric analogs, example (a) yields functions related to the Dirichlet kernel whereas examples (b) and (c) come from de la Vallée Poussin means and from an orthogonalization procedure applied to translates thereof, respectively.

For the sake of good time localization for the wavelet and wavelet packet bases to be constructed we generally suggest that the coefficients $a_M^l(k)$ should decrease monotonically with increasing k and should be normalized such that $a_M^l(-M)=1$.

Based on the examples (b) and (c), the following graphs represent the size of the coefficients a_M^l with respect to their distribution in the frequency domain (see also the definition of the wavelet packet functions on p. 425) and thus illustrate some of the many possibilities for constructing the spaces $W_{N,l}^M$, for $l=1,\dots,2^\eta-1$ and $\eta=2$ (left) and $\eta=3$ (right). The graphs for $\eta=2$ also depict one doubling of N and M.



Example (b)



Example (c)

Now we study the spaces defined above and show that under certain conditions on the coefficients they span

$$V_{2N}^{2M} := \operatorname{span}(\{T_k : k=0,\dots,2N-2M\} \\ \cup \{a_{2M}^0(k-2M)T_{2N-2M+k} + a_{2M}^0(2M-k)T_{2N+2M-k} : k=1,\dots,2M\}).$$

THEOREM 2.1. *For any real coefficients* $a_M^l(k)$ *and* $a_{2M}^0(k)$, *it holds that*

$$V_N^M \cup W_N^M \subset V_{2N}^{2M}, \qquad V_N^M \perp W_N^M,$$

and

$$W_{N,l}^M \subset W_N^M, \qquad \text{for } l = 1, \ldots, 2^\eta - 1.$$

If

$$a_M^l(M) = 0, \qquad \text{for } l = 1, \ldots, 2^\eta - 2, \tag{2}$$

then

$$W_{N,l_1}^M \perp W_{N,l_2}^M, \qquad \text{for } 1 \leqq l_1 < l_2 \leqq 2^\eta - 1,$$

and if moreover the coefficients satisfy for all $l = 0, \ldots, 2^\eta - 2$

$$\begin{aligned} (a_M^l(k))^2 + (a_M^l(-k))^2 &> 0, && \text{for all } k = 0, \ldots, M, \\ (a_{2M}^0(k))^2 + (a_{2M}^0(-k))^2 &> 0, && \text{for all } k = 0, \ldots, 2M-1, \end{aligned} \tag{3}$$

then we have

$$V_{2N}^{2M} = V_N^M \oplus W_N^M \tag{4}$$

and

$$W_N^M = \bigoplus_{l=1}^{2^\eta - 1} W_{N,l}^M. \tag{5}$$

PROOF. The inclusions $V_N^M \cup W_N^M \subset V_{2N}^{2M}$ and $W_{N,l}^M \subset W_N^M$ follow directly from the definition of the spaces.

Using (1) and (2) the orthogonality $V_N^M \perp W_N^M$ and $W_{N,l_1}^M \perp W_{N,l_2}^M$ can be easily checked. In particular, for $l = l_1 = l_2 - 1$ we obtain for any $k = 0, \ldots, M-1$

$$\begin{aligned} &\langle a_M^l(-k) T_{N+2Ml-k} + a_M^l(k) T_{N+2Ml+k}\,, a_M^l(k) T_{N+2Ml-k} - a_M^l(-k) T_{N+2Ml+k} \rangle \\ &\quad = a_M^l(-k)\, a_M^l(k)\, (\langle T_{N+2Ml-k}\,, T_{N+2Ml-k} \rangle - \langle T_{N+2Ml+k}\,, T_{N+2Ml+k} \rangle) \\ &\quad = 0. \end{aligned}$$

For $|l_1 - l_2| > 1$ the orthogonality $W_{N,l_1}^M \perp W_{N,l_2}^M$ is evident.

Let us now prove (4) and (5). From (3) it follows that both of $a_M^l(-k)$ and $a_M^l(k)$ cannot vanish. Hence,

$$\begin{aligned} &\dim \operatorname{span}\, \{a_M^l(-k) T_{N+2Ml+k} - a_M^l(k) T_{N+2Ml-k} : \\ &\qquad\qquad k = 1, \ldots, M-1\} = M - 1, \\ &\dim \operatorname{span}\, \{a_M^l(k-M) T_{N+M(2l-1)+k} + a_M^l(M-k) T_{N+M(2l+1)-k} : \end{aligned}$$

$$k=0,\dots,M\}=M+1\,,$$
$$\dim\operatorname{span}\,\{a^0_{2M}(k-2M)T_{2N-2M+k}+a^0_{2M}(2M-k)T_{2N+2M-k}:\\ k=1,\dots,2M\}=2M\,.$$

Then, for the dimensions of the spaces we obtain

$$\dim V^M_N=N+1\,,\qquad \dim W^M_N=N\,,\qquad \dim W^M_{N,l}=2M\,,$$

for $l=1,\dots,2^\eta-2$, and

$$\dim W^M_{N,2^\eta-1}=4M\,.$$

Hence

$$\dim V^{2M}_{2N}=\dim V^M_N+\dim W^M_N$$

and

$$\dim W^M_N=\sum_{l=1}^{2^\eta-1}\dim W^M_{N,l}\,.$$

Together with the imbedding and orthogonality relations this proves the assertion. □

Following [7] one can define scaling functions and wavelets in terms of Chebyshev polynomials as generalized Chebyshev shifts of one function.

We define scaling functions for $s=0,\dots,N$, by

$$\phi^M_{N,s}:=\frac{1}{2}T_0+\sum_{k=1}^{N-M}\cos\frac{ks\pi}{N}T_k+\sum_{k=N-M+1}^{N+M-1}a^0_M(k-N)\cos\frac{ks\pi}{N}T_k\,,$$

and wavelets, for $s=1,\dots,N$, by

$$\psi^M_{N,s}:=\sum_{k=N-M+1}^{N+M-1}a^0_M(N-k)\cos\frac{k(2s-1)\pi}{2N}T_k+\sum_{k=N+M}^{2N-2M}\cos\frac{k(2s-1)\pi}{2N}T_k\\ +\sum_{k=2N-2M+1}^{2N+2M-1}a^0_{2M}(k-2N)\cos\frac{k(2s-1)\pi}{2N}T_k\,.$$

In this paper we introduce corresponding wavelet packet functions, for $p=$

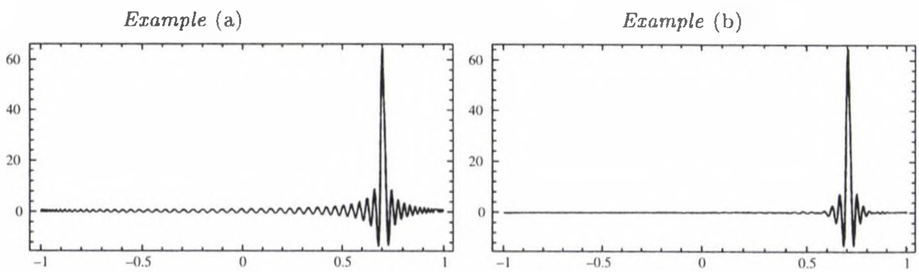
$1, \ldots, 2^{\eta-1} - 1$ and $s = 1, \ldots, 2M$, by

$$\psi^M_{N,2p-1,s} := \sum_{k=N+(4p-5)M+1}^{N+(4p-3)M-1} a_M^{2p-2}(N + (4p-4)M - k)\, \sin \frac{k(2s-1)\pi}{4M}\, T_k$$
$$+ \sum_{k=N+(4p-3)M}^{N+(4p-1)M-1} a_M^{2p-1}(k - N - (4p-2)M)\, \sin \frac{k(2s-1)\pi}{4M}\, T_k,$$
$$\psi^M_{N,2p,s} := \sum_{k=N+(4p-3)M+1}^{N+(4p-1)M-1} a_M^{2p-1}(N + (4p-2)M - k)\, \cos \frac{k(2s-1)\pi}{4M}\, T_k$$
$$+ \sum_{k=N+(4p-1)M}^{N+(4p+1)M-1} a_M^{2p}(k - N - 4pM)\, \cos \frac{k(2s-1)\pi}{4M}\, T_k\,,$$

and for $s = 1, \ldots, 4M$, by

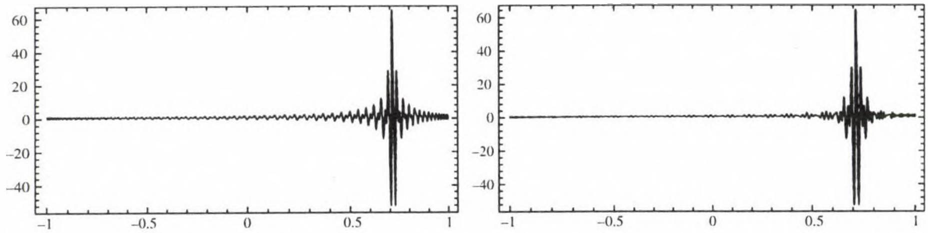
$$\psi^M_{N,2^\eta-1,s} := \sum_{k=2N-5M+1}^{2N-3M-1} a_M^{2^\eta-2}(k - 2N - 4M)\, \cos \frac{k(2s-1)\pi}{8M}\, T_k$$
$$+ \sum_{k=2N-3M}^{2N-2M} \cos \frac{k(2s-1)\pi}{8M}\, T_k$$
$$+ \sum_{k=2N-2M+1}^{2N+2M-1} a_{2M}^0(k - 2N)\, \cos \frac{k(2s-1)\pi}{8M}\, T_{2N-k}\,.$$

In order to illustrate their time localization, we have drawn corresponding functions for the coefficients from the examples (a) and (b) for $N = 128$ and $M = 16$. The corresponding functions for the example (c) are not shown; they would be quite similar to those for (b).

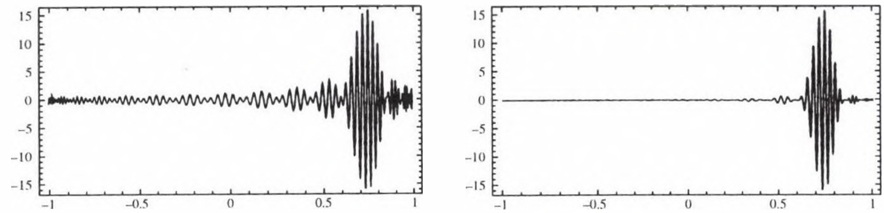


Scaling functions $\phi^{16}_{128,32}$

Example (a) *Example* (b)



Wavelets $\psi^{16}_{128,32}$



Wavelet packet functions $\psi^{16}_{128,1,8}$

We can show that the functions defined above build bases of the previously defined spaces.

THEOREM 2.2. *For* $N, M \in \mathbb{N}$, *with* $N = 2^{\eta+1}M$, *we have*

$$V_N^M = \text{span}\ \{\phi_{N,s}^M : s = 0, \ldots, N\},$$
$$W_N^M = \text{span}\ \{\psi_{N,s}^M : s = 1, \ldots, N\},$$
$$W_{N,l}^M = \text{span}\ \{\psi_{N,l,s}^M : s = 1, \ldots, 2M\},$$

for $l = 1, \ldots, 2^\eta - 2$, *and*

$$W_{N,2^\eta-1}^M = \text{span}\ \{\psi_{N,2^\eta-1,s}^M : s = 1, \ldots, 4M\}.$$

PROOF. Let us write the basis used in the definition of V_N^M in the order given there as a column vector $\underline{\mathbf{v}}_N^M$, that is,

$$\underline{\mathbf{v}}_N^M = \Big(T_0, \ldots, T_{N-M}, a_M^0(1-M)T_{N-M+1} + a_M^0(M-1)T_{N+M-1},$$
$$\ldots, a_M^0(-1)T_{N-1} + a_M^0(1)T_{N+1},\ 2a_M^0(0)T_N\Big)^T.$$

Let us similarly represent the given basis of W_N^M as a column vector $\underline{\mathbf{w}}_N^M$ and the given basis of $W_{N,l}^M$ as a column vector $\underline{\mathbf{w}}_{N,l}^M$.

Now we can write

$$\left(\phi_{N,s}^M\right)_{s=0}^N = \mathbf{A}_N\, \underline{\mathbf{v}}_N^M\,, \qquad \left(\psi_{N,s}^M\right)_{s=1}^N = \mathbf{B}_N\, \underline{\mathbf{w}}_N^M$$

where

$$\mathbf{A}_N = \left(\frac{2-\delta_{k,0}-\delta_{k,N}}{2} \cos \frac{ks\pi}{N} \right)_{s,k=0}^{N},$$

$$\mathbf{B}_N = \left(\frac{2-\delta_{k,2N}}{2} \cos \frac{k(2s-1)\pi}{2N} \right)_{s=1,k=N+1}^{N,2N}.$$

Also, for $l=1,\ldots,2^\eta-2$, we have

$$\left(\psi_{N,l,s}^M\right)_{s=1}^{2M} = \mathbf{C}_{N,l}^M\,\underline{\mathbf{w}}_{N,l}^M\,, \qquad \left(\psi_{N,2^\eta-1,s}^M\right)_{s=1}^{4M} = \mathbf{C}_{N,2^\eta-1}^M\,\underline{\mathbf{w}}_{N,2^\eta-1}^M\,,$$

where, for $p=1,\ldots,2^{\eta-1}-1$,

$$\mathbf{C}_{N,2p}^M = \left(\frac{2-\delta_{k,N+4pM}}{2} \cos \frac{k(2s-1)\pi}{4M} \right)_{s=1,k=N+(4p-2)M+1}^{2M,N+4pM},$$

$$\mathbf{C}_{N,2p-1}^M = \left(\frac{2-\delta_{k,N+(4p-2)M}}{2} \sin \frac{k(2s-1)\pi}{4M} \right)_{s=1,k=N+(4p-4)M+1}^{2M,N+(4p-2)M},$$

and

$$\mathbf{C}_{N,2^\eta-1}^M = \left(\frac{2-\delta_{k,2N}}{2} \cos \frac{k(2s-1)\pi}{8M} \right)_{s=1,k=2N-4M+1}^{4M,2N}.$$

The proof of the theorem is now completed by noting that the regularity of these matrices is well-known and follows directly from (see Tasche [12])

$$\mathbf{A}_N\mathbf{A}_N^T = \left(\frac{N}{2}\delta_{s,k} \right)_{s,k=0}^{N},$$

$$\mathbf{B}_N^T\mathbf{B}_N = \left(\frac{N}{2}\delta_{s,k}(2-\delta_{k,N}) \right)_{s,k=1}^{N}$$

and

$$(\mathbf{C}_{N,2p}^M)^T\mathbf{C}_{N,2p}^M = (\mathbf{C}_{N,2p-1}^M)^T\mathbf{C}_{N,2p-1}^M = \left(M\delta_{s,k}(2-\delta_{k,2M})\right)_{s,k=1}^{2M}. \qquad \square$$

Note that in the above proof the transformation matrices between corresponding bases of the scaling function spaces, wavelet spaces, and wavelet packet spaces are given. The transformation from one basis to another can also be carried out by use of fast algorithms (cf. [12, 7]).

3. Orthogonal bases

Here we further impose orthogonality of the bases given in Theorem 2.2. It turns out to be guaranteed by a certain condition on the coefficients a_M^l.

THEOREM 3.1. *If*

$$a_M^l(M) = 0\,, \qquad \textit{for } l = 1, \ldots, 2^\eta - 2\,, \tag{6}$$

and

$$\begin{aligned} (a_M^l(k))^2 + (a_M^l(-k))^2 &= 1 \qquad \textit{for } k = 0, \ldots, M\,, \quad l = 0, \ldots, 2^\eta - 2\,, \\ (a_{2M}^0(k))^2 + (a_{2M}^0(-k))^2 &= 1 \qquad \textit{for } k = 0, \ldots, 2M-1\,, \end{aligned} \tag{7}$$

then we have the orthogonality properties

$$\langle \phi_{N,r}^M\,, \phi_{N,s}^M \rangle = N\delta_{r,s} \frac{1 + \delta_{s,0} + \delta_{s,N}}{2}\,, \qquad \textit{for } r, s = 0, \ldots, N\,, \tag{8}$$

$$\langle \psi_{N,r}^M\,, \psi_{N,s}^M \rangle = N\delta_{r,s}\,, \qquad \textit{for } r, s = 1, \ldots, N\,, \tag{9}$$

$$\langle \psi_{N,l,r}^M\,, \psi_{N,l,s}^M \rangle = M\delta_{r,s} \qquad \textit{for all } l = 1, \ldots, 2^\eta - 1 \textit{ and } r, s = 1, \ldots, 2M\,, \tag{10}$$

$$\langle \psi_{N,2^\eta-1,r}^M\,, \psi_{N,2^\eta-1,s}^M \rangle = 2M\delta_{r,s} \qquad \textit{for } r, s = 1, \ldots, 4M\,. \tag{11}$$

Notice the connection between the conditions (3) giving linear independence and (7) giving orthogonality.

PROOF. For the proof, we will use the orthogonality properties (1) of the Chebyshev polynomials T_k. In order to show (8), we note that

$$\begin{aligned} \langle \phi_{N,r}^M\,, \phi_{N,s}^M \rangle &= \frac{1}{2} + \sum_{k=1}^{N-M} \cos\frac{kr\pi}{N} \cos\frac{ks\pi}{N} \\ &\quad + \sum_{k=-M+1}^{M-1} \frac{(2M-k)^2}{2M^2 + 2(M-k)^2} \cos\frac{(N-k)r\pi}{N} \cos\frac{(N-k)s\pi}{N} \\ &= \frac{1 + (-1)^{r-s}}{2} + \sum_{k=1}^{N-1} \cos\frac{kr\pi}{N} \cos\frac{ks\pi}{N} \\ &= \frac{2 + (-1)^{r-s} + (-1)^{r+s}}{4} + \frac{1}{2} \sum_{k=1}^{N-1} \left(\cos\frac{k(r-s)\pi}{N} + \cos\frac{k(r+s)\pi}{N} \right) \end{aligned}$$

$$= N\delta_{r,s}\frac{1+\delta_{s,0}+\delta_{s,N}}{2},$$

where we used that

$$\frac{1}{2}+\frac{(-1)^r}{2}+\sum_{k=1}^{N-1}\cos\frac{kr\pi}{N}=N\delta_{r,0\,\mathrm{mod}\,2N}\,.$$

The proof of (9)–(11) follows the same ideas. □

The conditions (6)–(7) hold for our example (c). For this special case, the functions are

$$\phi^M_{N,s}=\frac{1}{2}T_0+\sum_{k=1}^{N-M}\cos\frac{ks\pi}{N}T_k+\sum_{k=1}^{2M-1}\frac{2M-k}{\sqrt{2M^2+2(M-k)^2}}\cos\frac{(N-M+k)s\pi}{N} \times T_{N-M+k},$$

$$\begin{aligned}\psi^M_{N,s}=&\sum_{k=1}^{2M-1}\frac{k}{\sqrt{2M^2+2(M-k)^2}}\cos\frac{(N-M+k)(2s-1)\pi}{2N}T_{N-M+k}\\&+\sum_{k=N+M}^{2N-2M}\cos\frac{k(2s-1)\pi}{2N}T_k\\&+\sum_{k=1}^{4M-1}\frac{4M-k}{\sqrt{8M^2+2(k-2M)^2}}\cos\frac{(2N-2M+k)(2s-1)\pi}{2N}T_{2N-2M+k},\end{aligned}$$

$$\psi^M_{N,2p,s}=\sum_{k=-2M+1}^{2M-1}\frac{2M-|k|}{\sqrt{2M^2+2(M-|k|)^2}}\cos\left((k-M)\frac{(2s-1)\pi}{4M}\right)\times T_{N+(4p-1)M+k},$$

$$\psi^M_{N,2p-1,s}=\sum_{k=-2M+1}^{2M-1}\frac{2M-|k|}{\sqrt{2M^2+2(M-|k|)^2}}\sin\left((k-3M)\frac{(2s-1)\pi}{4M}\right)\times T_{N+(4p-3)M+k},$$

and

$$\begin{aligned}\psi^M_{N,2^\eta-1,s}=&\sum_{k=1}^{2M-1}\frac{k}{\sqrt{2M^2+2(M-k)^2}}\cos\frac{k(2s-1)\pi}{8M}T_k\\&+\sum_{k=2N-3M}^{2N-2M}\cos\frac{k(2s-1)\pi}{8M}T_k\end{aligned}$$

$$+\sum_{k=1}^{4M-1} \frac{4M-k}{\sqrt{8M^2+2(k-2M)^2}} \cos\frac{(2N-2M+k)(2s-1)\pi}{8M} \times T_{2N-2M+k}.$$

With appropriate choices of N and M given by successive doubling of certain initial N and M, the pairwise orthogonal wavelet packet functions just described can be used to define an orthogonal Schauder basis for $C[-1,1]$ consisting of polynomials of optimal degree at most $n(1+\epsilon)$. For this construction, the initial values of N and M are determined by the given value of ϵ. We have mentioned this problem already in the introduction; the details are given in Kilgore, Prestin, and Selig [5].

REFERENCES

[1] Depczynski, U., Konstruktion von waveletartigen Zerlegungen auf kompakten Intervallen mit Hilfe der Eigenlösungen Sturm-Liouvillescher Randwertprobleme, Dissertation, Universität-Gesamthochschule Duisburg, Duisburg, 1995.

[2] Depczynski, U. and Jetter, K., Multiscale decompositions on the interval based on DCT-I transforms, *Advanced topics in multivariate approximation* (Montecatini, 1995), Ed. by F. Fontanella, K. Jetter and P.-J. Laurent, World Scientific, Singapore, 1996, 57–70.

[3] Fischer, B. and Prestin, J., Wavelets based on orthogonal polynomials, *Math. Comp.* (to appear).

[4] Kilgore, T. and Prestin, J., Polynomial wavelets on the interval, *Constr. Approx.* **12** (1996), 95–110. *MR* **97b**:41003

[5] Kilgore, T., Prestin, J. and Selig, K., Orthogonal algebraic polynomial Schauder bases of optimal degree, *J. Fourier Anal. Appl.* **2** (1996), 597–610.

[6] Narcowich, F. J. and Ward, J. D., Wavelets associated with periodic basis functions, *Appl. Comput. Harmon. Anal.* **3** (1996), 40–56. *MR* **96j**:42022

[7] Plonka, G., Selig, K. and Tasche, M., On the construction of wavelets on a bounded interval, *Adv. Comput. Math.* **4** (1995), 357–388. *MR* **96m**:42057

[8] Prestin, J. and Selig, K., Interpolatory and orthonormal trigonometric wavelets, *Signal and image representation in combined spaces*, Ed. by J. Zeevi and R. Coifman (to appear).

[9] Privalov, A. A., An orthogonal trigonometric basis, *Mat. Sb.* **182** (1991), 384–394 (in Russian). *MR* **92f**:42005; English translation: *Math. USSR Sbornik* **72** (2), 363–372.

[10] Rösler, M. and Voit, M., An uncertainty principle for ultraspherical expansions, *J. Math. Anal. Appl.* **209** (1997), 624–634.

[11] Selig, K., Trigonometric wavelets and the uncertainty principle, *Approximation theory* (Witten, 1995), Math. Research, Vol. 86, Akademie-Verlag, Berlin, 1995, 293–304. *MR* **97c**:42058

[12] Tasche, M., Polynomial wavelets on $[-1,1]$, *Approximation theory, wavelets and applications* (Maratea, 1994), Ed. by S. P. Singh, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 454, Kluwer Academic Publ., Dordrecht, 1995, 497–512. *MR* **96c**:42073

(Received September 26, 1995)

DEPARTMENT OF MATHEMATICS
AUBURN UNIVERSITY
AUBURN, AL 36849
U.S.A.

kilgota@mallard.duc.auburn.edu

INSTITUT FÜR BIOMATHEMATIK UND BIOMETRIE
GSF – FORSCHUNGSZENTRUM FÜR
UMWELT UND GESUNDHEIT, GmbH
INGOLSTÄDTER LANDSTRASSE 1
D–85764 NEUHERBERG
GERMANY

prestin@gsf.de

INSTITUT FÜR ANGEWANDTE MATHEMATIK UND STATISTIK
UNIVERSITÄT HOHENHEIM
D–70593 STUTTGART
GERMANY

selig@uni-hohenheim.de

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 433–440

ZUR GEOMETRIE DER TRIDENS-KURVEN DER ISOTROPEN EBENE

J. TÖLKE

Abstract

D. Palman [2] and H. Sachs [6] studied cubic curves in an isotropic plane with the property that the linear power of a point P with respect to a given curve is independent from the straight line through P. There are three classes of such curves. We show that the class of the Tridens-curves can be generated by special circumscribed tangent foursides of isotropic circles. So every Tridens-curve is a focal curve.

Die zirkulären Kurven 3. Ordnung spielen eine wichtige Rolle bei den Untersuchungen von H. Sachs [4] über oskulierende und hyperoskulierende Kegelschnittbüschel. Die Unterklasse der vollständig zirkulären Kurven 3. Ordnung wurde von D.Palman [2, 3] und H. Sachs [6] untersucht. Speziell wurde in [2] die Geometrie der sogenannten *Tridens-Kurven* behandelt.

Vorliegende Note gibt dazu Ergänzungen. Wir zeigen, daß jeder Tridens-Kurve ein zulässiges Sehnenvierseit eingeschrieben werden kann. Es ist zugleich Tangentenvierseit eines isotropen Kreises. Durch solche zulässigen Haupttangentenvierseite isotroper Kreise lassen sich alle Tridens-Kurven erfassen. Hiermit gelingt eine neue (elementare) Deutung für den Radius des asymptotischen Kreises der Tridens-Kurve. Erwähnt sei noch, daß jede Tridens-Kurve Fokalkurve der einem zulässigen Sehnenvierseit einbeschriebenen Kegelschnittschar ist.

1

Bezeichne $\{\xi, \eta\}$ affine Koordinaten in der isotropen Ebene I_2. Sind $(x : y : t)$ die zugehörigen homogenen Koordinaten, so wird die *absolute Gerade f* durch $t = 0$ und der *absolute Punkt F* durch $(0 : 1 : 0)$ beschrieben. Die zugrundeliegende Fundamentalgruppe ist die dreiparametrige Bewegungsgruppe B_3 [5]. Die Geometrie der *Tridens-Kurven* 3. Ordnung

$$\xi\eta = \alpha\xi^3 + \alpha_1\xi^2 + \alpha_2\xi + \alpha_3 \tag{1}$$

wurde von D. Palman [2] behandelt. Für die Bewegungsinvariante α – bzw. den Radius des asymptotischen Kreises von (1) – fand H. Sachs eine geometrische Deutung [6, S. 381]. Wir wollen eine weitere *elementare* Deutung

1991 *Mathematics Subject Classification.* Primary 51N25; Secondary 51N15.

Key words and phrases. Isotropic plane, Tridens curve, completely circular curves of third order, focal curve.

0081–6906/97/$ 5.00

aufzeigen. Sind $t_1, \dots, t_4$ vier Geraden von I_2, so sagen wir, daß die vier Geraden t_i ein *zulässiges Vierseit* bilden, wenn keine drei Geraden kopunktal sind, keine zwei Geraden parallel sind und keine Gerade isotrop ist.

LEMMA. *Jeder Tridens-Kurve läßt sich ein zulässiges Sehnenvierseit einbeschreiben.*

Das soll heißen, daß die drei Gegenpunktepaare des Vierseits Punkte der Tridens-Kurve sind.

BEWEIS. (a) Wir wählen eine Gerade, welche die Tridens-Kurve (1) in drei verschiedenen reellen Punkten schneidet, als neue ξ-Achse. Dann folgt mit den Schnittpunktsabszissen a_1, a_2, a_3

$$\xi\eta = \alpha(\xi - a_1)(\xi - a_2)(\xi - a_3). \tag{1'}$$

Sollte es i, j, k mit $\{i, j, k\} = \{1, 2, 3\}$ geben derart, daß $a_i + a_j = 0$, so drehen wir die ξ-Achse so, daß die neue Lage wieder drei reelle, verschiedene Schnittpunkte hat

$$\eta = p\alpha\bar{\xi} + \bar{\eta}, \quad \xi = \bar{\xi}, \quad p \neq 0.$$

Dann folgt

$$\bar{\xi}\bar{\eta} = \alpha[-p\bar{\xi}^2 + (\bar{\xi} - a_k)(\bar{\xi}^2 - a_i^2)].$$

Wären auch jetzt noch zwei Punkte $\bar{\xi} = \pm c, \bar{\eta} = 0$ Tridenskurvenpunkte, so müßte $c(c^2 - a_i^2) = 0$, wegen $p \neq 0$, also $c = 0$ sein. Also läßt sich (1) durch eine isotrope Bewegung auf die Form

$$\bar{\xi}\bar{\eta} = \alpha(\bar{\xi} - \bar{a})(\bar{\xi} - \bar{A})(\bar{\xi} - B) \tag{1''}$$

mit

$$(\bar{a} + \bar{A})(\bar{a} + B)(\bar{A} + B)(\bar{a} - \bar{A})(\bar{a} - B)(\bar{A} - B)B \neq 0$$

bringen, wobei o.B.d.A. $\bar{a} < \bar{A}$ sei. Die Schiebung $\xi = \bar{\xi} - B, \quad \eta = \bar{\eta}$ liefert schließlich

$$\eta(\xi + B) = \alpha\xi(\xi - a)(\xi - A) \tag{2}$$

mit

$$\begin{aligned} &a = \bar{a} - B, \quad A = \bar{A} - B, \quad aA(a - A) \neq 0, \quad a + A + 2B \neq 0, \\ &B \neq 0, \quad aA + 2B(a + A + 2B) = (\bar{a} + B)(\bar{A} + B) \neq 0, \quad a < A. \end{aligned} \tag{3}$$

(b) Wir betrachten eine Gerade $\eta = \alpha\rho\xi$ mit $\rho \neq 0$, für die

$$X^2 := (a + A + \rho)^2 - 4aA + 4B\rho \tag{4}$$

positiv ist und die keine Tangente von (2) im Punkte (0,0) ist. Diese Gerade schneidet (2) in den reellen, verschiedenen Punkten[1] $(i = 1, 2)$

$$\xi_i = 1/2\{(a + A + \rho) + (-1)^i X\}, \quad \eta_i = \alpha\rho\xi_i, \quad X := +\sqrt{X^2} \tag{5}$$

[1] Da F Doppelpunkt ist, gilt $(\xi_1 - a)(\xi_2 - A) \neq 0$.

und es gilt ($\xi_i \neq 0$)

$$aA - B\rho \neq 0. \tag{6}$$

Die Verbindungsgerade von $(a,0)$ mit (ξ_1, η_1) bzw. die von $(A,0)$ mit (ξ_2, η_2) schneidet die Tridens-Kurve (2) im weiteren Punkt (ξ_a, η_a) bzw. (ξ_A, η_A) und es gilt

$$\xi_a(\xi_1 - a) - \{\rho\xi_1 - (\xi_1 - a)(\xi_1 - A)\} = 0,$$

bzw.

$$\xi_A(\xi_2 - A) - \{\rho\xi_2 - (\xi_2 - a)(\xi_2 - A)\} = 0.$$

Wegen (5) ist $\xi_a = \xi_A$ äquivalent mit

$$(A - a - X)\{\rho^2 + 2\rho(a + A) - X^2 + (a - A)^2\} = 0,$$

was mit (4) auf

$$(A - a - X)B\rho = 0 \tag{7}$$

führt. Wegen $\rho B \neq 0$ folgt also unter den gemachten Voraussetzungen

$$\xi_a = \xi_A \text{ äquivalent } X = A - a. \tag{8}$$

Aus $X = A - a$ folgt mit (4)

$$2(a + A + 2B) + \rho = 0. \tag{9}$$

Setzen wir umgekehrt (9) voraus, so gilt

(i) Aus (3) folgt $\rho \neq 0$;

(ii) Aus (4) folgt $X^2 = (A - a)^2$, also mit (3) $X^2 > 0$;

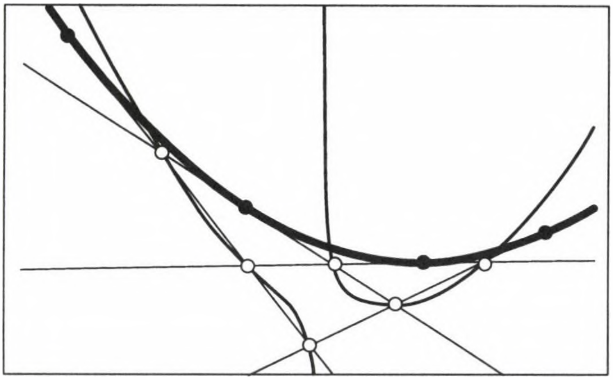
(iii) $aA - B\rho = aA + 2B(a + A + 2B)$, was nach (3) von Null verschieden ist, sodaß (6) gilt.

Damit sind die Voraussetzungen, die zu (8) führten, erfüllt. Mit (9) verifizieren wir $X = A - a$, d.h. $\xi_a = \xi_A$. □

2

Für ein zulässiges Sehnenvierseit einer Tridens-Kurve ist keine der Diagonalen isotrop. Allgemein nennen wir zulässige Vierseite, für die keine Diagonale isotrop ist, *zulässige Hauptvierseite*.

SATZ 1. *Ein zulässiges Hauptvierseit ist genau dann zulässiges Sehnenvierseit einer Tridens-Kurve, wenn es zulässiges Haupttangentenvierseit eines isotropen Kreises ist.*



Beweis. (a) Wahl des Koordinatensystems.
Sei $t_1, \ldots, t_4$ ein zulässiges Vierseit. Wir setzen

$$T_{ik} := t_i \wedge t_k = T_{ki}.$$

Damit ist z.B. $\Delta(T_{13}, T_{14}, T_{34})$ ein *zulässiges* Dreieck [5, S. 22]. Durch eine isotrope Bewegung können wir erreichen, daß die genannten Dreieckseckpunkte die Koordinaten $(A, B, \lambda \in \mathbb{R})$

(10) $\quad T_{13} = (0,0), \quad T_{14} = (A,0), \quad T_{34} = (\lambda B, B) \quad \text{mit} \quad \lambda(\lambda B - A)AB \neq 0$

bekommen. Damit folgt $(a, b \in \mathbb{R})$

(11) $\quad T_{12} = (a,0), \quad T_{23} = (\lambda b, b) \text{ mit } (a-A)(b-B)(bA-aB)(a-\lambda b)ab \neq 0,$

und es gilt für die Koordinaten von $T_{24} =: (\xi_t, \eta_t)$

(12) $$(aB - bA)\xi_t = \beta, \quad (aB - bA)\eta_t = bB(a - A),$$

wobei abkürzend gesetzt wurde

(13) $$\beta := -aA(b - B) + \lambda bB(a - A).$$

(b) Die durch T_{ik} bestimmte Tridens-Kurve.

Unter einer *vollständig zirkulären Kurve 3. Ordnung* versteht man eine Kurve 3. Ordnung $k^{(3)}$, für die F ein dreifacher Schnittpunkt von $k^{(3)}$ mit der absoluten Geraden f ist [2, 6]. Für die durch die Punkte T_{ik} bestimmte – eventuell reduzible – vollständig zirkuläre Kurve 3. Ordnung findet man nach einfacher, längerer Zwischenrechnung[2] die Darstellung

(14) $$\xi(\xi - a)(\xi - A) + \gamma\xi\eta - \lambda\delta\eta^2 + \lambda(\lambda^2 bB - aA)\eta = 0$$

[2] Man verwendet zweckmäßig zunächst das angepaßte affine Koordinatensystem $x = \xi - \lambda\eta$, $y = \eta$.

mit den Abkürzungen

(15) $$(aB - bA)\delta := (a - b\lambda)(B\lambda - A)\sigma, \quad \sigma := a - A - \lambda(b - B)$$

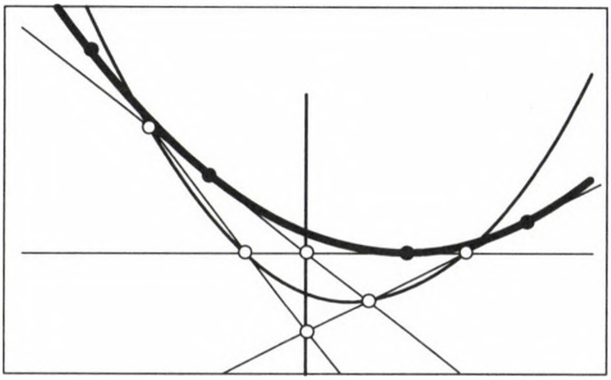
und

(16) $$\gamma - \delta := \lambda[a + A - \lambda(b + B)].$$

Damit ist $\delta = 0$ mit $\sigma = 0$ äquivalent. Wegen $aB - bA \neq 0$ und

(17) $$(b - B)(\gamma - \delta) = \lambda\{(b + B)\sigma - 2(aB - bA)\}$$

kann der Fall $\gamma = \delta = 0$ nicht eintreten.



Für $\sigma = 0$[3] zerfällt die Kubik (14) genau dann, wenn die isotrope Richtung zu einer der Diagonalen $T_{23} \vee T_{14}$, $T_{12} \vee T_{34}$, $T_{13} \vee T_{24}$ parallel ist. Sie zerfällt dann in die betreffende isotrope Diagonale und den isotropen Kreis durch die anderen beiden Gegenpunktepaare des zulässigen Vierseits. Also ist für $\sigma = 0$ die Kubik (14) genau für zulässige Hauptvierseite nicht zerfallend. Der absolute Punkt ist dann ein Doppelpunkt. Nach (14), (15) und (17) sind die absolute Gerade und die isotrope Gerade

(18) $$2(b - B)\xi + AB - ab = 0$$

die Doppelpunktstangenten. Die Kurve ist also eine *Tridens-Kurve* [2].

(c) Die geometrische Deutung von $\sigma = 0$.

Für die isotropen Linienkoordinaten $w : u : v \quad (w\eta = u\xi + v)$ der Geraden t_i unseres zulässigen Vierseits gilt gemäß Teil (a) im festgelegten Koordinatensystem

(19) $$t_1:\ u = v = 0, \quad t_2:\ w = a - \lambda b,\ u = -b, v = ab,$$
$$t_3:\ w = \lambda,\ u = 1,\ v = 0, \quad t_4:\ w = A - \lambda B,\ u = -B,\ v = AB.$$

[3] Für den Fall $\sigma \neq 0$ sei auf [8] verwiesen.

Damit folgt die Darstellung der dem zulässigen Vierseit $t_1, \ldots, t_4$ *einbeschriebenen Parabel* zu

$$\lambda(aB-bA)u^2-(aB-bA)uw+\sigma uv+(b-B)vw=0. \tag{20}$$

Für ihren Berührpunkt mit der absoluten Geraden ($u=w=0$) gilt daher

$$(b-B)w+\sigma u=0.$$

Somit gilt genau dann $\sigma=0$, wenn die dem zulässigen Vierseit einbeschriebene Parabel ein isotroper Kreis ist. □

3

H. Sachs hat in [6, S. 381] eine geometrische Deutung des Koeffizienten α der Tridens-Kurve (1) angegeben. Wir wollen eine weitere, *elementargeometrische* Deutung für α zeigen. Nach D. Palman [2, S. 39] gilt für den Radius R des *asymptotischen Kreises* $R=\alpha$.

SATZ 2. *Sei $t_1, \ldots, t_4$ ein zulässiges Sehnenvierseit einer Tridens-Kurve und bezeichne $T_{ik}:=t_i \wedge t_k=T_{ki}$ die Schnittpunkte der Sehnen t_j. Dann gilt für den Radius R des asymptotischen Kreises*

$$R=\frac{\sphericalangle(t_i,t_j)}{d(T_{ik},T_{jk})+d(T_{il},T_{jl})} \tag{21}$$

mit $\{i,j,k,l\}=\{1,2,3,4\}$. Dabei bezeichne d bzw. $\sphericalangle$ den isotropen Abstand bzw. den isotropen Winkel.

BEWEIS. Wir zeigen etwa die Fälle $i=4, j=2$ und $i=3, j=1$. Im Koordinatensystem von Abschnitt 2 gilt nach (12)

$$-R^{-1}=\gamma=\frac{2\lambda(bA-aB)}{b-B}=\lambda[a+A-\lambda(b+B)],$$

wobei λ vermöge (15) und $\sigma=0$ definiert ist. Mit (10) und (11) folgt

$$d(T_{41},T_{21})+d(T_{43},T_{23})=2\lambda(b-B),$$
$$d(T_{32},T_{12})+d(T_{34},T_{14})=a+A-\lambda(b+B).$$

Nach (19) gilt

$$\sphericalangle(t_2,t_4)=\frac{bA-aB}{(a-\lambda b)(A-\lambda B)}=\frac{(b-B)^2}{bA-aB},\quad \sphericalangle(t_1,t_3)=\frac{1}{\lambda}.$$ □

BEMERKUNG. Nach (20) und [5, S. 27] gilt für den Radius R^* des dem zulässigen Sehnenvierseit einbeschriebenen isotropen Kreises

$$R^*=R/2,$$

was eine weitere geometrische Deutung des asymptotischen Kreisradius ergibt.

4

Um eine weitere Eigenschaft der Tridens-Kurven aufzudecken, betrachten wir die einem zulässigen Vierseit $t_1, \ldots, t_4$ einbeschriebene Kegelschnittschar. Im Koordinatensystem von Abschnitt 2 folgt mit einem Scharparameter κ nach (20) und (12)

$$\begin{aligned}\lambda(aB-bA)u^2+(\sigma-\kappa\beta)uv-\kappa(aB-bA)v^2-(aB-bA)uw\\ +[b-B+\kappa bB(a-A)]vw=0.\end{aligned} \tag{22}$$

Für die *isotropen Hauptachsen* [7, S. 389] $w:u:v$ gilt demnach das Gleichungssystem

$$2\lambda(aB-bA)u+(\sigma-\kappa\beta)v-(aB-bA)w=0$$

$$(\sigma-\kappa\beta)u-2\kappa(aB-bA)v+[b-B+\kappa bB(a-A)]w=0.$$

Somit hüllen die isotropen Hauptachsen den Kegelschnitt

$$\begin{gathered}(aB-bA)\{2\lambda\beta u^2+4\lambda(aB-bA)uv+2\sigma v^2\\ -[2\lambda bB(a-A)+\beta]uw+bB(a-A)w^2\}\\ -\{\beta(b-B)+\sigma bB(a-A)+2(aB-bA)^2\}vw=0\end{gathered}$$

ein. Genau für $\sigma=0$ ist die Enveloppe der isotropen Hauptachsen somit eine Parabel P. Für ihre isotropen Tangenten gilt

$$u\{\beta u+2(aB-bA)v\}=0, \quad w=0.$$

Wegen (13) und (15) ist also der eigentliche isotrope Brennstrahl von P nach (18) die Asymptote der Tridens-Kurve (14).

Nach M. Greiner [1, S. 33 und S. 39] ist die Fokalkurve der Kegelschnittschar (22) eine Kurve 3. Ordnung durch die 3 Paare von Gegenpunkten des zulässigen Vierseits. Sie hat (1.c) im absoluten Punkt einen Doppelpunkt. Die Doppelpunktstangenten sind die Brennstrahlen der Enveloppe der isotropen Hauptachsen. Damit ist gezeigt:

SATZ 3. *Jede Tridens-Kurve ist Fokalkurve der einem sie erzeugenden zulässigen Sehnenvierseit einbeschriebenen Kegelschnittschar.*

LITERATURVERZEICHNIS

[1] GREINER, M., Über den Ort der Berührungspunkte der Tangenten von einem Punkt an die Kegelschnitte einer Schar oder eines Büschels, *Arch. Math. Physik* **69** (1888), 30–43.

[2] Palman, D., Über zirkuläre Kurven 3. Ordnung der isotropen Ebene, *Rad. Jugoslav Akad. Znan. Umjet.* no. **444** (1989), 37–46. *MR* **92g**:51021

[3] Palman, D., Eine Verallgemeinerung des Potenzbegriffes in der isotropen Ebene und im einfach isotropen Raum, *Rad. Jugoslav Akad. Znan. Umjet.* no. **456** (1991), 81–95.

[4] Sachs, H., Oskulierende und hyperoskulierende Kegelschnittbüschel der isotropen Ebene, *Österreich. Akad. Wiss. Math. Natur. Kl. Sitzungsber. II* **196** (1987), no. 8–10, 337–375. *MR* **90c**:51030

[5] Sachs, H., *Ebene isotrope Geometrie*, Vieweg-Verlag, Wiesbaden, 1987.

[6] Sachs, H., Vollständige zirkuläre Kurven n-ter Ordnung der isotropen Ebene, *Studia Sci. Math. Hungar.* **24** (1989), 377–383. *MR* **92a**:51028

[7] Strubecker, K., Geometrie in einer isotropen Ebene, *Math. Naturwiss. Unterricht* **15** (1962/63), 297–306, 343–351. *MR* **26**#5465. III. ibid. 385–395. *MR* **27** #631

[8] Tölke, J., Zu den einem Vierseit umschriebenen zirkulären Kurven 3. Ordnung der isotropen Ebene (in Vorbereitung.)

(Eingegangen am 6. November 1995)

UNIVERSITÄT-GESAMTHOCHSCHULE SIEGEN
FACHBEREICH 6 – MATHEMATIK
HÖLDERLINSTRASSE 3
D–57068 SIEGEN
GERMANY

schuetz@mathematik.uni-siegen.d400.de

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 441–454

A NONLINEAR EIGENVALUE PROBLEM RELATED TO GABRIELLA BOGNÁR'S CONJECTURE

W. PIELICHOWSKI

Abstract

In this paper we show the simplicity and some other properties of the principal eigenvalue of the Dirichlet problem for a quasilinear second order elliptic operator in a non-smooth domain.

1. Introduction

Let Ω be a bounded domain (i.e., an open and connected set) in the space $\mathbb{R}^N$ $(N \geqq 1)$ with the closure $\overline{\Omega}$ and boundary $\partial\Omega$. We shall be concerned with the eigenvalue problem of the form

$$(1)\quad \begin{cases} -\sum_{i=1}^{N} D_i(a_i(x)[D_i u(x)]_p) + a_0(x)[u(x)]_p = \lambda m(x)[u(x)]_p, & \forall x \in \Omega, \\ u(x) = 0, & \forall x \in \partial\Omega, \end{cases}$$

where $p \in (1, \infty)$ and $[u]_p := |u|^{p-2}u$ for $u \in \mathbb{R}$. In what follows the coefficients a_i $(i = 1, \dots, N)$, a_0 and m are supposed to be essentially bounded in Ω and satisfy the following conditions:

$$(2)\qquad \exists \alpha > 0: \ a_i(x) \geqq \alpha \quad \text{for a.e. } x \in \Omega \quad (i = 1, \dots, N),$$

$$(3)\qquad a_0(x) \geqq 0 \quad \text{for a.e. } x \in \Omega,$$

$$(4)\qquad m_+ := \max(m, 0) \neq 0 \text{ a.e. in } \Omega.$$

Putting $a_i(x) \equiv 1$, $a_0(x) \equiv 0$ and $m(x) \equiv 1$, we obtain the eigenvalue problem

$$(5)\qquad \begin{cases} -\sum_{i=1}^{N} D_i\left([D_i u]_p\right) = \lambda [u]_p \quad \text{in } \Omega, \\ u|_{\partial\Omega} = 0. \end{cases}$$

1991 *Mathematics Subject Classification*. Primary 35P30, 35J60.
Key words and phrases. First eigenvalue, quasilinear elliptic operator.

0081-6906/97/$ 5.00

The problem (5) was investigated (in the case $N=2$) by G. Bognár [4], who conjectured that the smallest eigenvalue of (5) is equal to the infimum of the Rayleigh quotient

$$R(u) := \frac{\int_\Omega \sum_{i=1}^{N} |D_i u|^p \, dx}{\int_\Omega |u|^p \, dx}$$

taken over an appropriate space of functions, contained in $W_0^{1,p}(\Omega)$. Bognár also conjectured that the smallest eigenvalue is simple and has a positive eigenfunction. She partly proved her conjecture under some additional assumptions, including the hypothesis that $p \geqq 2$. The main purpose of this paper is to obtain similar results for the first eigenvalue of problem (1), which include those conjectured by Bognár. It is worth noting that we make no regularity assumptions on the domain Ω and also the function m, usually called the *weight* function, may change sign in Ω. Our approach is based on the methods developed in the papers [3], [8], [7], [5] and [9], which were devoted to the study of the first eigenvalue of the pseudo-Laplacian $\Delta_p = \operatorname{div}([\nabla]_p)$ and some related quasilinear elliptic operators.

2. Preliminaries and main results

In this note we use real function spaces only. Let us recall that $W^{1,p}(\Omega)$, with $1<p<\infty$, denotes the space of all functions which together with their derivatives (in the distribution sense) $D_i u$ ($i=1,\dots,N$) belong to $L^p(\Omega)$. We define the norm in $W^{1,p}(\Omega)$ by setting

$$\|u\|_{W^{1,p}} := \left\{ \int_\Omega \left(|u|^p + \sum_{i=1}^{N} |D_i u|^p \right) dx \right\}^{1/p}, \qquad \forall u \in W^{1,p}(\Omega).$$

As usual, the symbol $W_0^{1,p}(\Omega)$ stands for the subspace of $W^{1,p}(\Omega)$ obtained by closing the set of all C^∞-functions with compact support in Ω. The space $W_0^{1,p}(\Omega)$ inherits the norm from $W^{1,p}(\Omega)$.

Any nontrivial function u is said to be an eigenfunction of problem (1) if and only if

$$(\mathrm{E}_\lambda) \qquad \begin{cases} u \in W_0^{1,p}(\Omega), \\ \int_\Omega \left(\sum_{i=1}^{N} a_i [D_i u]_p D_i \varphi + a_0 [u]_p \varphi \right) dx = \lambda \int_\Omega m\, [u]_p \varphi \, dx, \\ \qquad\qquad \forall \varphi \in W_0^{1,p}(\Omega). \end{cases}$$

We shall make use of the following notations:

$$A(v) := \int_\Omega \left(\sum_{i=1}^{N} a_i \, |D_i v|^p + a_0 \, |v|^p \right) dx, \qquad \forall v \in W_0^{1,p}(\Omega),$$

$$M(v) := \int_\Omega m \, |v|^p \, dx, \qquad \forall v \in W_0^{1,p}(\Omega),$$

$$\nu_1 := \sup \left\{ \frac{M(v)}{A(v)} \mid v \in W_0^{1,p}(\Omega) \setminus \{0\} \right\}, \qquad \lambda_1 := \frac{1}{\nu_1},$$

$$J_\lambda(u) := A(u) - \lambda M(u), \qquad \forall \lambda \in \mathbb{R}, \ \forall u \in W_0^{1,p}(\Omega).$$

The main result of this paper is the following:

THEOREM 1. *Suppose that the functions* a_i $(i = 1, \ldots, N)$, a_0 *and* m *satisfy all the assumptions of Section 1. Then the following assertions hold:*

(A_1) $0 < \lambda_1 < \infty$,

(A_2) *for* $\lambda \in [0, \lambda_1)$ *the problem* (E_λ) *admits no eigenfunction,*

(A_3) $J_{\lambda_1}(u) \geqq 0$ *for all* $u \in W_0^{1,p}(\Omega)$,

(A_4) *a nontrivial function* $u \in W_0^{1,p}(\Omega)$ *is an eigenfunction of problem* (E_{λ_1}) *if and only if* $J_{\lambda_1}(u) = 0$,

(A_5) *there exists a function* $u_1 \in W_0^{1,p}(\Omega) \cap C(\Omega)$ *such that* $u_1(x) > 0$ *for all* $x \in \Omega$, u_1 *is an eigenfunction of problem* (E_{λ_1}) *and the set of all solutions of* (E_{λ_1}) *is of the form* $\{tu_1 \mid t \in \mathbb{R}\}$,

(A_6) *the problem* (E_λ) *with* $\lambda > 0$ *admits a nonnegative eigenfunction if and only if* $\lambda = \lambda_1$,

(A_7) *the first eigenvalue* λ_1 *is isolated.*

The proof of Theorem 1 will be given in Section 5. It is preceded by some auxiliary results.

3. Regularity of eigenfunctions

In this section we state two lemmas concerning the regularity of eigenfunctions of problem (E_λ).

LEMMA 1. *Suppose that* $u \in W_0^{1,p}(\Omega)$ *is an eigenfunction of problem* (E_λ). *Then* $u \in L^\infty(\Omega)$.

PROOF. It is sufficient to show that there exists a constant $C > 0$ and a sequence of real numbers $\{\varkappa_n\}$, satisfying

(6) $$\lim_{n \to \infty} \varkappa_n = +\infty \qquad \text{and} \qquad \|u\|_{L^{\varkappa_n}} \leqq C, \quad \forall n \in \mathbb{N}.$$

In this connection see [1], Theorem 2.8. Our proof is a slight modification of that of [5], Lemma 3.2. For the sake of completeness we present it here in detail.

We first introduce a family of auxiliary functions $u_T : \Omega \to \mathbb{R}$ $(T \in \mathbb{R}^+)$ defined by

$$u_T(x) := \min(u_+(x), T), \qquad \forall x \in \Omega.$$

It is easy to see that $u_T \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega)$ and hence

$$u_T^\beta := (u_T)^\beta \in W_0^{1,p}(\Omega) \cap L^\infty(\Omega), \qquad \forall \beta \geqq 1 \tag{7}$$

(compare with [6], Lemma A.3). Next, we use the functions $\varphi = u_T^{kp+1}$ (where $k \geqq 0$) as test functions in (E_λ), which is legitimate because of (7). As a result we obtain the equality

$$\sum_{i=1}^N \int_\Omega a_i [D_i u]_p D_i(u_T^{kp+1})\, dx = \int_\Omega (\lambda m - a_0)[u]_p u_T^{kp+1}\, dx, \qquad \forall k \geqq 0,\ \forall T > 0. \tag{8}$$

Let us note that the left-hand side of (8) can be estimated as follows

$$\begin{aligned}
\sum_{i=1}^N \int_\Omega a_i [D_i u]_p D_i(u_T^{kp+1})\, dx &= \sum_{i=1}^N \int_\Omega a_i [D_i u]_p (kp+1) u_T^{kp} D_i(u_T)\, dx \\
&= \sum_{i=1}^N \int_\Omega a_i (kp+1) u_T^{kp} |D_i(u_T)|^p\, dx \\
&\geqq \alpha(kp+1) \sum_{i=1}^N \int_\Omega u_T^{kp} |D_i(u_T)|^p\, dx \\
&= \frac{\alpha(kp+1)}{(k+1)^p} \sum_{i=1}^N \int_\Omega |D_i(u_T^{k+1})|^p\, dx \\
&= \frac{\alpha(kp+1)}{(k+1)^p} \sum_{i=1}^N \|D_i(u_T^{k+1})\|_{L^p}^p .
\end{aligned} \tag{9}$$

Here we have used inequality (2). Moreover, according to the Poincaré inequality (see for instance [1], Section 6.26) we have the estimate

$$\|v\|_{W_0^{1,p}}^p \leqq C_1 \sum_{i=1}^N \|D_i v\|_{L^p}^p, \qquad \forall v \in W_0^{1,p}(\Omega), \tag{10}$$

where $C_1 > 0$ is a constant independent of v. From the Sobolev embedding theorem (see, e.g., [1], Theorem 5.4) it follows that there is a real number q, greater than p (for example $q = 2p$ when $p \geqq N$ and $q = \frac{Np}{N-p}$ when $1 < p < N$), such that $W_0^{1,p}(\Omega) \subset L^q(\Omega)$ and the embedding is continuous, so that

$$\|v\|_{L^q} \leqq C_2 \|v\|_{W_0^{1,p}}, \qquad \forall v \in W_0^{1,p}(\Omega), \tag{11}$$

where $C_2 > 0$ is a constant independent of v. Combining inequalities (9), (10) and (11) we conclude that

$$\sum_{i=1}^{N} \int_\Omega a_i [D_i u]_p D_i(u_T^{kp+1})\, dx \geqq \frac{\alpha(kp+1)}{(k+1)^p C_3} \left\{ \int_\Omega (u_T^{k+1})^q\, dx \right\}^{p/q} \tag{12}$$

for all $k \geqq 0$ and $T > 0$, where $C_3 := C_1 C_2^p$ is a constant independent of k and T.

We now estimate the right-hand side of (8):

$$\begin{aligned} \int_\Omega (\lambda m - a_0)[u]_p u_T^{kp+1}\, dx &\leqq \int_\Omega |\lambda m - a_0|\, |u|^{p-1} u_T^{kp+1}\, dx \\ &\leqq \int_\Omega |\lambda m - a_0|\, u_+^{(k+1)p}\, dx \\ &\leqq C_4 \int_\Omega u_+^{(k+1)p}\, dx, \qquad \forall k \geqq 0,\ \forall T > 0, \end{aligned} \tag{13}$$

where $C_4 := \|\lambda m - a_0\|_{L^\infty}$ (since the integrals are taken over the support of the function u_+).

The inequalities (12), (13) yield the estimate

$$\left\{ \int_\Omega (u_T^{k+1})^q\, dx \right\}^{p/q} \leqq C_5 \frac{(k+1)^p}{kp+1} \int_\Omega u_+^{(k+1)p}\, dx, \quad \forall k \geqq 0,\ \forall T > 0,$$

where $C_5 := C_3 C_4 \alpha^{-1}$. Equivalently,

$$\|u_T\|_{L^{(k+1)q}} \leqq C_6^{\frac{1}{k+1}} \left\{ \frac{k+1}{(kp+1)^{1/p}} \right\}^{\frac{1}{k+1}} \|u_+\|_{L^{(k+1)p}}, \quad \forall k \geqq 0,\ \forall T > 0, \tag{14}$$

where $C_6 := C_5^{1/p}$.

Since $p<q$, we may choose $k_0>0$ in such a way that $(k_0+1)p=q$. Hence the inequality (14) with $k:=k_0$ gives

$$\|u_T\|_{L^{(k_0+1)q}} \leqq C_6^{\frac{1}{k_0+1}} \left\{ \frac{k_0+1}{(k_0p+1)^{1/p}} \right\}^{\frac{1}{k_0+1}} \|u_+\|_{L^q}, \qquad \forall T>0.$$

Letting $T\to\infty$, we see by Fatou's lemma that

$$\|u_+\|_{L^{(k_0+1)q}} \leqq C_6^{\frac{1}{k_0+1}} \left\{ \frac{k_0+1}{(k_0p+1)^{1/p}} \right\}^{\frac{1}{k_0+1}} \|u_+\|_{L^q}.$$

Now we can choose $k_1>0$ so that $(k_1+1)p=(k_0+1)q=q^2/p$. Putting $k:=k_1$ in (14), we obtain the estimate

$$\|u_T\|_{L^{(k_1+1)q}} \leqq C_6^{\frac{1}{k_1+1}} \left\{ \frac{k_1+1}{(k_1p+1)^{1/p}} \right\}^{\frac{1}{k_1+1}} \|u_+\|_{L^{(k_0+1)q}}, \qquad \forall T>0,$$

which as before leads to the conclusion that

$$\|u_+\|_{L^{(k_1+1)q}} \leqq C_6^{\frac{1}{k_1+1}} \left\{ \frac{k_1+1}{(k_1p+1)^{1/p}} \right\}^{\frac{1}{k_1+1}} \|u_+\|_{L^{(k_0+1)q}}.$$

By induction we obtain the estimate

$$\|u_+\|_{L^{(k_n+1)q}} \leqq C_6^{\frac{1}{k_n+1}} \left\{ \frac{k_n+1}{(k_np+1)^{1/p}} \right\}^{\frac{1}{k_n+1}} \|u_+\|_{L^{(k_{n-1}+1)q}}, \qquad \forall n\in\mathbb{N},$$

where $k_n+1=(q/p)^{n+1}$ $(n=0,1,\ldots)$. Finally, we arrive at the inequality

$$\begin{aligned}\|u_+\|_{L^{(k_n+1)q}} &\leqq C_6^{\sum_{i=0}^{n}\frac{1}{k_i+1}} \prod_{i=0}^{n} \left\{ \frac{k_i+1}{(k_ip+1)^{1/p}} \right\}^{\frac{1}{k_i+1}} \|u_+\|_{L^q} \\ &= C_6^{\sum_{i=0}^{n}\frac{1}{k_i+1}} \prod_{i=0}^{n} \left[\left\{ \frac{k_i+1}{(k_ip+1)^{1/p}} \right\}^{\frac{1}{\sqrt{k_i+1}}} \right]^{\frac{1}{\sqrt{k_i+1}}} \|u_+\|_{L^q}, \qquad \forall n\in\mathbb{N}.\end{aligned} \tag{15}$$

Since

$$\lim_{y\to\infty} \left\{ \frac{y+1}{(yp+1)^{1/p}} \right\}^{\frac{1}{\sqrt{y+1}}} = 1,$$

there exists a constant C_7 such that

$$\left\{ \frac{k_i+1}{(k_ip+1)^{1/p}} \right\}^{\frac{1}{\sqrt{k_i+1}}} \leqq C_7, \qquad \forall i\in\mathbb{N}. \tag{16}$$

From (15) and (16) we infer the estimate

$$\|u_+\|_{L^{(k_n+1)q}} \leqq C_6^{\sum_{i=0}^{n} \frac{1}{k_i+1}} C_7^{\sum_{i=0}^{n} \frac{1}{\sqrt{k_i+1}}} \|u_+\|_{L^q} \tag{17}$$

$$\leqq C_6^{\sum_{i=0}^{\infty} \frac{1}{k_i+1}} C_7^{\sum_{i=0}^{\infty} \frac{1}{\sqrt{k_i+1}}} \|u_+\|_{L^q}, \qquad \forall n \in \mathbb{N}.$$

By similar computations we get the inequality

$$\|u_-\|_{L^{(k_n+1)q}} \leqq C_6^{\sum_{i=0}^{\infty} \frac{1}{k_i+1}} C_7^{\sum_{i=0}^{\infty} \frac{1}{\sqrt{k_i+1}}} \|u_-\|_{L^q}, \qquad \forall n \in \mathbb{N}. \tag{18}$$

It follows from (17) and (18) that

$$\|u\|_{L^{(k_n+1)q}} \leqq 2\, C_6^{\sum_{i=0}^{\infty} \frac{1}{k_i+1}} C_7^{\sum_{i=0}^{\infty} \frac{1}{\sqrt{k_i+1}}} \|u\|_{L^q}, \qquad \forall n \in \mathbb{N}.$$

Note that

$$\sum_{i=0}^{\infty} \frac{1}{k_i+1} = \sum_{i=0}^{\infty} (p/q)^{n+1} < \infty$$

and that

$$\sum_{i=0}^{\infty} \frac{1}{\sqrt{k_i+1}} = \sum_{i=0}^{\infty} \left(\sqrt{p/q}\right)^{n+1} < \infty,$$

since $p < q$. Thus we can put $\varkappa_n := (k_n+1)q$ $(n = 1, 2, \dots)$ and

$$C := 2\, C_6^{\sum_{i=0}^{\infty} \frac{1}{k_i+1}} C_7^{\sum_{i=0}^{\infty} \frac{1}{\sqrt{k_i+1}}} \|u\|_{L^q}$$

to see that the conditions (6) are satisfied. This completes the proof. □

LEMMA 2. *Let u be an eigenfunction of problem* (E_λ). *Then for any compact $K \subset \Omega$ there exists $\alpha \in (0,1)$ such that $u \in C^{0,\alpha}(K)$. Furthermore, if $u \geqq 0$ in Ω, then $u(x) > 0$ for all $x \in \Omega$.*

PROOF. On the base of Lemma 1, the $C^{0,\alpha}$-regularity of the eigenfunction u can be derived from [10], Theorem 2.2. Now the positivity of u follows from the Harnack type inequality of Trudinger [10], Theorem 1.1. □

4. Inequalities of the Anane-Lindqvist type

Following the ideas of Lindqvist [7], which are continuation of those introduced by Anane [2], we formulate some inequalities in Lemma 4 below. They are one-dimensional versions of inequalities implicitly used in [7] and [5]. In order to prove them we need the following

LEMMA 3. *If* $p \geqq 2$, *then there exists a constant* $C(p) > 0$ *such that*

$$|t_2|^p - |t_1|^p - p|t_1|^{p-2}t_1(t_2-t_1) \geqq C(p)|t_2-t_1|^p, \qquad \forall t_1, t_2 \in \mathbb{R}. \tag{19a}$$

If $1 < p < 2$, *then there exists a constant* $C(p) > 0$ *such that*

$$|t_2|^p - |t_1|^p - p|t_1|^{p-2}t_1(t_2-t_1) \geqq C(p)\frac{|t_2-t_1|^2}{(1+|t_1|+|t_2|)^{2-p}}, \qquad \forall t_1, t_2 \in \mathbb{R}. \tag{19b}$$

The proof of Lemma 3 can be found in [7], Section 4, where it is explained that one can choose

$$C(p) = \begin{cases} \frac{1}{2^{p-1}-1} & \text{for } p \geqq 2, \\ \frac{3}{16}p(p-1) & \text{for } 1 < p < 2. \end{cases} \qquad \square$$

Now, let us introduce the set

$$S := \{u \in W^{1,p}(\Omega) \cap L^\infty(\Omega) \mid \exists \varepsilon > 0 : \ u \geqq \varepsilon \text{ in } \Omega\}.$$

It is easy to see that for every $u, v \in S$ the functions

$$\frac{u^p - v^p}{u^{p-1}} \quad \text{and} \quad \frac{v^p - u^p}{v^{p-1}}$$

belong to $W^{1,p}(\Omega)$. Moreover,

$$D_i\left(\frac{u^p - v^p}{u^{p-1}}\right) = \left\{1 + (p-1)\left(\frac{v}{u}\right)^p\right\} D_i u - p\left(\frac{v}{u}\right)^{p-1} D_i v \quad (i = 1, \dots, N) \tag{20}$$

and

$$D_i\left(\frac{v^p - u^p}{v^{p-1}}\right) = \left\{1 + (p-1)\left(\frac{u}{v}\right)^p\right\} D_i v - p\left(\frac{u}{v}\right)^{p-1} D_i u \quad (i = 1, \dots, N). \tag{21}$$

LEMMA 4. *If* $p \geqq 2$, *then there is a constant* $C(p) > 0$ *such that for every* $u, v \in S$ *the inequality*

$$\begin{aligned} &[D_i u]_p D_i\left(\frac{u^p - v^p}{u^{p-1}}\right) + [D_i v]_p D_i\left(\frac{v^p - u^p}{v^{p-1}}\right) \\ &\geqq C(p)\left(\frac{1}{u^p} + \frac{1}{v^p}\right)|vD_i u - uD_i v|^p \end{aligned} \tag{22a}$$

holds almost everywhere in Ω $(i=1,\dots,N)$.

If $1<p<2$, *then there is a constant* $C(p)>0$ *such that for every* $u,v\in S$ *the inequality*

$$\text{(22b)}\qquad \begin{aligned}&[D_iu]_pD_i\left(\frac{u^p-v^p}{u^{p-1}}\right)+[D_iv]_pD_i\left(\frac{v^p-u^p}{v^{p-1}}\right)\\ &\geqq C(p)\left(\frac{1}{u^p}+\frac{1}{v^p}\right)\frac{|vD_iu-uD_iv|^2}{(uv+|vD_iu|+|uD_iv|)^{2-p}}\end{aligned}$$

holds almost everywhere in Ω $(i=1,\dots,N)$.

PROOF. To begin with we transform the left-hand side of (22). Using formulae (20) and (21), we obtain

$$\begin{aligned}L:=&[D_iu]_pD_i\left(\frac{u^p-v^p}{u^{p-1}}\right)+[D_iv]_pD_i\left(\frac{v^p-u^p}{v^{p-1}}\right)\\ &=\left\{1+(p-1)\left(\frac{v}{u}\right)^p\right\}|D_iu|^p+\left\{1+(p-1)\left(\frac{u}{v}\right)^p\right\}|D_iv|^p\\ &-p\left(\frac{v}{u}\right)^{p-1}|D_iu|^{p-2}D_iuD_iv-p\left(\frac{u}{v}\right)^{p-1}|D_iv|^{p-2}D_ivD_iu\\ &=(u^p-v^p)(|t_2|^p-|t_1|^p)-pv^p|t_2|^{p-2}t_2(t_1-t_2)-pu^p|t_1|^{p-2}t_1(t_2-t_1)\\ &=u^p\{|t_2|^p-|t_1|^p-p|t_1|^{p-2}t_1(t_2-t_1)\}+v^p\{|t_1|^p-|t_2|^p-p|t_2|^{p-2}t_2(t_1-t_2)\},\end{aligned}$$

where $t_1:=\dfrac{D_iv}{v}$ and $t_2:=\dfrac{D_iu}{u}$. In the case $p\geqq 2$ inequality (19a) gives

$$\begin{aligned}L&\geqq C(p)(u^p+v^p)|t_2-t_1|^p\\ &=C(p)\left(\frac{1}{u^p}+\frac{1}{v^p}\right)|vD_iu-uD_iv|^p,\end{aligned}$$

as desired. In the case $1<p<2$ inequality (19b) implies that

$$\begin{aligned}L&\geqq C(p)(u^p+v^p)\frac{|t_2-t_1|^2}{(1+|t_1|+|t_2|)^{2-p}}\\ &=C(p)\left(\frac{1}{u^p}+\frac{1}{v^p}\right)\frac{|vD_iu-uD_iv|^2}{(uv+|vD_iu|+|uD_iv|)^{2-p}},\end{aligned}$$

and the lemma follows. □

5. Proof of Theorem 1

Since the proof of assertions (A_1)–(A_4) is quite similar to that of [8], we can omit the details here. Therefore we pass to the proof of the statement (A_5). Let us observe that the functional $A:W_0^{1,p}(\Omega)\to\mathbb{R}$, defined in Section 2, is weakly lower semicontinuous and it is coercive on the set

$$\mathcal{M}:=\{u\in W_0^{1,p}(\Omega)\mid M(u)=1\},$$

by the assumptions (2), (3). Thus we can find in $\mathcal{M}$ a minimizer $u_0 \neq 0$, characterized by

$$A(u_0) = \inf\{A(u) \mid u \in \mathcal{M}\}$$

(see for instance [3], Theorem 6.1.1). Hence $A(u_0) = \lambda_1$ and so $J_{\lambda_1}(u_0) = 0$. According to the assertion (A$_4$), u_0 is an eigenfunction of problem (E$_{\lambda_1}$). Since $J_{\lambda_1}(|u_0|) = J_{\lambda_1}(u_0) = 0$, the function $u_1 := |u_0|$ is also an eigenfunction of (E$_{\lambda_1}$). Now Lemma 2 shows that $u_1 \in C(\Omega)$ and $u_1(x) > 0$ for all $x \in \Omega$. Next, let us suppose that v is another eigenfunction of problem (E$_{\lambda_1}$). Observe that v cannot change sign in Ω, because otherwise the function $|v|$ would be a nonnegative eigenfunction of (E$_{\lambda_1}$) vanishing somewhere in Ω, which contradicts Lemma 2. Thus we can assume that $v \in C(\Omega)$ and $v(x) > 0$ for all $x \in \Omega$. To simplify our notation, we put $u := u_1$. Since, by Lemma 1, $u, v \in L^\infty(\Omega)$, we see that the functions

$$\varphi := \frac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}} \qquad \text{and} \qquad \psi := \frac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}},$$

where $u_\varepsilon = u + \varepsilon$ and $v_\varepsilon = v + \varepsilon$ ($\varepsilon > 0$), belong to the space $W_0^{1,p}(\Omega)$. We now use the test function φ in (E$_{\lambda_1}$) to conclude that

$$\sum_{i=1}^{N} \int_\Omega a_i [D_i u]_p D_i \left(\frac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}} \right) dx = \int_\Omega (\lambda_1 m - a_0)[u]_p \frac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}}\, dx. \tag{23}$$

Analogously, putting ψ in (E$_{\lambda_1}$), we get the equality

$$\sum_{i=1}^{N} \int_\Omega a_i [D_i v]_p D_i \left(\frac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}} \right) dx = \int_\Omega (\lambda_1 m - a_0)[v]_p \frac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}}\, dx. \tag{24}$$

Note that

$$D_i u = D_i u_\varepsilon, \quad D_i v = D_i v_\varepsilon, \qquad \forall \varepsilon > 0 \quad (i = 1, \ldots, N). \tag{25}$$

It follows from (23), (24) and (25) that

$$\sum_{i=1}^{N} \int_\Omega a_i \left\{ [D_i u_\varepsilon]_p D_i \left(\frac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}} \right) + [D_i v_\varepsilon]_p D_i \left(\frac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}} \right) \right\} dx$$
$$= \int_\Omega (\lambda_1 m - a_0) \left\{ \left(\frac{u}{u_\varepsilon} \right)^{p-1} - \left(\frac{v}{v_\varepsilon} \right)^{p-1} \right\} (u_\varepsilon^p - v_\varepsilon^p)\, dx.$$

Given any $\varepsilon > 0$, we see that $u_\varepsilon, v_\varepsilon \in S$ and hence we can use Lemma 4, which together with the assumption (2) implies that

$$\int_\Omega (\lambda_1 m - a_0)\left\{\left(\frac{u}{u_\varepsilon}\right)^{p-1} - \left(\frac{v}{v_\varepsilon}\right)^{p-1}\right\}(u_\varepsilon^p - v_\varepsilon^p)\,dx$$
$$\geqq \alpha C(p)\sum_{i=1}^N \int_\Omega \left(\frac{1}{u_\varepsilon^p} + \frac{1}{v_\varepsilon^p}\right)|v_\varepsilon D_i u_\varepsilon - u_\varepsilon D_i v_\varepsilon|^p\,dx$$

when $p \geqq 2$, and that

$$\int_\Omega (\lambda_1 m - a_0)\left\{\left(\frac{u}{u_\varepsilon}\right)^{p-1} - \left(\frac{v}{v_\varepsilon}\right)^{p-1}\right\}(u_\varepsilon^p - v_\varepsilon^p)\,dx$$
$$\geqq \alpha C(p)\sum_{i=1}^N \int_\Omega \left(\frac{1}{u_\varepsilon^p} + \frac{1}{v_\varepsilon^p}\right)\frac{|v_\varepsilon D_i u_\varepsilon - u_\varepsilon D_i v_\varepsilon|^2}{(u_\varepsilon v_\varepsilon + |v_\varepsilon D_i u_\varepsilon| + |u_\varepsilon D_i v_\varepsilon|)^{2-p}}\,dx$$

in the case $1 < p < 2$. Hence, letting $\varepsilon \to 0$, we get by the Fatou lemma the equalities

$$|vD_iu - uD_iv| = 0 \quad \text{in } \Omega \quad (i = 1, \ldots, N).$$

Thus there exists a constant $t \in \mathbb{R}$ such that $v = tu$ in Ω, and so part (A_5) is proved.

In a similar way we prove assertion (A_6). We first assume that $m \geqq 0$ in Ω. Let u be a nonnegative eigenfunction of problem (E_{λ_1}). Suppose, contrary to our claim, that there is a nonnegative eigenfunction v of problem (E_λ) with $\lambda > \lambda_1$. The case $\lambda < \lambda_1$ is excluded by assertion (A_2). Then we can use $\dfrac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}}$ as a test function in (E_{λ_1}), and $\dfrac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}}$ as a test function in (E_λ) to obtain the equality

$$\sum_{i=1}^N \int_\Omega a_i\left\{[D_iu_\varepsilon]_p D_i\left(\frac{u_\varepsilon^p - v_\varepsilon^p}{u_\varepsilon^{p-1}}\right) + [D_iv_\varepsilon]_p D_i\left(\frac{v_\varepsilon^p - u_\varepsilon^p}{v_\varepsilon^{p-1}}\right)\right\}dx$$
$$= \int_\Omega \left\{\lambda_1\left(\frac{u}{u_\varepsilon}\right)^{p-1} - \lambda\left(\frac{v}{v_\varepsilon}\right)^{p-1}\right\} m(u_\varepsilon^p - v_\varepsilon^p)\,dx$$
$$- \int_\Omega a_0\left\{\left(\frac{u}{u_\varepsilon}\right)^{p-1} - \left(\frac{v}{v_\varepsilon}\right)^{p-1}\right\}(u_\varepsilon^p - v_\varepsilon^p)\,dx.$$

It follows from Lemma 4 that the above expression is nonnegative. Letting $\varepsilon \to 0$, we find that

$$\int\limits_{\Omega} (\lambda_1 - \lambda) m (u^p - v^p)\, dx \geqq 0.$$

Hence

$$0 \leqq \int\limits_{\Omega} m u^p\, dx \leqq \int\limits_{\Omega} m v^p\, dx.$$

Since v may be replaced by $\frac{1}{2}v$, $\frac{1}{3}v$, etc., we conclude that

$$\int\limits_{\Omega} m u^p\, dx = 0.$$

Consequently, $m u^p \equiv 0$ in Ω. Thus the positivity of the eigenfunction u contradicts assumption (4). This completes the proof of (A_6) under the temporary assumption that $m \geqq 0$. Now we can take any function $m \in L^\infty(\Omega)$ such that $m_+ \not\equiv 0$ a.e. in Ω and argue as in [8], with obvious modifications, to complete the proof of part (A_6).

The proof of assertion (A_7) is based on ideas of Anane [2]. We begin by showing that there is a continuous function $\mathbb{R} \ni \lambda \mapsto B(\lambda) \in \mathbb{R}^+$ such that if v is any eigenfunction of problem (E_λ), for which $v_- \not\equiv 0$, then

$$|\Omega^-| \geqq B(\lambda), \tag{26}$$

where $\Omega^- = \{x \in \Omega \mid v(x) < 0\}$. To see this, we apply the test function v_- in (E_λ), which yields

$$\sum_{i=1}^{N} \int\limits_{\Omega} a_i |D_i v_-|^p\, dx = \int\limits_{\Omega^-} (\lambda m - a_0) |v_-|^p\, dx. \tag{27}$$

Note that

$$\sum_{i=1}^{N} \int\limits_{\Omega} a_i |D_i v_-|^p\, dx \geqq \frac{\alpha}{C_3} \|v_-\|^p_{L^q}, \tag{28}$$

where q and C_3 are the constants introduced in Lemma 1. On the other hand, the inequality $p < q$ implies that

$$\begin{aligned} \int\limits_{\Omega^-} (\lambda m - a_0) |v_-|^p\, dx &\leqq B_1(\lambda) \int\limits_{\Omega^-} |v_-|^p\, dx \\ &\leqq B_1(\lambda)\, |\Omega^-|^{\frac{q-p}{q}} \|v_-\|^p_{L^q}, \end{aligned} \tag{29}$$

where $B_1(\lambda) := 1 + \|\lambda m - a_0\|_{L^\infty} > 0$. From (27), (28) and (29) we get the estimate (26) with

$$B(\lambda) := \left\{ \frac{\alpha}{C_3 B_1(\lambda)} \right\}^{\frac{q}{q-p}}.$$

Now, let us suppose that λ_1 is not isolated and so there is a sequence $\{\lambda_n\}$ of eigenvalues $\lambda_n > \lambda_1$ $(n = 2, 3, \dots)$ such that $\lim_{n\to\infty} \lambda_n = \lambda_1$. Let $\{u_n\}$ denote the sequence of corresponding eigenfunctions scaled in such a way that $M(u_n) = 1$ $(n = 2, 3, \dots)$. Using u_n as a test function in (E_{λ_n}) we conclude that $A(u_n) = \lambda_n M(u_n) = \lambda_n$. Consequently, the sequence $\{u_n\}$ is bounded in $W_0^{1,p}(\Omega)$. Passing to a subsequence if necessary, we can assume that the sequence $\{u_n\}$ is weakly convergent in $W_0^{1,p}(\Omega)$. We denote the limit function by u_0. Observe that

$$A(u_0) \leqq \liminf_{n\to\infty} A(u_n) = \lim_{n\to\infty} \lambda_n = \lambda_1. \tag{30}$$

Since the embedding $W_0^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ is compact (see, e.g., [1], Theorem 6.2), the sequence $\{u_n\}$ converges in $L^p(\Omega)$. Therefore $M(u_0) = 1$ and hence $u_0 \not\equiv 0$ in Ω. The inequality (30) shows that u_0 is an eigenfunction of (E_{λ_1}). By assertion (A_5), the function u_0 does not change sign. Without loss of generality we may assume that $u_0(x) > 0$ for all $x \in \Omega$. We can assume, after passing to a subsequence, that $\{u_n\}$ converges almost everywhere in Ω. According to the Jegorov theorem, the sequence $\{u_n\}$ converges almost uniformly in Ω to the positive function u_0, which contradicts the estimate (26) applied to u_n. This completes the proof. □

Acknowledgements. The author wishes to thank Professor J. Bochenek for drawing the author's attention to the paper [4] as well as Professor P. Drábek for sending a copy of the preprint [5].

REFERENCES

[1] Adams, R. A., *Sobolev spaces*, Pure and Applied Mathematics, Vol. 65, Academic Press, New York, 1975. *MR* **56** #9247

[2] Anane, A., Simplicité et isolation de la première valeur propre du p-laplacien avec poids, *C. R. Acad. Sci. Paris. Sér. I Math.* **305** (1987), 725–728. *MR* **89e**:35124

[3] Berger, M. S., *Nonlinearity and functional analysis*, Lectures on nonlinear problems in mathematical analysis, Pure and Applied Mathematics, Academic Press, New York, 1977. *MR* **58** #7671

[4] Bognár, G., Eigenvalue problem for some nonlinear elliptic partial differential equation, *Studia Sci. Math. Hungar.* **29** (1994), 213–231. *MR* **95g**:35061

[5] Drábek, P., Strongly nonlinear degenerated and singular elliptic problems, Preprint 74, University of West Bohemia, Pilsen, 1995.

[6] Kinderlehrer, D. and Stampacchia, G., *An introduction to variational inequalities and their applications*, Pure and Applied Mathematics, 88, Academic Press, New York, 1980. *MR* **81g**:49013; Russian translation: Vvedenie v variatsionnye neravenstva i ikh prilozheniya, Mir, Moscow, 1983. *MR* **85b**:49007

[7] Lindqvist, P., On the equation $\mathrm{div}(|\nabla u|^{p-2}\nabla u) + \lambda |u|^{p-2}u = 0$, *Proc. Amer. Math. Soc.* **109** (1990), 157–164. *MR* **90h**:35088

[8] Otani, M. and Teshima, T., On the first eigenvalue of some quasilinear elliptic equations, *Proc. Japan Acad. Ser. A Math. Sci.* **64** (1988), 8–10. *MR* **89h**:35257

[9] Pielichowski, W., On the first eigenvalue of a quasilinear elliptic operator, *Selected problems of mathematics*, Anniversary issue, Vol. 6, Cracow University of Technology, Cracow, 1995, 235–241.

[10] Trudinger, N. S., On Harnack type inequalities and their application to quasilinear elliptic equations, *Comm. Pure Appl. Math.* **20** (1967), 721–747. *MR* **37** #1788

(Received December 1, 1995)

INSTYTUT MATEMATYKI
POLITECHNIKA KRAKOWSKA
IM. TADEUSZA KOŚCIUSZKI
UL. WARSZAWSKA 24
PL–31-155 KRAKÓW
POLAND

wpielich@usk.pk.edu.pl

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 455–458

ON A CHARACTERIZATION OF ABELIAN GROUPS BY SUMS AND DIFFERENCES

KATALIN KOVÁCS

Let $G=\{g_1,g_2,\dots,g_n\}$ denote an abelian group of n elements, $(a)=(a_1,a_2,\dots,a_n)\in G^n$, $A=\{a_1,a_2,\dots,a_n\}\subseteqq G$ and $\|A\|:=$ the number of the different elements of A.

THEOREM 1. (i) *There exists an* $(a)\in G^n$ *such that*

$$\{a_1+a_2,a_2+a_3,\dots;a_n+a_1\}=G \tag{1}$$

if and only if $2\not|n$.

(ii) *If* $G\cong Z_n$, *then* $\max\|A\|=n-1$ *if* $4\mid n$, $\max\|A\|=n$ *if* n *is odd and* $\min\|A\|=O(\sqrt{n})$ *if* $n\equiv 3$ *or* $4 \bmod 4$.

CONJECTURE. *If* $n\equiv 1 \bmod 4$, *then* $\min\|A\|\ll n$.

THEOREM 2. (i) *There exists* $(a)\in G^n$ *such that*

$$\{a_2-a_1,a_3-a_2,\dots,a_1-a_n\}=G \tag{2}$$

if and only if the number of the elements of order 2 *in* G *is not* 1.

(ii) *If* $G\cong Z_n$ *and* $2\nmid n$, *then* $\max\|A\|=n-1$ *and* $\min\|A\|=O(\sqrt{n})$.

PROOF OF THEOREM 1. (i) For $n=4k+2$ there exists a $g\in G$ such that $o(g)=2$ and $\sum_{i=1}^{n} g_i=g$. If there exists an (a) such that (1) is satisfied then

$$2y:=2\sum_{i=1}^{n} g_i=g, \tag{3}$$

i.e. $4y=0$. Therefore $o(y)|\text{g.c.d.}(4,|G|)=2$ which implies $y=g$ or $y=0$.

Replacing in (3) $g=0$ contradicts the existence of (a).

We assume that $n\neq 4k+2$. We show a sequence (a) such that (1) is satisfied. Let us see the construction in the case $G=Z_n$. Suitable sequences are for example the following ones:

$$(1,2,\dots,n) \text{ if } n=2k+1$$

1991 *Mathematics Subject Classification.* Primary 11A99; Secondary 20C99.

Key words and phrases. Characterization of abelian groups by sums (differences).

0081-6906/97/$ 5.00 ©1997 *Akadémiai Kiadó, Budapest*

$$(0,1,1,2,2,\dots,k-1,k-1,3k,k,k,\dots,2k-1,2k-1) \text{ if } n=4k$$
$$(0,0,1,1,2,2,\dots,k-1,k-1,3k,k,k,\dots,2k-1,2k-1) \text{ if } n=4k+1$$
$$(0,1,1,2,2,\dots,k-1,k-1,k,k+1,k+1,\dots,2k-1,2k-1,2k) \text{ if } n=4k-1$$
$$(0,0,1,1,2,2,\dots,k-1,k-1,k,k+1,k+1,\dots,2k-1,2k-1,2k) \text{ if } n=4k.$$

If $4 \mid n$ then $G \cong Z_2 \oplus Z_2 \oplus T$ or $G \cong Z_{2^\alpha} \oplus T$ with $\alpha \geqq 2$ and $T=\{t_1,t_2,\dots,t_s\}$. For $Z_2 \oplus Z_2$ we have

$$(q)=(q_1,\dots,q_4)=((0,0),(1,1),(1,0),(1,0))$$

as a suitable sequence. For the cyclic group Z_{2^γ}, we showed a suitable $(q)=(q_1,\dots,q_{2^\gamma})$. If $2^\alpha \,\|\, |G|$ then the construction for G is as follows:

$$\begin{gathered}((q_1,t_1),(q_2,0),(q_3,t_1),(q_4,0),\dots,(q_{2^{\alpha-1}},t_1),(q_{2^\alpha},0),\\ (q_1,t_2),(q_2,0),(q_3,t_2),(q_4,0),\dots,(q_{2^{\alpha-1}},t_2),(q_{2^\alpha},0),\\ \vdots\\ (q_1,t_s),(q_2,0),(q_3,t_s),(q_4,0),\dots,(q_{2^{\alpha-1}},t_s),(q_{2^\alpha},0)).\end{gathered}$$

For odd n we prove by induction. Then $G \cong \bigoplus_{i=1}^{k-1} T_i \oplus Z_m$. We assume that $(t_1,\dots,t_s)$ is a suitable choice for the direct sum. Then a suitable sequence for G is the following one:

$$\begin{gathered}((t_1,1),(t_2,0),(t_3,1),(t_4,0),\dots,(t_{s-1},0),(t_s,1),\\ (t_1,2),(t_2,0),(t_3,2),(t_4,0),\dots,(t_{s-1},0),(t_s,2),\\ \vdots\\ (t_1,0),(t_2,0),(t_3,0),(t_4,0),\dots,(t_{s-1},0),(t_s,0)).\end{gathered}$$

(ii) Maximum for Z_n: For odd n we may choose $a_i=i$ for $i=1,\dots,n$.

For $n=4t$ $a_{2s}=4t-s+1$ for $s=1,\dots,t$, $a_{2s}=4t-s$ for $s=t+1,\dots,2t$ and $a_{2s-1}=2t-s$ for $s=1,\dots,2t$ is a maximal sequence.

Minimum for Z_n: We can choose less than n different sums $a+b$ such that $a,b \in [0,\sqrt{n}]$. So $\sqrt{n}$ is a trivial lower bound for the minimum of $\|A\|$.

For an arbitrary $n \in N$ $n \equiv 2,3 \bmod 4$ there exists a $d \in N$ such that

$$4d^2+4d \leqq n = 4d^2+4d+4k+i < 4(d+1)^2+4(d+1)-6$$

with $i=0$ or $i=-1$ and $0 \leqq k \leqq 2d$. So $d=O(\sqrt{n})$ and $k=O(\sqrt{n})$.

The integers in $[0,4d^2+4d-1]$ can be obtained by the following construction:

$$\begin{matrix} \dots & 0, & 4jd, & 1, & 4jd+2d, & \dots,\\ \vdots & & & & & \\ \dots & 2i-2, & 4jd, & 2i-1, & 4jd+2d, & \dots,\\ \vdots & & & & & \\ \dots & 2d-2, & 4jd, & 2d-1, & 4jd+2d, & \dots, \end{matrix}$$

where j runs over the integers of the interval $[0,d]$. We needed only $2d+(2d+2)=O(\sqrt{n})$ integers. Let us continue the previous sequence by

$$0,\dots,0,4d^2+4d+4j,1,4d^2+4d+4j+2,0,\dots,1,4d^2+4d+4j+4k+2.$$

We have the integers in $[0,4d^2+4d+4k+3]$ (case $4|n$) adding only $2k+2=O(\sqrt{n})$ numbers. Deleting the first 0 we get the same result for $n=4k+3$.

PROOF OF THEOREM 2. (i) If there is only one element g of order 2 in G, then

$$0=(a_2-a_1)+(a_3-a_2)+\dots+(a_1-a_n)=a_1+a_2+\dots+a_n=g$$

excludes this case. Otherwise $a_i=\sum_{j=1}^{i} g_j$ is a suitable sequence.

(ii) Maximum for Z_k: We may assume that $a_i\in[1,n]$. One of the differences has to be 0, so $\|A\|\leqq n-1$. For $n=4k+1$ the maximum can be obtained by the following sequence:

$$(0,0,1,-1,\dots,k,-k,-(k+1),(k+1),-(k+2),(k+2),\dots,-(2k-1),2k-1,-2k).$$

If $n=4k+3$ we continue the previous sequence by $2k$ and $-(2k+1)$.

Minimum: The minimum cannot be below $\sqrt{n}$ as $[\sqrt{n}]$ numbers have less than

$$16r^2\leqq n<16r^2+1+2k<16(r+1)^2.$$

Here $r=O(\sqrt{n})$ and $k<16n+8=O(\sqrt{n})$. We write the blocks

$$\{m,1,-m,1,m,2,-m,2,m\dots,m,4r-1,-m,4r-1,m,0\}$$

for $m=4r,12r,\dots,4r(2r-1)$ and the block

$$\{8r,\quad 0,\quad 16r,\quad 0,\quad \dots\quad 8r(r-1),\quad 0,\quad 8r^2,\quad 0,\quad 0\}.$$

So we needed less than $8r=O(\sqrt{n})$ numbers to cover all the integers in $[-8r^2,8r^2]$. For $k>0$ we add a further block:

$$\{8r^2+1,\quad 0,\quad 8r^2+2,\quad 0,\quad \dots,\quad 8r^2+k,\quad 0\}.$$

We needed to add $k=O(\sqrt{n})$ new integers to the previous ones to cover all the integers in $[-8r^2+k,8r^2+k]$.

PROBLEMS. (a) What is the minimum of $\|A\|$ if G is not cyclic?

(b) In general: Determine a lower bound c_1 and an upper bound c_2 for $\|A\|$. Let k be an integer in $[c_1,c_2]$. Is there a sequence (a) such that $\|A\|=k$?

(Received December 19, 1995)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT
URBANA-CHAMPAIGN
273 ALTGELD HALL
1409 WEST GREEN STREET
URBANA, IL 61801
U.S.A.

Present address:

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ALGEBRA ÉS SZÁMELMÉLET TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

pkovacsk@cs.elte.hu

Studia Scientiarum Mathematicarum Hungarica 33 (1997), 459–476

STABILITY OF STOCHASTIC POPULATION MODELS

M. H. VELLEKOOP and G. HÖGNÄS

Abstract

In this paper we derive sufficient conditions for the stability of a wide class of stochastic population models in discrete time. The class we study is the largest possible in some sense if one takes into account two fundamental biological premises which population models have to satisfy. The conditions for this stochastic stability that we obtain are of some interest, since studying certain statistical characteristics of these stochastic population processes is only possible if the process converges to an invariant distribution.

1. Introduction

One of the most fundamental recent discoveries in theoretical ecology is the possibility of extremely complex dynamics in even the simplest single-species population models [7]. This observation is particularly important since in most models some parameters, especially the ones which describe the influence of an environment on the population, vary around a certain fixed value in a stochastic way. This may pose a serious problem for the study of population models, since large scale stochastic simulations have to be used to study the variations but it is not clear on beforehand (and in some cases indeed not true) that these simulation results are valid, unless some 'stochastic stability' conditions are met. Only then can we guarantee that certain statistical properties converge to stationary values, which can then be found by appropriate measurements. Some of the stability questions for stochastic population models have been addressed earlier [1, 2] but the results are only valid for one particular model due to Ricker [10]. In this paper we will show that the results obtained there are a special case of conditions for a more general model that we will derive here. Indeed, we try to show that stochastic stability is an intrinsic characteristic of *all* population models which satisfy two fundamental biological premises.

This paper was presented at the Conference on Stochastic Differential and Difference Equations, Győr, Hungary, August 1996.

1991 *Mathematics Subject Classification.* Primary 60F10.

Key words and phrases. Markov processes, recurrence, invariant measures, population dynamics.

0081–6906/97/$ 5.00

2. A general population model

In [3] Hassell has introduced a systematic approach to the development of so-called density-dependent population models for a single species of the form

$$X_{t+1} = g(X_t).$$

Here X_t represents the population size at time t and g is a known function which is usually assumed to be continuous. Hassell argued that these population models should be based on two fundamental premises:

(P1) The population should have the *potential* to increase exponentially for small populations.

(P2) There should be a density-dependent feedback which reduces the *actual* rate of increase as the population grows.

Ecologists usually look at the 'mortality' as a measure of this density-dependent feedback, which is characterized by the following *density dependence function*:

$$\ln X_t \to \ln \frac{X_t}{X_{t+1}}. \tag{2.1}$$

In experimental data [3], it is often found that there is a pronounced density-dependence for large X_t which becomes negligible as X_t decreases, and that the density-dependence for large X_t is characterized either by a fixed slope b (in so called *contest models*) or that this slope increases rapidly with increasing X_t (in *scramble models*). A simple linear contest model describing this situation is for example

$$\ln \frac{X_t}{X_{t+1}} = -\ln r + b \ln X_t \tag{2.2}$$

with $r > 1$ and $b > 1$, resulting in

$$X_{t+1} = r X_t^{(1-b)}.$$

Clearly, this model is not very realistic for small populations since it predicts that

$$\lim_{X_t \to 0} X_{t+1} = \infty,$$

which means that for very small populations there is an infinite capacity to grow. A simple correction, as proposed in [12] is:

$$X_{t+1} = \begin{cases} r\lambda X_t^{(1-b)} & X_t > X_c \\ \lambda X_t & X_t \leqq X_c \end{cases} \tag{2.3}$$

with $\lambda > 1$ a constant and X_c a critical population value. This means that there is either density-dependence as before for populations which are larger than the critical population, or exponential growth for populations which are smaller than the critical population. However, this model is not smooth at X_c and has some other properties which are unsatisfactory from an ecological point of view [3].

Two famous examples exist of models which 'smoothen' this model. The first one is the *Hassell model* [3]:

$$X_{t+1} = \frac{rX_t}{(1+X_t)^b}$$

with $r > 1$ and $b > 1$. It is clear that for this model

$$X_{t+1} \approx \begin{cases} rX_t & X_t \approx 0 \\ rX_t^{(1-b)} & X_t \gg 0 \end{cases}$$

and it can thus indeed be interpreted as a smoothed version of model (2.3). Note that the density dependence function is given by

$$\ln \frac{X_t}{X_{t+1}} = -\ln r + b\ln(1+X_t)$$

which means that the density dependence relation (2.1) will be approximately linear with fixed slope b for large X_t. As mentioned before, we also need *scramble* models in which this slope increases rapidly as the population X_t increases. An example of such a model is due to *Ricker* [10]:

$$X_{t+1} = rX_t e^{-bX_t}.$$

It satisfies

$$\ln \frac{X_t}{X_{t+1}} = -\ln r + be^{\ln X_t},$$

so the density dependence increases exponentially when X_t goes to infinity. Remark that in all these models r represents the exponential growth factor for very small populations, the *natural* growth in ideal circumstances, and b the density-dependent feedback because of limited environmental conditions.

In order to smoothen the model (2.3) we propose a more general one:

$$\ln \frac{X_t}{X_{t+1}} = -\ln r + b\ln f(X_t) \tag{2.4}$$

with $r > 1$, $b > 0$ and in which the function $f : [0,\infty[\to \mathbb{R}$ is called the *density function*. We require it to satisfy:

(A1) $f : [0,\infty[\to \mathbb{R}$ is strictly positive and continuously differentiable on its domain.

(A2) $f(0) = 1$.

(A3) The function $L\colon [0,\infty[\to \mathbb{R}$ defined by $L(x) = \dfrac{xf'(x)}{f(x)}$ is strictly increasing.

Condition (A1) guarantees that the transitions are indeed smooth. Condition (A2) makes sure that the model satisfies the first fundamental assumption in ecological models (P1): for small populations exponential growth should be possible

$$X_t \approx 0 \qquad \Rightarrow \qquad X_{t+1} \approx rX_t.$$

Remark that it can always be satisfied by choosing suitable r and b, as long as $r > 1$, $b > 0$. The last condition (A3) represents the second fundamental assumption in population models (P2). It guarantees that for $b > 0$ the density dependence increases for increasing population size since the slope of the density is:

$$\frac{d \ln \frac{X_t}{X_{t+1}}}{d \ln X_t} = bL(X_t).$$

As mentioned in the preceding section, models with $\lim_{x\to\infty} L(x) < \infty$ are usually called *contest models* while models where L has no finite limit are called *scramble models*. We will see that there is an important difference between the two types when we consider stochastic perturbations of the models.

Note that the Hassell model and the Ricker model are both special cases of our general framework, with $f(x) = 1 + x$ and $f(x) = e^x$, respectively. The first one is a typical contest model and the second one a typical scramble model. Also remark that the linear model (2.2) does *not* belong to our class of models since it violates the second condition: it does not permit exponential growth for small population sizes.

Our general model can be rewritten in the transition form

$$X_{t+1} = g(X_t) := rX_t \left[f(X_t)\right]^{-b}, \tag{2.5}$$

which clearly shows the capacity for exponential growth (with exponential growth factor $r > 1$) and the density-dependent feedback that reduces growth (and of which the intensity is characterized by the positive parameter b). The model (2.4) together with the conditions on f can thus be seen as a general population model for non-overlapping generations which unifies models like the Hassell and Ricker models, taking into account the premises of ecological population models and the experimental observations on density-dependence (see for example [3]).

In the following lemma some elementary but useful properties of the functions f satisfying (A1)–(A3) and g are derived, that will be needed later on.

LEMMA 2.1. *All f satisfying* (A1)–(A3) *and g as defined in equation* (2.5) *derived from these satisfy:*

1. *f is strictly increasing and $f(x) \geqq 1$ for all $x \geqq 0$.*
2. *For all $x > 0$:*

$$\ln f(e^x) \leqq f(1) - 1 + xL(e^x). \tag{2.6}$$

3. *If the function L has a finite limit $\bar{L} = \lim_{x\to\infty} L(x)$ then*

$$\lim_{x\to\infty} x^{-1} \ln f(e^x) = \bar{L}.$$

PROOF. 1. Since $L(0)=0$ and $L(x)$ is strictly increasing, we have $L(x)>0$ for $x > 0$ and since f is strictly positive, condition (A3) implies that f' is strictly positive. Therefore we have that $f(x) \geqq f(0) = 1$.

2. Since $\ln f(0) = 0$ we have

$$\ln f(e^x) = \int_0^{e^x} \frac{f'(s)}{f(s)}\, ds = \int_0^1 \frac{f'(s)}{f(s)}\, ds + \int_1^{e^x} \frac{f'(s)}{f(s)}\, ds. \tag{2.7}$$

Because of the first part of the lemma we find

$$\int_0^1 \frac{f'(s)}{f(s)}\, ds \leqq \int_0^1 f'(s)\, ds = f(1) - f(0) = f(1) - 1$$

and for the second integral

$$\int_1^{e^x} \frac{f'(s)}{f(s)}\, ds = \int_1^{e^x} \frac{L(s)}{s}\, ds = \int_0^x L(e^\tau)\, d\tau \leqq xL(e^x), \tag{2.8}$$

where we have used condition (A3) in deriving the last inequality.

3. We use de l'Hospital's rule:

$$\lim_{x\to\infty} \ln f(e^x)/x = \lim_{x\to\infty} f'(e^x)e^x/f(e^x) = \lim_{x\to\infty} L(e^x). \qquad \square$$

The general population model (2.4) has two free parameters once the function f has been specified: a *natural growth* parameter $r > 1$ which is the exponential growth factor in an ideal environment and an environmental parameter $b > 1$ which represents the rate of density-dependent feedback of the environment. We now want to consider population models in which the environment (that is, the parameter b) varies *stochastically*. That is, we now take

$$X_{t+1} = G(X_t, \gamma_t) = \frac{r X_t}{f(X_t)^{\gamma_t}} \tag{2.9}$$

in which $X_0 > 0$ can be chosen freely and $\{\gamma_t \mid t \in \mathbb{N}\}$ is a set of independent identically distributed random variables with a distribution function which satisfies some mild conditions that will be specified later on. Remark that this implies that $\{X_t \mid t \in \mathbb{N}\}$ is a stochastic Markov process. We will denote the probability space for this process by $(\Omega, \mathcal{F}, P)$ with Ω the sample space, $\mathcal{F}$ an appropriate σ-field and P the probability measure. We use the notation $\mathbb{P}(A) = P(\{\omega \mid A(\omega)\})$ and take the symbol $X \sim \pi$ to state that the stochastic variable X has probability distribution π. The symbol $\mathbb{E}$ denotes the usual expectation operator. The indicator function of a set A will be denoted by $\mathbf{1}_A$; it has the value one on A and the value zero outside A.

We will now address the question under what conditions there exists an invariant distribution π, that is, a distribution satisfying

$$X_t \sim \pi \qquad \Rightarrow \qquad X_{t+1} \sim \pi.$$

It is easy to see that an initial distribution of this kind does not change in time and it is therefore important if we are interested in the limiting behaviour of our stochastic population process. The invariant distribution represents a 'stochastic equilibrium' and if the probability distribution of the process converges to a stationary distribution, this distribution has to be invariant. Moreover, studying certain statistical properties of the distribution function of the population only makes sense if the distribution is stationary in time.

We will now state some preliminary results on discrete time Markov processes on a continuous state-space that we will need in the sequel. All results are stated without proofs. These can be found in standard textbooks about the theory of Markov processes, for example [11] or the references in [1, 2]. We are interested in *transition probabilities*

$$\mathbb{P}\left(X_t \in A \mid X_0 = x\right)$$

for $x \in]0, \infty[$ and A a Borel-set. When the distribution of the noise $\{\xi_t \mid t \in \mathbb{N}\}$ has a positive density everywhere on $\mathbb{R}^+$, we have

$$\mathbb{P}\left(X_t \in A \mid X_0 = x\right) = \int_A p_t(\tau \mid x)\, d\tau$$

with $p_t(\tau \mid x)$ the *t-step transition density* which is strictly positive for all values of x and $\tau \in]0, xr^t[$. This implies that the process is *Lebesgue irreducible*: every set in the state space with positive Lebesgue measure can be reached from any initial value X_0. Furthermore, we may conclude that the process is *aperiodic*: there are no cyclic subsets in the state space. For h a bounded measurable function on the state space, the mapping

$$h \to \mathbb{E}[h(X_1) \mid X_0 = x]$$

is called the *transition probability operator* of the process. Under our assumptions, the transition probability operator maps bounded continuous functions to bounded continuous functions.

An irreducible Markov process is called *recurrent* if the expectation of the number of returns to A is infinite:

$$\sum_{t=1}^{\infty} \mathbb{P}(X_t \in A \mid X_0 = x) = \infty$$

for all sets A with positive Lebesgue measure. Otherwise the process is called *transient*. Every recurrent process has a σ-finite invariant measure π, satisfying

$$\pi(A) = \int_{\mathbb{R}^+} \mathbb{P}(X_t \in A \mid X_0 = x)\, \pi(dx)$$

for any Borel set A. If the measure π is finite:

$$\int_{\mathbb{R}^+} \pi(dx) < \infty$$

and we can therefore make it a probability measure by normalizing it, the process is called *positively recurrent*, otherwise it is called *null-recurrent*. In the latter case there exists at least one set A with positive Lebesgue measure such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \mathbb{P}(X_t \in A \mid X_0 = x) = 0.$$

It is clear from these results that it is interesting from a biological point of view to know whether a certain biological process is positively recurrent or not: if we are interested in an invariant probability measure, we must first prove positive recurrence. We will use the following theorem to obtain results for our class of stochastic population models (see [9, 11]):

THEOREM 2.1. *Suppose that a Markov process* $\{X_t \mid t \in \mathbb{N}\}$ *is Lebesgue irreducible and that its one step transition probability maps continuous functions to continuous functions. If there exists a compact set* K*, an* $\epsilon > 0$*,* $C > 0$ *and a function* V *such that*

1. $(\forall x \in \mathbb{R})$ $V(x) \geqq 0$,
2. $(\forall x \in K^c)$ $\mathbb{E}[V(X_1) \mid X_0 = x] - V(x) < -\epsilon$,
3. $(\forall x \in K)$ $\mathbb{E}[V(X_1) \mid X_0 = x] \leqq C$

then the process $\{X_t \mid t \in \mathbb{N}\}$ *is positively recurrent.*

The conditions of the last theorem are called *mean drift conditions* and they roughly state that the process $V(X_t)$ is a supermartingale outside some compact set, with an expectation which decreases uniformly. The function V is often called a *Stochastic Lyapunov Function.* Note that the Lyapunov function is not required to be continuous.

3. Sufficient conditions for stationarity

Using the results of the last section we will now give sufficient conditions for the stochastic population models to have an invariant probability measure. We will assume, as stated earlier, that the model has the form

$$X_{t+1} = G(X_t, \gamma_t) = \frac{rX_t}{f(X_t)^{\gamma_t}} \tag{3.1}$$

and that the following conditions are satisfied:

(B1) $X_0 > 0$ (initial population not trivial) and $r > 1$ (there is natural growth),

(B2) $\{\gamma_t \mid t \in \mathbb{N}\}$ are independent identically distributed stochastic variables with finite first and second order moments,

(B3) the random variables $\{\gamma_t \mid t \in \mathbb{N}\}$ have an absolutely continuous distribution, with a probability density which is positive on the whole $\mathbb{R}^+$.

Under these assumptions, the transition probability operator maps continuous functions to continuous functions. Indeed, if h is a bounded continuous function on $\mathbb{R}^+$ and ϕ is the distribution of the noise γ then

$$\mathbb{E}[h(X_1) \mid X_0 = x] = \int_{\mathbb{R}} h(rxf(x)^{-\gamma})\phi(d\gamma)$$

and this is continuous in x by Lebesgue's dominated convergence theorem, since f is continuous.

Our result is split into two separate theorems, one for the contest models and one for the scramble models, since the analysis for these two cases is different.

THEOREM 3.1. *Consider the stochastic process given by* (3.1) *satisfying conditions* (B1)–(B3) *and with the function* f *satisfying conditions* (A1)–(A3). *If*

$$\lim_{x\to\infty} L(x) = \bar{L} < \infty$$

and

$$\mathbb{E}\left((1-\bar{L}\gamma_t)\mathbf{1}_{\{\gamma_t \leqq \frac{1}{\bar{L}}\}}\right) < 1 \tag{3.2}$$

then there exists a unique invariant probability distribution π *on* $\mathbb{R}^+$ *for this process.*

PROOF. Define the process $Y_t = \ln X_t$, then we have that $\{Y_t \mid t \in \mathbb{N}\}$ positively recurrent if and only if $\{X_t \mid t \in \mathbb{N}\}$ positively recurrent, and obviously

$$Y_{t+1} = Y_t + \ln r - \gamma_t \ln f(e^{Y_t}). \tag{3.3}$$

Since the probability density is positive on the whole $\mathbb{R}^+$ and $\ln r > 0$ the process is irreducible. We will therefore prove that the function $V : \mathbb{R} \to \mathbb{R}^+$ defined by

$$V(y) = \begin{cases} y, & y \geqq 0 \\ -ay, & y < 0 \end{cases}$$

with $a > 0$ a suitably chosen constant that we will specify later on, is a Lyapunov function for the process $\{Y_t\}$. We introduce the functions

$$B(y) = \frac{y + \ln r}{\ln f(e^y)}$$

and

$$p(y) = \mathbb{P}(\gamma_t \leqq B(y)) = \mathbb{P}(Y_{t+1} \geqq 0 \mid Y_t = y).$$

Because of Lemma 2.1 we have

$$\lim_{y\to\infty} B(y) = 1/\bar{L}. \tag{3.4}$$

It is also easy to establish that we must have

$$\lim_{y\to-\infty} p(y) = 0. \tag{3.5}$$

Evaluating the expected growth of V in a point y we find

$$\begin{aligned}\mathbb{E}[V(Y_1) \mid Y_0 = y] - V(y) &= \mathbb{E}(y + \ln r - \gamma_t \ln f(e^y))\mathbf{1}_{\{\gamma_t \leqq B(y)\}} \\ &\quad + \mathbb{E}(-ay - a\ln r + a\gamma_t \ln f(e^y))\mathbf{1}_{\{\gamma_t > B(y)\}} - V(y).\end{aligned}$$

Rearranging terms gives

$$\begin{aligned}
&\mathbb{E}[V(Y_1)\,|\,Y_0=y]-V(y)\\
&=(y+\ln r)p(y)-(\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t\leqq B(y)\}})\ln f(e^y)\\
&+(-ay-a\ln r)(1-p(y))+a(\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t>B(y)\}})\ln f(e^y)-V(y).
\end{aligned}\tag{3.6}$$

Now for $y\ll 0$ the first term will become equal to or smaller than zero because of (3.5), the second and the fourth term both go to zero since $\ln f(e^y)\to 0$ as $y\to-\infty$ and the expectations have upper bounds which do not depend on y. The third term is asymptotically equal to $-ay-a\ln r$ and the last term is equal to $-(-ay)=ay$. Therefore we have

$$\mathbb{E}[V(Y_1)\,|\,Y_0=y]-V(y)<-\tfrac{1}{2}a\ln r,\qquad (y\ll 0).\tag{3.7}$$

For $y\gg 0$ we find, using equation (3.6) and the fact that in this case $V(y)=y$, that

$$\begin{aligned}
\mathbb{E}[V(Y_1)\,|\,Y_0=y]-V(y)&=ay\left(p(y)-1+\frac{\ln f(e^y)}{y}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t>B(y)\}}\right)\\
&+y\left(p(y)-1-\frac{\ln f(e^y)}{y}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t\leqq B(y)\}}\right)\\
&+p(y)\ln r-a(1-p(y))\ln r.
\end{aligned}$$

The last two terms have an upper bound which is independent of y. Because of dominated convergence and Lemma 2.1 the bracketed expression in the second term will converge as $y\to\infty$:

$$p(y)-1-\frac{\ln f(e^y)}{y}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t\leqq B(y)\}}\overset{y\to\infty}{\longrightarrow}\mathbb{E}(1-\bar{L}\gamma_t)\mathbf{1}_{\{\gamma_t\leqq 1/\bar{L}\}}-1$$

which implies, because of (3.2) that for y big enough, the second term will be smaller than $-\nu|y|$ for some $\nu>0$. Analogously, the bracketed expression in the first term converges according to the dominated convergence theorem:

$$p(y)-1+\frac{\ln f(e^y)}{y}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t>B(y)\}}\overset{y\to\infty}{\longrightarrow}-\mathbb{P}(\gamma_t>1/\bar{L})+\bar{L}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t>1/\bar{L}\}},$$

so for y sufficiently large it will become smaller than the positive value $D:=\bar{L}\mathbb{E}\gamma_t\mathbf{1}_{\{\gamma_t>1/\bar{L}\}}$. Therefore, if we take $a<\frac{1}{2}\nu/D$, the first term will become smaller than $\frac{1}{2}\nu y$, and therefore we find for the total in (3.6):

$$\mathbb{E}[V(Y_1)\,|\,Y_0=y]-V(y)<\hat{C}-\tfrac{1}{2}\nu y\tag{3.8}$$

for y big enough, with $\hat{C}$ a constant which is independent of y. Taking (3.7) and (3.8) together we see that there exist positive y_1, y_2 such that if we take $K=[-y_1,y_2]$ we have for $y\in\mathbb{R}\setminus K$:

$$\mathbb{E}[V(Y_1)\,|\,Y_0=y]-V(y)<-\epsilon$$

with

$$\epsilon = \min\{\frac{1}{2}\nu y_2 - \hat{C},\ \frac{1}{2}a \ln r\}.$$

This shows that the conditions of Theorem 2.1 are satisfied since it is easy to see, using (3.6), that for every $y \in K$ we have

$$\mathbb{E}[V(Y_1) \mid Y_0 = y] \leqq y_2 + \ln r + (1+a)|\mathbb{E}\gamma_t| \ln f(e^{y_2}) + ay_1 + y_1 := C.$$

The result now follows from Theorem 2.1. □

We have thus found a sufficient condition for the existence of a stationary distribution in contest models. Remark that this condition will automatically be satisfied if the probability density of the environment γ_t is taken small enough (or even zero) for negative values since then

$$\mathbb{E}(1 - \bar{L}\gamma_t)\mathbf{1}_{\{\gamma_t \leqq \frac{1}{\bar{L}}\}} \approx \mathbb{E}(1 - \bar{L}\gamma_t)\mathbf{1}_{\{0 \leqq \gamma_t \leqq \frac{1}{\bar{L}}\}} < 1.$$

The following example illustrates this.

EXAMPLE 3.1. *Contest models with Gaussian distributed environments.* Suppose we have a stochastic contest model where the $\{\gamma_t \mid t \in \mathbb{N}\}$ are independent Gaussian random variables with mean μ and variance σ^2. Then the process X_t as defined in (3.1) has a unique invariant probability measure when

$$\mu > \frac{1}{\bar{L}}, \qquad \sigma < \frac{\sqrt{2\pi}}{\bar{L}}.$$

Indeed, evaluating the conditional expectation in (3.2) gives:

$$\int_{-\infty}^{\frac{1}{\bar{L}}} \frac{(1 - \bar{L}x)}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \leqq \bar{L} \int_{-\infty}^{\frac{1}{\bar{L}}} \frac{1/\bar{L} - x}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\bar{L}^{-1}}{\sigma})^2} dx.$$

Introducing a new integration variable $s = \frac{1}{\bar{L}} - x$ we find that the conditional expectation is smaller than

$$\frac{\bar{L}}{\sigma\sqrt{2\pi}} \int_0^\infty s e^{-\frac{1}{2}(\frac{s}{\sigma})^2} ds = \frac{\sigma\bar{L}}{\sqrt{2\pi}} < 1$$

which proves that there exists indeed an invariant probability measure for this case, since the other requirements of the theorem are trivially satisfied for Gaussian random variables. □

Having established a sufficient condition for stationarity when L has a finite limit, we now turn to the scramble models, where L diverges to infinity. The following theorem is an extension of results in [1], where a proof is given for the specific case of the Ricker model.

THEOREM 3.2. *Consider the stochastic process* (3.1) *satisfying conditions* (B1)–(B2) *and with the function* f *satisfying conditions* (A1)–(A3). *If*

$$\lim_{x\to\infty} L(x) = \infty$$

and the $\{\gamma_t \mid t \in \mathbb{N}\}$ *have a probability density function which is positive on* $]0,\infty[$, *zero on* $]-\infty, 0[$, *but with possibly an atom at the origin satisfying*

$$\mathbb{P}(\gamma_t = 0) < \left(\limsup_{x\to\infty} \frac{L(rx)}{L(x)}\right)^{-1} \tag{3.9}$$

then there exists a unique invariant probability measure π *on* $\mathbb{R}^+$ *for this process.*

PROOF. We use the Lyapunov function

$$V(y) = \begin{cases} yL(e^y) & y \geqq 0 \\ -ay & y < 0 \end{cases}$$

for the process $Y_t := \ln X_t$. Here $a > 0$ will be specified later on. We have

$$\begin{aligned} &\mathbb{E}[V(Y_1) \mid Y_0 = y] - V(y) \\ &= \mathbb{E}(y + \ln r - \gamma_t \ln f(e^y))L(e^{y+\ln r - \gamma_t \ln f(e^y)})\mathbf{1}_{\{\gamma_t \leqq B(y)\}} \\ &\quad + \mathbb{E}(-ay - a\ln r + a\gamma_t \ln f(e^y))\mathbf{1}_{\{\gamma_t > B(y)\}} - V(y). \end{aligned} \tag{3.10}$$

For $y \ll 0$ this will become smaller than $-\frac{1}{2}a\ln r$ for $y < -y_1$ with $y_1 > 0$ large enough, since $\gamma_t \geqq 0$, $L > 0$ and $\ln f(e^y) \to 0$ as $y \to -\infty$ together imply that the above expression is asymptotically equivalent to

$$(-ay - a\ln r)(1 - p(y)) - V(y) \approx -ay - a\ln r - (-ay)$$

as y goes to $-\infty$. For $y \gg 0$ we find that (3.10) is smaller than

$$\begin{aligned} &(y + \ln r)L(e^{y+\ln r})p(y) - a(y + \ln r)(1 - p(y)) \\ &+ a|\mathbb{E}\gamma_t \mathbf{1}_{\{\gamma_t > B(y)\}}|\,[f(1) - 1 + yL(e^y)] - yL(e^y) \end{aligned}$$

because of (2.6). Rearranging gives

$$\begin{aligned} &yL(e^y)[p(y)\frac{L(re^y)}{L(e^y)}(1 + \frac{\ln r}{y}) - 1] + ay[p(y) - 1 + |\mathbb{E}\gamma_t \mathbf{1}_{\{\gamma_t > B(y)\}}|L(e^y)] \\ &+ a[(f(1) - 1)|\mathbb{E}\gamma_t \mathbf{1}_{\{\gamma_t > B(y)\}}| + \ln r(p(y) - 1)]. \end{aligned}$$

Since $\lim_{x\to\infty} L(x) = \infty$ we have that $\lim_{y\to\infty} B(y) = 0$ so

$$\lim_{y\to\infty} p(y) = \lim_{y\to\infty} \mathbb{P}(\gamma_t \leqq B(y)) = \mathbb{P}(\gamma_t = 0)$$

so for $y \geqq y_2$, with $y_2 > 0$ large enough we have, because of (3.9)

$$yL(e^y)[p(y)\frac{L(re^y)}{L(e^y)}(1+\frac{\ln r}{y})-1] < -\nu y L(e^y)$$

for a certain positive ν and choosing

$$a < \frac{\nu}{2(1+|\mathbb{E}\gamma_t \mathbf{1}_{\{\gamma_t>0\}}|)}$$

gives

$$\mathbb{E}\,[V(Y_1)\,|\,y_0=y] - V(y) < \hat{C} - \frac{1}{2}\nu y L(e^y)$$

with $\hat{C}$ a constant which is independent of y. So again, if we take $K = [-y_1, y_2]$ the conditions of Theorem 2.1 are satisfied outside K and on K we find

$$\begin{aligned}\mathbb{E}\,[V(Y_1)\,|\,Y_0=y] = {} & \mathbb{E}\ (y+\ln r-\gamma_t \ln f(e^y))\ L(e^{y+\ln r-\gamma_t \ln f(e^y)})\mathbf{1}_{\gamma_t \leqq B(y)}\\ & -\mathbb{E}\,(ay+a\ln r - a\gamma_t \ln f(e^y))\mathbf{1}_{\{\gamma_t > B(y)\}}\\ & \leqq (y_2+\ln r)L(re^{y_2}) + ay_1 + ay_2\mathbb{E}\gamma_t \ln f(e^{y_2})\end{aligned}$$

and thus the theorem has been proven. □

Remark that there is a fundamental difference between the class of stochastic processes we use in these two results: in the second one we assume that γ_t can *only* attain nonnegative values.

Now that we have found sufficient and almost necessary conditions for the existence of an invariant probability measure for the stochastic population models, we will address the question of stochastic stability. That is, we would like to know under what conditions an arbitrary initial probability distribution on the state space converges to this invariant probability distribution and what can be said about the speed of convergence. This is of some importance, because if we cannot guarantee that an arbitrary initial population will converge to a stationary random variable in time, some attempts to measure statistical properties of the population over time may be useless. Moreover, we can only compute the invariant distribution numerically by iterating an arbitrary initial distribution and hoping that this converges to the invariant one.

Since we want to prove a stronger result than simple pointwise convergence of the distributions, we introduce the following *total variation norm* for signed measures ν on $\mathbb{R}^+$:

$$\|\nu\| = \sup_{\{g:|g|\leqq 1\}} |\nu(g)| = \sup_{A\in\mathcal{B}(\mathbb{R}^+)} \nu(A) - \inf_{A\in\mathcal{B}(\mathbb{R}^+)} \nu(A)$$

with $\mathcal{B}(\mathbb{R}^+)$ the Borel sets in $\mathbb{R}^+$. If we define P_x^t to be the probability measure of X_t given the initial condition $X_0 = x$, and π the stationary probability measure, we can now state precisely the property we would like to establish:

DEFINITION 3.1. The Markov process $\{X_t \mid t \in \mathbb{N}\}$ with invariant probability distribution π is called *ergodic* if for every $x \in \mathbb{R}^+$

$$\lim_{t\to\infty} \|P_x^t - \pi\| = 0.$$

The question of stochastic stability is completely resolved by the following result, which is a special case of results proven in [9]:

LEMMA 3.1. *Suppose that the irreducible Markov process* $\{X_t \mid t \in \mathbb{N}\}$ *admits a stationary probability distribution and is* Harris recurrent, *that is, it visits every Borel set in its domain infinitely often:*

$$(\forall A \in \mathcal{B}(\mathbb{R}^+))\ (\forall X_0 \in A) \qquad \mathbb{P}\Big(\sum_{t=1}^{\infty} \mathbf{1}_{\{X_t \in A\}} = \infty\Big) = 1.$$

Then the process is ergodic.

Since conditions (B1)–(B3) imply already that every positive Borel set will be visited infinitely often, we immediately have from Lemma 3.1:

COROLLARY 3.1. *All processes* (3.1) *satisfying conditions* (B1)–(B3) *and with the function f satisfying conditions* (A1)–(A3) *which admit a stationary probability measure, are ergodic.*

So in all cases where we have proven that a stationary distribution exists, every initial population will converge to a random variable having this distribution. That is, the population models are stochastically stable.

4. Examples of applications

We will now illustrate our results by some numerical examples.

EXAMPLE 4.1.. *Stochastic stability of a contest model.* We consider a stochastic perturbation of the Hassell model. We take the environmental parameters identical independently distributed Gaussian random variables with mean 3 and variance 1, and a natural growth rate $r = 2$:

$$X_{t+1} = \frac{rX_t}{(1+X_t)^{\gamma_t}}, \qquad r = 2, \qquad \gamma_t \sim \text{i.i.d. } N(3,1).$$

For this model we have

$$L(x) = \frac{x}{x+1} \quad \Rightarrow \quad \bar{L} = \lim_{x\to\infty} L(x) = 1.$$

So for this particular choice of the environmental random variables, condition (3.2) for the existence of an invariant distribution becomes

$$\mathbb{E}\,(1-\bar{L}\gamma_t)\mathbf{1}_{\{\gamma_t \leqq \frac{1}{L}\}} \approx 0.0084 \ll 1$$

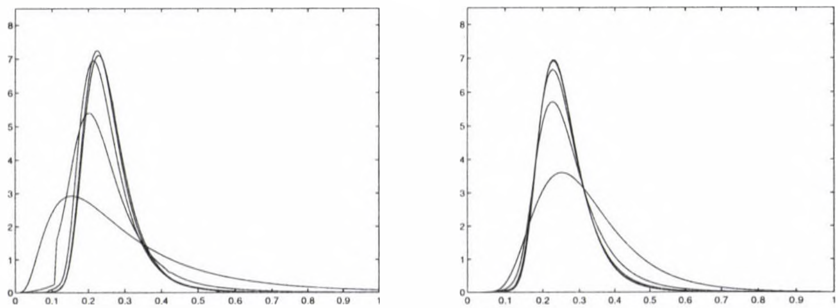
and therefore the distribution of the stochastic process $\{X_t\}$ should converge to a stationary distribution π. Figure 1a shows the distribution of X_t for $t=1,2,\ldots,5$ when we take the initial condition $X_0=1$, and Figure 1b when $X_0=0.5$. We see that the distribution converges indeed to the same stationary distribution in both cases. Taking the same stochastic perturbation model but now with environmental variables which have zero mean:

$$X_{t+1}=\frac{rX_t}{(1+X_t)^{\gamma_t}}, \qquad r=2, \qquad \gamma_t \sim \text{ i.i.d. } N(0,1)$$

we find

$$\mathbb{E}\,(1-\bar{L}\gamma_t)\mathbf{1}_{\{\gamma_t \leqq \frac{1}{L}\}} \approx 1.083 > 1$$

so our condition is not satisfied here. Looking at the distribution functions in Figure 2, we see that in this case the system is not stochastically stable. The initial distribution 'wanders off' in the positive direction, and it will eventually spread out over the entire positive axis, while converging pointwise to zero in every single point. These two examples clearly show the importance of our conditions for stochastic stability in the analysis of populations in stochastic environments.



(a) $X_0=1.0$ (b) $X_0=0.5$

Fig. 1. Hassell model, $r=2$, $b\sim N(3,1)$

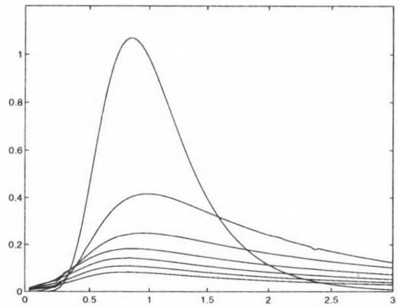


Fig. 2. Hassell model, $r=2$, $b\sim N(0,1)$, $X_0=0.5$

EXAMPLE 4.2. *Stochastic stability of a scramble model.* We now take a look at a typical stochastic perturbation of a scramble model:

$$X_{t+1}=2X_t e^{-X_t\gamma_t}, \qquad \begin{array}{c}\mathbb{P}(\gamma_t=0)=p_0\\ \mathbb{P}(\gamma_t=|Z_t|)=1-p_0\end{array}$$

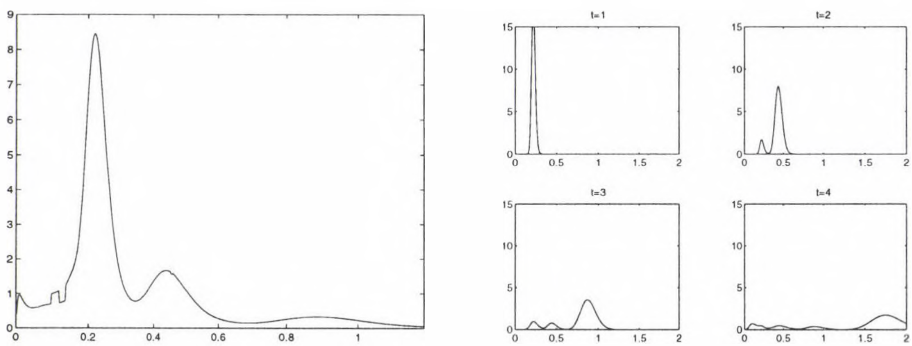
with $Z_t\sim$ i.i.d. $N(3,\frac{1}{2})$. For the Ricker model we have

$$f(x)=e^x \Rightarrow L(x)=x$$

and according to Theorem 3.2 this means that a stationary distribution exists if

$$p_0<\left(\limsup_{x\to\infty}\frac{L(rx)}{L(x)}\right)^{-1}=\left(\limsup_{x\to\infty}\frac{rx}{x}\right)^{-1}=\frac{1}{r}=0.5.$$

First the distributions for $p_0=0.2$ were calculated; after seven iterations the distribution function showed no significant changes any more and the stationary distribution of Figure 3a was obtained. Remark that peaks are found at distances which differ by a factor r due to the atom at zero in the environment which implies a population growth with factor r.



(a) $p_0=0.2<\frac{1}{r}$ (b) $p_0=0.8>\frac{1}{r}$

Fig. 3. Distributions for Ricker model, $r=2$

For $p_0 = 0.8$ we find the distribution functions of Figure 3b. The distribution spreads on the entire positive axis, with peaks travelling to infinity due to multiplication by r in every iteration, while the distribution converges to zero in every single point.

5. Conclusions

In this paper we have addressed the question of stability in single species discrete population models. A sufficient condition for stability was derived, which turned out to be satisfied in all cases where the environmental influence was purely adversary, but also if there was a limited possibility of favourable environments in the contest case.

The most surprising aspect of the results presented here is their very general nature. The two intuitively appealing premises stated in the first paragraph apparently contain sufficient information about the 'nature' of the dynamics to prove stability in the stochastic model for favourable environments. Therefore one may state that the possibility of a stable stationary probability distribution in their stochastic perturbations is a consequence of the ecological principles behind population models, and not of the details of a particular model. Since it is impossible in practice to describe any population exactly with one particular model, this should be a reassuring point for those who use population models to fit experimental data. If the stability of a stationary distribution of a stochastic population model can be proven for a particular model, but not for a model which is 'close' to this one, then the calculation of mean, variance and other statistical characteristics would be useless from a mathematical point of view. In this way the framework provided here can be used not only as a theoretically interesting generalization, but also as a rigorous justification for the use of population models to obtain theoretical statements about experimental data.

The question arises naturally if the conditions we have found are not only sufficient but even necessary for stochastic stability. The simulation results presented in the previous section clearly indicate that this might be true, and this topic is currently under investigation.

REFERENCES

[1] Gyllenberg, M., Högnäs, G. and Koski, T., Population models with environmental stochasticity, *J. Math. Biol.* **32** (1994), 93–108.

[2] Gyllenberg, M., Högnäs, G. and Koski, T., Null recurrence in a stochastic Ricker model, in: M. Gyllenberg and L. E. Persson (eds.): *Analysis, algebra and computers in mathematical research (Lulea, 1992)*, Proceedings of the Twenty-first Nordic Congress of Mathematicians, Lecture Notes in Pure and Appl. Math., **156**, Marcel Dekker, New York, 1994, 147–164. *MR* **95j**:60102

[3] Hassell, M. P., Density-dependence in single-species populations, *J. Animal Ecology* **44** (1974), 283–296.

[4] HASSELL, M. P., LAWTON, J. H. and MAY, R. M., Patterns of dynamical behaviour in single species population models, *J. Animal Ecology* **45** (1976), 471–486.
[5] KARLIN, T. and TAYLOR, H. M., *A first course in stochastic processes*, 2nd edition, Academic Press, New York, 1975. *MR* **50** #8668
[6] LATTO, J., The analysis of density dependence in insect populations, Department of Pure and Applied Biology, Imperial College, London, 1989 (Thesis).
[7] MAY, R. M., Simple mathematical models with very complicated dynamics, *Nature* **261** (1976), 459–467.
[8] MAY, R. M. and OSTER, G. F., Bifurcations and dynamic complexity in simple ecological models, *Amer. Nature* **110** (1976), 573–599.
[9] MEYN, S. P. and TWEEDIE, R. L., *Markov chains and stochastic stability*, Communications and Control in Engineering Series, Springer-Verlag London Ltd., London, 1993. *MR* **95j**:60103
[10] RICKER, W. E., Stock and recruitment, *J. Fisheries Research Board of Canada* **11** (1954), 559–623.
[11] TWEEDIE, R. L., Criteria for classifying general Markov chains, *Adv. Appl. Probab.* **8** (1976), 737–771. *MR* **56** #9695
[12] VARLEY, G. C., GRADWELL, G. R. and HASSELL, M. P., *Insect population ecology*, Blackwell Scientific Publications, Oxford, 1973.
[13] VELLEKOOP, M. H., Analysis of ecological systems: chaos, stochastics and optimization, Department of Applied Mathematics, University of Twente, Enschede, 1993 (internal report).
[14] VELLEKOOP, M. H. and HÖGNÄS, G., A unifying framework for chaos and stochastic stability in discrete population models, *J. Math. Biology* **35** (1997), 557–588.

(Received September 1, 1996)

CENTRE FOR PROCESS SYSTEMS ENGINEERING
IMPERIAL COLLEGE
EXHIBITION ROAD
LONDON SW7 2BT
UNITED KINGDOM

m.vellekoop@ic.ac.uk

DEPARTMENT OF MATHEMATICS
ÅBO AKADEMI
FIN–20500 ÅBO
FINLAND

ghognas@ra.abo.fi

BOOK REVIEWS

Logic and Algebra, Eds. A. Ursini and P. Aglianò, Proc. Conf. Siena, 1994, Lecture Notes in Pure and Applied Mathematics, vol. 180, Marcel Dekker, New York, 1996, 728 pp. ISBN 0 8247 9606 3. $ 175.00

This volume is the Proceedings of a conference held in memory of Roberto Magari (1934–1994) in Pontignano (Siena, Italy), April 26–30, 1994. The conference was originally planned to celebrate Magari's 60th birthday by putting together researchers from mathematical logic and general algebra. Here is the list of invited papers:

S. Artemov and A. Chuprina: Logic of proofs with complexity operators.

M. Comini and G. Levi: Beyond the s-semantics: A theory of observables.

D. Finberg, M. Mainetti and G.-C. Rota: The logic of commuting equivalence relations.

J.-Y. Girard: Proof-nets: The parallel syntax for proof-theory.

P. Hájek: Magari and others on Gödel's ontological proof.

L. Hendriks and D. de Jongh: Finitely generated Magari algebras and arithmetic.

J. Lambek: The butterfly and the serpent.

F. W. Lawvere: Adjoints in and among bicategories.

A. Macintyre: Exponential algebra.

R. McKenzie: An algebraic version of categorical equivalence for varieties and more general algebraic categories.

A. F. Pixley: Boolean universal algebra.

R. Wille: Restructuring mathematical logic: An approach based on Peirce's pragmatism.

G. Zappa: The development of research in algebra in Italy from 1850 to 1940.

In addition, the volume contains 20 research papers, mainly on model theory and universal algebra, including a posthumous paper of Magari, jointly with G. Simi, entitled 'A revision of the mathematical part of Magari's paper on "Introduction to metamorality"'.

Many papers in this volume are valuable contributions to mathematical logic, universal algebra, and category theory, and the volume is recommended to researchers in these fields.

L. Márki (Budapest)

Mathematical methods for scientists and engineers: Linear and nonlinear systems, Peter B. Kahn, Wiley Science Paperback Series, John Wiley and Sons, New York, 1990, XIX, 469 pp. with illustrations. ISBN 0 471 62305 9 (cloth), 0 471 16611 0 (pbk). £ 29,95

This book is meant as a one-year course in mathematical methods for advanced undergraduates or beginning graduate students in applied mathematics, physics and engineering. The first half of the book deals with linear systems to prepare the reader for the second half which deals with nonlinear systems, especially with weakly nonlinear oscillatory systems and nonlinear difference equations. It is very easy to read the book, since the author

Akadémiai Kiadó, Budapest

introduces the methods by giving examples with a lot of introductory, accompanying and concluding remarks.

It is not easy, however, to see the purpose of the author by writing of this book. For example, the text treats only the 2×2 matrices, so in the reviewer's opinion, somebody with a minimum knowledge of mathematics would not much learn from this book.

The content of the book is as follows.

Part I. Linear systems. — Matrix theory — The Gamma and related functions — Elements of asymptotics — Evaluation of sums: The Euler–Maclaurin sum expansion — Evaluation of integrals: The Laplace method — Differential equations.

Part II. Nonlinear systems. — The simple harmonic oscillator and the logistic equation — Aspects of harmonic motion and the concept of secular terms — Equilibrium points and the phase plane — Conservative systems — Nonconservative systems — The method of averaging — The method of multiple times scales — Higher-order calculations — Error analysis — One-dimensional iterative maps and the onset of chaos.

Appendix. A discussion of Euler's constant.

T. Fényes (Budapest)

Multigraph. Version 1.0c from IntelliPro, Wiley College Software, John Wiley and Sons, Inc., Chichester, 1996, ISBN 0 471 10618 6. £ 22,50

If you need a lightweight and easy-to-use software to make some mathematical analysis, Multigraph is a reasonable choice for you. This software is not for professional users and professional mathematicians, but for students (or anyone who likes maths and computer graphics). You can make nice and useful drawings of simple or complicated 2-variable functions with some simple steps.

It is a good idea that the software does not leave the user alone with a cryptic collection of menu points, but draws sample graphs with one mouse click, and from this state you can build the graphic you need by simply changing parameters step-by-step.

You can create 3-dimensional colour surface graphs, 2-dimensional vector fields even with a line integral along a user-specific path, or contour diagrams. You can set everything by an easy-to-use graphical interface and see the resulting graphics immediately. Pictures then can be printed out or saved — but only in a special format used by the software, so you cannot export them to other applications.

G. Sipőczy (Budapest)

Studia Scientiarum Mathematicarum Hungarica

Volume 33

Akadémiai Kiadó, Budapest

1997

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 33

CONTENTS

RECENTLY ACCEPTED PAPERS

SEN, K. and GUPTA, R., Time dependent analysis of T-policy $M/M/1$ queues — a new approach

KIYOMURA, J., KUSANO, T. and NAITO, M., Positive solutions of second order quasilinear ordinary differential equations with general nonlinearities

GUO, S. and QI, Q., On the simultaneous approximation of functions and their derivatives by the Gamma operator

MOLNÁR, E., PROK, I. and SZIRMAI, J., Classification of solid transitive simplex tilings. II. Metric realization of the maximal simplex tilings

WU, J. and LI, R., The hypocontinuity and uniform boundedness for bilinear maps

TÓTH, L., Asymptotic formulae concerning arithmetical functions defined by cross-convolution. III. On the function τ_k

CZLAPIŃSKI, T. and KAMONT, Z., Generalized solutions of local initial problems for quasilinear hyperbolic functional differential systems

ARGYROS, I. K., Approximating solutions of operator equations and applications using modified contractions

GUO, S. and QI, Q., Pointwise estimates for Bernstein-type operators

FÉNYES, T., Integral representations of the Laplace transform and moments of the modified Bessel function with respect to the order

EL MOUMNI, S., Optimal packings of unit squares in a square

CSÖRGŐ, S., Rényi confidence bands

MÓRI, T. F., Asymptotics of periodic permanents

CHUDAK, F. and GRIGGS, J., A new extension of Lubell's inequality to the lattice of divisors

MEGYESI, Z., Coverage probabilities of Rényi confidence bands

BAUER, C., On the exceptional set for the sum of a prime and the k-th power of a prime

Manuscripts should be submitted in duplicate, typed in double spacing on only one side of the paper with wide margins. Only original papers will be published and a copy of the Publishing Agreement will be sent to the authors of papers accepted for publication. Manuscripts will be processed only upon receipt of the signed copy of the agreement.

Authors are encouraged to submit their papers electronically. All common dialects of TEX are welcome. The electronic file of an article should always be accompanied by a hardcopy, the printed version exactly corresponding to the electronic one.

Figures should be submitted on separate sheets, drawn either in India ink at their expected final size, or as printouts and matching files processed electronically, preferably as encapsulated PostScript (EPS) ones.

CONTENTS

HU ISSN 0081-6906