Studia Scientiarum Mathematicarum Hungarica

VOLUME 34
NUMBERS 1–3
1998

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 1 318 2875 Fax: +36 1 317 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 August, 1998
Vol: 34 **Pages**: 1-400
Numbers: 1-3 **Whole**: 66

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 1–13

A LIMIT THEOREM FOR LACUNARY SERIES $\sum f(n_k x)$

I. BERKES[1] and W. PHILIPP

To the memory of Alfréd Rényi

Abstract

Let $f: R \to R$ be a Lebesgue measurable function satisfying

$$f(x+1) = f(x), \quad \int_0^1 f(x)dx = 0, \quad \int_0^1 f^2(x)dx = 1.$$

Several authors investigated the asymptotic properties of lacunary series $\sum c_k f(n_k x)$ under the Hadamard gap condition

$$n_{k+1}/n_k \geqq q > 1 \qquad (k = 1, 2, \dots)$$

and the behaviour of such series is well known. On the other hand, very little is known on the properties of $\sum c_k f(n_k x)$ if (n_k) grows slower than exponentially. The purpose of this paper is to prove an asymptotic result for such series.

1. Introduction

Let $f: R \to R$ be a Lebesgue measurable function satisfying

$$(1.1) \qquad f(x+1) = f(x), \qquad \int_0^1 f(x)dx = 0, \qquad \int_0^1 f^2(x)dx = 1.$$

The asymptotic properties of lacunary series $\sum c_k f(n_k x)$ have been investigated by many authors and are known to be very similar to those of independent random variables. For example, Takahashi proved ([13], [14]) that if f is a Lipschitz function satisfying (1.1) and (n_k) is a sequence of positive integers satisfying

$$(1.2) \qquad n_{k+1}/n_k \to \infty$$

1991 *Mathematics Subject Classification.* Primary 42A55, 60F15.
Key words and phrases. Lacunary series, law of the iterated logarithm.

[1]Research supported by Hungarian National Foundation for Scientific Research Grants T 16384 and T 19346.

0081–6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

then

$$\lim_{N\to\infty} \mu\{0 \leqq x \leqq 1 : \sum_{k\leqq N} f(n_k x) < t\sqrt{N}\} = (2\pi)^{-1/2} \int_{-\infty}^{t} e^{-u^2/2} du \tag{1.3}$$

and

$$\limsup_{N\to\infty} (2N \log\log N)^{-1/2} \sum_{k\leqq N} f(n_k x) = 1 \qquad \text{a.e.,} \tag{1.4}$$

where μ is the Lebesgue measure. As an example of Erdős and Fortet (see [8], p. 646) shows, the CLT (1.3) and the LIL (1.4) become generally false if instead of (1.2) we assume only the Hadamard gap condition

$$n_{k+1}/n_k \geqq q > 1 \qquad (k = 1, 2, \dots). \tag{1.5}$$

Indeed, let $f(x) = \cos 2\pi x + \cos 4\pi x$, $n_k = 2^k - 1$. Then, as it is not difficult to show,

$$\lim_{N\to\infty} \mu\{0 \leqq x \leqq 1 : \sum_{k\leqq N} f(n_k x) < t\sqrt{N}\} = (2\pi)^{-1/2} \int_0^1 ds \int_{-\infty}^{t/\sqrt{2}|\cos \pi s|} e^{-u^2/2} du$$

and

$$\limsup_{N\to\infty} (2N \log\log N)^{-1/2} \sum_{k\leqq N} f(n_k x) = \sqrt{2} \cos \pi x \qquad \text{a.e..}$$

On the other hand, Kac [7] showed that if f is smooth and $n_k = 2^k$ then the CLT (1.3) is valid, with the $N(0,1)$ distribution on the right-hand side replaced by $N(0, \sigma^2)$ for some $\sigma \geqq 0$. Thus we see that under (1.5) the asymptotic behaviour of $f(n_k x)$ depends not only on the growth speed of (n_k), but also on its arithmetic properties. This interesting phenomenon was investigated in detail by Gaposkin [6] who gave a characterization of sequences (n_k) satisfying the CLT (1.3) for all sufficiently smooth f. His results imply, e.g., that (1.3) holds if the ratios n_{k+1}/n_k are all integers, or if $n_{k+1}/n_k \to \beta$ where β^r is irrational for all positive integers r. For extensions and further limit theorems for $f(n_k x)$ see Gaposkin [5], Berkes [1], Berkes and Philipp [3]. It is interesting to note that if we assume only (1.5) then the upper half of the LIL still holds for $f(n_k x)$, i.e.,

$$\limsup_{N\to\infty} (2N \log\log N)^{-1/2} \left| \sum_{k\leqq N} f(n_k x) \right| \leqq C \qquad \text{a.e.} \tag{1.6}$$

for some constant C (see Takahashi [12], Philipp [9]). For further limit theorems for $f(n_k x)$ assuming only (1.5) see Berkes [1].

While under the Hadamard gap condition (1.5) the asymptotic properties of $f(n_k x)$ are fairly well known, very few results exist in the case when (n_k) grows slower than the exponential speed required by (1.5). For certain "nice" sequences (n_k) the LIL (1.6) still holds: Philipp [10] proved that this is the case if (n_k) is the sequence consisting of all integers of the form $q_1^{\alpha_1} \cdots q_r^{\alpha_r}$ ($\alpha_i \geqq 0$ integers), arranged in increasing order, where $\{q_1, \dots, q_r\}$ is a finite set of coprime integers. But, as Berkes and Philipp [4] proved, the LIL (1.6) is generally false for subexponential (n_k): for any $\varrho_k \to 0$ there exists a sequence (n_k) of positive integers satisfying

$$n_{k+1}/n_k \geqq 1 + \varrho_k \qquad (k = 1, 2, \dots)$$

such that

$$\limsup_{N\to\infty} (2N \log\log N)^{-1/2} \sum_{k \leq N} f(n_k x) = +\infty \qquad \text{a.e.}$$

with $f(x) = x - [x] - 1/2$. (Here, and in the sequel, $[x]$ denotes the largest integer not exceeding x.) The examples in [4] also show that the asymptotic properties of $\sum_{k \leqq N} f(n_k x)$ in the subexponential domain depend on the growth speed of (n_k), but no analogues of the LIL (1.6) exist in the literature for subexponential (n_k). The purpose of this paper is to prove a first result in this direction. Indeed, we shall prove the following

THEOREM. *Let $f\colon R \to R$ be a Lebesgue measurable function satisfying* (1.1) *and assume that f is of bounded variation on* $(0, 1)$. *Let (n_k) be a sequence of positive integers satisfying*

$$n_{k+1}/n_k \geqq 1 + \varrho_k \qquad (k = 1, 2, \dots) \tag{1.7}$$

where (ϱ_k) is nonincreasing with $\varrho_k \to 0$ and

$$\varrho_k \geqq k^{-\alpha} \qquad \text{for some} \quad 0 < \alpha < 1/2. \tag{1.8}$$

Then

$$\limsup_{N\to\infty} \frac{\left|\sum_{k \leqq N} f(n_k x)\right|}{\log \frac{1}{\varrho_{N^2}} \sqrt{N \log\log N}} < +\infty \qquad \textit{a.e.}. \tag{1.9}$$

As a comparison, we note that by a result of Berkes and Philipp [4], for any ϱ_k tending to 0 sufficiently slowly there exists a sequence (n_k) of integers satisfying (1.7) such that

$$\limsup_{N\to\infty} \frac{\left|\sum_{k \leqq N} f(n_k x)\right|}{\log\log \frac{1}{\varrho_N} \sqrt{N \log\log N}} \geqq c > 0 \qquad \text{a.e.} \tag{1.10}$$

with an absolute constant c, where $f(x) = x - [x] - 1/2$. Thus for subexponential (n_k) the growth speed of $\sum_{k\leqq N} f(n_k x)$ exceeds the classical LIL speed by a factor depending on the speed of convergence of n_{k+1}/n_k to 1. The upper bound (1.9) and the lower bound (1.10) are of similar character, but there is a gap between them and the precise order of magnitude of $\sum_{k\leqq N} f(n_k x)$ remains open.

2. Proof of the theorem

We use the method of our earlier paper [3]; the essential new element will be an estimate for integrals $\int_a^b (\sum f(n_k x))^2 dx$ for subexponentially growing (n_k) (cf. Lemma 3). Let s_n denote the n-th partial sum of the Fourier series of f. Since f is of bounded variation on $(0,1)$, the Fourier coefficients of f tend to zero as $O(1/k)$ (see Zygmund [15], p. 48) and thus

$$\|f - s_n\| = O(n^{-1/2}) \tag{2.1}$$

where $\|\cdot\|$ denotes the $L_2(0,1)$ norm. We now approximate the functions $f(n_k x)$ by stepfunctions $\varphi_k(x)$ as follows. Let $2^l \leqq n_k < 2^{l+1}$, put $m = [l + 120 \log k]$ and let φ_k denote the function in $[0,1)$ which takes the value $2^m \int_{i2^{-m}}^{(i+1)2^{-m}} f(n_k x)dx$ in the interval $[i2^{-m}, (i+1)2^{-m})$ $(0 \leqq i \leqq 2^m - 1)$. The assumptions made on f imply that $|f| \leqq C$ and consequently $|\varphi_k| \leqq C$ for some constant C. Hence using Lemma (3.1) in [1], p. 325 we have

$$\begin{aligned} \int_0^1 (f(n_k x) - \varphi_k(x))^4 dx &\leqq \text{const} \cdot \int_0^1 (f(n_k x) - \varphi_k(x))^2 dx \\ &\leqq \text{const} \cdot (2^m/n_k)^{-1/3} \leqq \text{const} \cdot \left(2^{120 \log k}\right)^{-1/3} \leqq \text{const} \cdot k^{-20}. \end{aligned} \tag{2.2}$$

Since $\alpha < 1/2$ we can choose β so that $\alpha/(1-\alpha) < \beta < 1$. Divide the set of positive integers into consecutive blocks $I_1, J_1, I_2, J_2, \ldots, I_k, J_k, \ldots$ (without gaps) such that

$$|I_k| = |J_k| = [k^\beta],$$

where $|A|$ denotes, for any set $A \subset R$, the number of integers in A. Set

$$T_k = \sum_{\nu \in I_k} f(n_\nu x), \qquad D_k = \sum_{\nu \in I_k} \varphi_\nu(x). \tag{2.3}$$

Then by (2.2)

$$\|D_k - T_k\|_4 \leqq C \sum_{\nu \in I_k} \nu^{-5} \leqq C \sum_{\nu \geqq k} \nu^{-5} \leqq C k^{-4}, \tag{2.4}$$

where $\|\cdot\|_p$ is the $L_p(0,1)$ norm and C denotes positive constants, possibly different at different places. Let $\mathcal{F}_k$ denote the σ-field generated by $D_1, \ldots, D_k$.

LEMMA 1. *We have**

$$|\mathbf{E}(D_k \mid \mathcal{F}_{k-1})| = O(k^{-2}) \qquad a.s. \quad as \quad k \to \infty. \tag{2.5}$$

PROOF. (2.4) shows that the expected value of $\mathbf{E}(|D_k - T_k| \mid \mathcal{F}_{k-1})$ is $\leqq Ck^{-4}$ and thus the Markov inequality and the Borel–Cantelli lemma imply

$$\mathbf{E}(|D_k - T_k| \mid \mathcal{F}_{k-1}) = O(k^{-2}) \qquad \text{a.s.} \quad \text{as} \quad k \to \infty.$$

Hence to prove (2.5) it suffices to show that

$$|\mathbf{E}(T_k \mid \mathcal{F}_{k-1})| \leqq Ck^{-2}. \tag{2.6}$$

We observe also that for any real $u < v$ and $\lambda > 0$ we have by the first two relations of (1.1)

$$\left| \int_u^v f(\lambda x)dx \right| = \left| \frac{1}{\lambda} \int_{u\lambda}^{v\lambda} f(t)dt \right| \leqq \frac{2}{\lambda} \int_0^1 |f(t)|dt. \tag{2.7}$$

Let $b = b(k)$ and $c = c(k)$ denote the largest integer of the block I_{k-1} and the smallest integer of the block I_k, respectively. Define the integer l by $2^l \leqq n_b < 2^{l+1}$ and put $w = [l + 120 \log b]$. From the definition of the φ_ν's it follows that every φ_ν, $1 \leqq \nu \leqq b$, takes a constant value on each interval of the form $I = [i2^{-w}, (i+1)2^{-w})$, $0 \leqq i \leqq 2^w - 1$ and thus each set of $\mathcal{F}_{k-1}$ is a union of intervals I of the above type. Hence to prove (2.6) it suffices to show that

$$\left| 2^w \int_{i2^{-w}}^{(i+1)2^{-w}} T_k dx \right| \leqq Ck^{-2} \qquad (0 \leqq i \leqq 2^w - 1). \tag{2.8}$$

Using (1.7), (1.8), $c \leqq 2k^{\beta+1}$ and $c - b = [(k-1)^\beta] + 1$ we get

$$\begin{aligned} \frac{n_c}{n_b} &\geqq \prod_{j=b}^{c-1} (1 + j^{-\alpha}) \geqq (1 + c^{-\alpha})^{c-b} \\ &\geqq \left(1 + \frac{1}{(2k^{\beta+1})^\alpha}\right)^{\frac{1}{2}k^\beta} \geqq \exp\left(\frac{1}{4}(2k^{\beta+1})^{-\alpha} k^\beta\right) \\ &= \exp(Ck^{\beta(1-\alpha)-\alpha}) \geqq \exp(k^\delta) \qquad (k \geqq k_0) \end{aligned} \tag{2.9}$$

* $\mathbf{E}$ and $\mathbf{P}$ denote expectation, resp. probability in the probability space $((0,1), \mathcal{B}, \mu)$ where $\mathcal{B}$ is the Borel σ-field in $(0,1)$ and μ is the Lebesgue measure.

for some $\delta > 0$ since $\beta > \alpha/(1-\alpha)$. Thus using (2.7), (2.9), $\beta < 1$ and $b \leqq 2k^{\beta+1}$ we get

$$\left|2^w \int_{i2^{-w}}^{(i+1)2^{-w}} T_k dx\right| = \left|2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \sum_{\nu \in I_k} f(n_\nu x) dx\right|$$
$$\leqq 2^w C \sum_{\nu \in I_k} \frac{2}{n_\nu} \leqq C \frac{2^w}{n_c} |I_k| \leqq C \frac{2^{l+120 \log b}}{n_c} k$$
$$\leqq C \frac{n_b}{n_c} b^{120} k \leqq C \exp(-k^\delta)(2k^{\beta+1})^{120} k \leqq C k^{-2}$$

proving (2.8).

LEMMA 2. *We have*

$$\mathbf{E}(D_k^4) \leqq C k^{2\beta+1-\tau} \log^4 k \tag{2.10}$$

for some constant $\tau > 0$.

PROOF. To simplify the formulas, we prove the lemma in the case when the Fourier series $f = \sum_{j=1}^{\infty} a_j \cos 2\pi jx$ of f is a pure cosine series; the general case can be treated similarly. (Here, and in the sequel, the convergence of trigonometric series is meant in L_2 norm.) In view of (2.4) and Minkowski's inequality, it suffices to prove (2.10) with D_k replaced by T_k. Let

$$f_1 = \sum_{j \leqq k^{10}} a_j \cos 2\pi jx, \qquad f_2 = \sum_{j > k^{10}} a_j \cos 2\pi jx$$
$$T_k^{(1)} = \sum_{\nu \in I_k} f_1(n_\nu x), \qquad T_k^{(2)} = \sum_{\nu \in I_k} f_2(n_\nu x)$$
$$T_{k,j} = \sum_{\nu \in I_k} \cos 2\pi j n_\nu x.$$

Using (1.7), (1.8) and Lemma (5.2) of [2], we get

$$\mathbf{E}(T_{k,j}^4) \leqq C\left(k^{2\beta} + \frac{k^{2\beta+1}}{k^{\frac{1}{2}-\alpha}}\right) \leqq C k^{2\beta+1-\tau} \tag{2.11}$$

for some constant $\tau > 0$. Also,

$$T_k^{(1)} = \sum_{j \leqq k^{10}} a_j T_{k,j}$$

and here $|a_j| = O(j^{-1})$ since f is of bounded variation (see [15], p. 48). Thus using the Minkowski inequality and (2.11) we get

$$\|T_k^{(1)}\|_4 \leqq \sum_{j \leqq k^{10}} |a_j| \|T_{k,j}\|_4 \leqq Ck^{(2\beta+1-\tau)/4} \sum_{j \leqq k^{10}} |a_j| \leqq Ck^{(2\beta+1-\tau)/4} \log k. \tag{2.12}$$

Since f is of bounded variation, the partial sums of the Fourier series of f are uniformly bounded (see [15], p. 90) and consequently $\|f_2\|_4 \leqq C\|f_2\|_2^{1/2}$. Hence another application of Minkowski's inequality gives

$$\|T_k^{(2)}\|_4 \leqq \|f_2\|_4 |I_k| \leqq C\|f_2\|_2^{1/2} k \leqq C\Big(\sum_{j > k^{10}} j^{-2}\Big)^{1/4} k \leqq Ck^{-1}. \tag{2.13}$$

Now (2.12) and (2.13) imply (2.10) with D_k replaced by T_k and thus Lemma 2 is proved.

LEMMA 3. *Let $1 \leqq m_1 < m_2 < \cdots < m_N$ be a sequence of positive numbers (not necessarily integers) such that for some $1 < q < 3/2$ we have*

$$m_{k+1}/m_k \geqq q \qquad (k = 1, \ldots, N-1) \tag{2.14}$$

$$m_{k+1} - m_k \geqq 2 \qquad (k = 1, \ldots, N-1). \tag{2.15}$$

Let $f = \sum_{k=1}^{\infty} (a_k \cos 2\pi kx + b_k \sin 2\pi kx)$ be a function with $|a_k| \leqq 1/k$, $|b_k| \leqq 1/k$ $(k = 1, 2, \ldots)$. Then for any real a we have

$$\int_a^{a+1} \Big(\sum_{k \leqq N} f(m_k x)\Big)^2 dx \leqq CN\Big(\log \frac{1}{q-1}\Big)^2, \tag{2.16}$$

where C is an absolute constant.

PROOF. To simplify the writing we assume that $f = \sum_{k=1}^{\infty} c_k \cos 2\pi kx$ is a pure cosine series. (The general case can be treated similarly.) We proceed in steps.

1. We first note that

$$\int_a^{a+1} \Big(\sum_{k \leqq N} \cos m_k x\Big)^2 dx \leqq 36N \tag{2.17}$$

for any real a. Indeed, the square root of the left-hand side of (2.17) equals

$$\Big\|\sum_{k\leqq N}\cos m_k(x+a)\Big\|_2\leqq\Big\|\sum_{k\leqq N}\cos m_kx\cos m_ka\Big\|_2+$$
$$+\Big\|\sum_{k\leqq N}\sin m_kx\sin m_ka\Big\|_2$$

and thus to prove (2.17) it suffices to show that

$$\Big\|\sum_{k\leqq N}\lambda_k\cos m_kx\Big\|_2\leqq 3\sqrt{N},\qquad \Big\|\sum_{k\leqq N}\lambda_k\sin m_kx\Big\|_2\leqq 3\sqrt{N} \tag{2.18}$$

for arbitrary $|\lambda_k|\leqq 1$. This, however, follows by observing that $\sin x/x\geqq 1/2$ for $0\leqq x\leqq 1$ and thus

$$\int_0^1\Big(\sum_{k\leqq N}\lambda_k\cos m_kx\Big)^2dx\leqq 4\int_0^1\Big(\frac{\sin x}{x}\Big)^2\Big(\sum_{k\leqq N}\lambda_k\cos m_kx\Big)^2dx$$
$$\leqq 4\int_{-\infty}^{+\infty}\Big(\frac{\sin x}{x}\Big)^2\Big(\sum_{k\leqq N}\lambda_k\cos m_kx\Big)^2dx<9N,$$

where the last inequality follows from (2.15) by expanding $\Big(\sum_{k\leqq N}\lambda_k\cos m_kx\Big)^2$ and using the fact that

$$\int_{-\infty}^{+\infty}(\sin x/x)^2dx=\pi\qquad\text{and}\qquad\int_{-\infty}^{+\infty}(\sin x/x)^2\cos ux\,dx=0\quad\text{for}\quad|u|\geqq 2.$$

The second relation of (2.18) follows similarly.

2. We prove now that under the conditions of Lemma 3 we have for any real a

$$\int_a^{a+1}\Big(\sum_{k\leqq N}f(m_kx)\Big)^2dx\leqq\frac{C\|f\|_2}{q-1}N \tag{2.19}$$

provided that $\|f\|_2\leqq 1$; here C is an absolute constant. Indeed, by Lemma (5.1) in [1], p. 338 we have

$$\Big|\int_a^{a+1}f(m_ix)f(m_jx)dx\Big|\leqq C_1\|f\|_2\Big(\frac{m_i}{m_j}\Big)^{1/2}\qquad(i<j),$$

where C_1 is an absolute constant. The last relation holds also for $i=j$ as it is seen by applying (2.7) with $\lambda=m_i$ and with $f^2-\|f\|^2$ instead of f. Now (2.19) follows upon noticing that by (2.14) and $1<q<3/2$ we have

$$\sum_{1\leqq i\leqq j\leqq N}\left(\frac{m_i}{m_j}\right)^{1/2}\leqq N\sum_{k=0}^{\infty}q^{-k/2}=\frac{N}{1-q^{-1/2}}\leqq\frac{4N}{q-1}.$$

3. Write

$$f=\sum_{l=1}^{\infty}c_l\cos 2\pi lx=\sum_{l\leqq T}+\sum_{l>T}=:f_1+f_2,$$

where $T\geqq 2$ will be chosen later. Letting $\|\cdot\|_{2,a}$ denote the $L_2(a,a+1)$ norm, we have by the statement proved in Step 1

$$\begin{aligned}\Big\|\sum_{k\leqq N}f_1(m_kx)\Big\|_{2,a}&\leqq\sum_{l\leqq T}|c_l|\Big\|\sum_{k\leqq N}\cos 2\pi l m_k x\Big\|_{2,a}\\&\leqq\sum_{l\leqq T}\frac{1}{l}6\sqrt{N}\leqq 6\sqrt{N}(\log T+1).\end{aligned}\tag{2.20}$$

On the other hand, $\|f_2\|_2\leqq\left(\sum_{l>T}l^{-2}\right)^{1/2}\leqq CT^{-1/2}$ and thus by the statement proved in Step 2 we have

$$\begin{aligned}\Big\|\sum_{k\leqq N}f_2(m_kx)\Big\|_{2,a}&\leqq\left(\frac{C\|f_2\|_2}{q-1}N\right)^{1/2}\\&\leqq\left(\frac{CT^{-1/2}}{q-1}N\right)^{1/2}.\end{aligned}\tag{2.21}$$

Now (2.16) follows from (2.20) and (2.21) upon choosing $T=1/(q-1)^2$ and noting that $\log 1/(q-1)\geqq 1/2$. This completes the proof of Lemma 3.

LEMMA 4. *We have*

$$\mathbf{E}(D_k^2\,|\,\mathcal{F}_{k-1})=O\left(|I_k|\left(\log\frac{1}{\varrho_{k^2}}\right)^2\right)\qquad a.s.\ as\quad k\to\infty.\tag{2.22}$$

PROOF. Let b,c,w denote the same as in the proof of Lemma 1 and let $d=d(k)$ denote the largest integer of the block I_k. Using (2.4) and $|T_k|\leqq Ck$, $|D_k|\leqq Ck$ we get

$$\|D_k^2-T_k^2\|_1\leqq Ck\|D_k-T_k\|_1\leqq Ck^{-3}$$

and thus by the Markov inequality and the Borel–Cantelli lemma it follows immediately that

$$\mathbf{E}(|D_k^2 - T_k^2| \mid \mathcal{F}_{k-1}) = O(k^{-1}) \qquad \text{a.s. as} \quad k \to \infty.$$

Hence similarly as in the proof of Lemma 1, (2.22) will follow if we show that for $k \geqq k_0$

$$(2.23)\quad 2^w \int_{i2^{-w}}^{(i+1)2^{-w}} \Big(\sum_{\nu \in I_k} f(n_\nu x)\Big)^2 dx \leqq C|I_k|\Big(\log \frac{1}{\varrho_{k^2}}\Big)^2 \qquad (0 \leqq i \leqq 2^w - 1).$$

Here

$$(2.24)\qquad 2^w \leqq 2^{l+120\log b} \leqq n_b b^{120}$$

and the substitution $t = 2^w x$ shows that the left-hand side of (2.23) equals

$$\int_i^{i+1} \Big(\sum_{\nu \in I_k} f(m_\nu t)\Big)^2 dt,$$

where $m_\nu = 2^{-w} n_\nu$. Now $d \leqq 2k^{\beta+1} \leqq k^2$ for $k \geqq k_0$ by $\beta < 1$ and thus using (1.7) and the monotonicity of ϱ_k we get for $k \geqq k_0$

$$m_{\nu+1}/m_\nu \geqq 1 + \varrho \qquad (\nu \in I_k),$$

where $\varrho = \varrho_{k^2}$. Clearly $\varrho < 1/2$ for sufficiently large k; on the other hand, for $\nu \in I_k$ we have by (2.9), (2.24), (1.7), (1.8) and $b \leqq c \leqq d \leqq k^2$

$$\begin{aligned} m_{\nu+1} - m_\nu &= (n_{\nu+1} - n_\nu)2^{-w} \geqq n_\nu \Big(\frac{n_{\nu+1}}{n_\nu} - 1\Big)\frac{1}{n_b b^{120}} \\ &\geqq \frac{n_\nu}{n_b b^{120}} \varrho_\nu \geqq \exp(k^\delta) k^{-240} k^{-2\alpha} \geqq 2 \end{aligned}$$

for $k \geqq k_0$ and thus (2.23) follows from Lemma 3.

Let now $\overline{D}_k = D_k - \mathbf{E}(D_k \mid \mathcal{F}_{k-1})$. Clearly $(\overline{D}_k, k \geqq 1)$ is a martingale difference sequence and Lemmas 1, 2 and 4 easily yield

$$(2.25)\qquad \mathbf{E}(\overline{D}_k^2 \mid \mathcal{F}_{k-1}) = O\left(|I_k|\Big(\log \frac{1}{\varrho_{k^3}}\Big)^2\right) \qquad \text{a.s. as} \quad k \to \infty$$

$$(2.26)\qquad \mathbf{E}(\overline{D}_k^4) \leqq C k^{2\beta+1-\tau} \log^4 k.$$

Relation (2.25) and the monotonicity of ϱ_k imply

$$\begin{aligned}\sum_{k\leqq n}\mathbf{E}(\overline{D}_k^2\mid\mathcal{F}_{k-1})&\leqq O(1)\sum_{k\leqq n}|I_k|\Big(\log\frac{1}{\varrho_{k^2}}\Big)^2\\&=O\left(n^{\beta+1}\Big(\log\frac{1}{\varrho_{n^2}}\Big)^2\right)\quad\text{a.s..}\end{aligned}\tag{2.27}$$

By the martingale version of the Skorohod representation theorem (see [11], Theorem 4.3) the sequence $(\overline{D}_k, k\geqq 1)$ can be redefined, without changing its distribution, on a suitable probability space together with a Wiener process W, nonnegative random variables $\tau_1,\tau_2,\dots$ and an increasing sequence $(\mathcal{H}_k, k\geqq 1)$ of σ-fields such that

$$\sum_{k\leqq n}\overline{D}_k=W\Big(\sum_{k\leqq n}\tau_k\Big)\tag{2.28}$$

further τ_k is $\mathcal{H}_k$ measurable and

$$\mathbf{E}(\tau_k\mid\mathcal{H}_{k-1})=\mathbf{E}(\overline{D}_k^2\mid\mathcal{F}_{k-1}),\qquad \mathbf{E}(\tau_k^2)\leqq C\mathbf{E}(\overline{D}_k^4).\tag{2.29}$$

By (2.29) and (2.26) we have

$$\sum_{k=1}^{\infty}\frac{\mathbf{E}(\tau_k-\mathbf{E}(\tau_k\mid\mathcal{H}_{k-1}))^2}{k^{2\beta+2}}\leqq\sum_{k=1}^{\infty}\frac{\mathbf{E}\tau_k^2}{k^{2\beta+2}}\leqq C\sum_{k=1}^{\infty}\frac{\mathbf{E}(\overline{D}_k^4)}{k^{2\beta+2}}<+\infty$$

and thus the martingale convergence theorem implies that the sum

$$\sum_{k=1}^{\infty}k^{-(\beta+1)}(\tau_k-\mathbf{E}(\tau_k\mid\mathcal{H}_{k-1}))$$

is a.s. convergent. Hence by the Kronecker lemma

$$\sum_{k\leqq n}(\tau_k-\mathbf{E}(\tau_k\mid\mathcal{H}_{k-1}))=o(n^{\beta+1})\quad\text{a.s..}$$

Moreover, (2.27) and the first relation of (2.29) yield

$$\sum_{k\leqq n}\mathbf{E}(\tau_k\mid\mathcal{H}_{k-1})=O\left(n^{\beta+1}\Big(\log\frac{1}{\varrho_{n^2}}\Big)^2\right)\quad\text{a.s.}$$

and consequently

$$\sum_{k\leqq n}\tau_k=O\left(n^{\beta+1}\Big(\log\frac{1}{\varrho_{n^2}}\Big)^2\right)\quad\text{a.s..}$$

Thus using (2.28), the law of the iterated logarithm for W and $\log\log xy \leqq \log\log x + \log\log y$ $(x \geqq 3,\ y \geqq 3)$ we get

$$(2.30)\qquad \begin{aligned}\sum_{k\leqq n} \overline{D}_k &= O\left(n^{(\beta+1)/2}\log\frac{1}{\varrho_{n^2}}\Big(\log\log n + \log\log\log\frac{1}{\varrho_{n^2}}\Big)^{1/2}\right)\\ &= O\left(n^{(\beta+1)/2}(\log\log n)^{1/2}\log\frac{1}{\varrho_{n^2}}\right)\qquad \text{a.s.}\end{aligned}$$

since $1/\varrho_{n^2} \leqq n$ by (1.8). Now (2.4), the Markov inequality and the Borel–Cantelli lemma imply $|D_k - T_k| = O(k^{-2})$ a.s. as $k \to \infty$ and thus by Lemma 1 we get

$$|\overline{D}_k - T_k| = O(k^{-2}) \qquad \text{a.s. as } k\to\infty.$$

Hence (2.30) implies

$$(2.31)\qquad \sum_{k\leqq n} T_k = O\left(n^{(\beta+1)/2}(\log\log n)^{1/2}\log\frac{1}{\varrho_{n^2}}\right)\qquad \text{a.s..}$$

Introducing

$$T_k' = \sum_{\nu\in J_k} f(n_\nu x)$$

we have, similarly to (2.31),

$$(2.32)\qquad \sum_{k\leqq n} T_k' = O\left(n^{(\beta+1)/2}(\log\log n)^{1/2}\log\frac{1}{\varrho_{n^2}}\right)\qquad \text{a.s.}$$

and thus setting $S_N = \sum_{\nu\leqq N} f(n_\nu x)$ and $N_k = \sum_{\nu\leqq k} 2[\nu^\beta] \sim ck^{\beta+1}$ it follows by adding (2.31) and (2.32) and using the monotonicity of ϱ_k that

$$(2.33)\qquad \begin{aligned}S_{N_k} &= O\left(k^{(\beta+1)/2}(\log\log k)^{1/2}\log\frac{1}{\varrho_{k^2}}\right) = O\left((N_k\log\log N_k)^{1/2}\log\frac{1}{\varrho_{k^2}}\right)\\ &= O\left((N_k\log\log N_k)^{1/2}\log\frac{1}{\varrho_{N_k^2}}\right)\qquad \text{a.s..}\end{aligned}$$

Now if $N_k \leqq N < N_{k+1}$ then by $|f| \leqq C$ and $\beta < 1$ we get

$$\begin{aligned}|S_N - S_{N_k}| &\leqq C(N_{k+1} - N_k) \leqq Ck^\beta \leqq CN_k^{\beta/(\beta+1)}\\ &\leqq CN_k^{1/2}\end{aligned}$$

and thus (2.33) implies

$$S_N = O\left((N\log\log N)^{1/2}\log\frac{1}{\varrho_{N^2}}\right)\qquad \text{a.s.}$$

completing the proof of our theorem.

REFERENCES

[1] BERKES, I., On the asymptotic behaviour of $\sum f(n_k x)$. I–II. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **34** (1976), 319–345, 347–365. *MR* **53** #4175, 4176

[2] BERKES, I., On the central limit theorem for lacunary trigonometric series, *Anal. Math.* **4** (1978), 159–180. *MR* **80c**:60032

[3] BERKES, I. AND PHILIPP, W., An a.s. invariance principle for lacunary series $f(n_k x)$, *Acta Math. Acad Sci. Hungar.* **34** (1979), 141–155. *MR* **80i**:60042

[4] BERKES, I. AND PHILIPP, W., The size of trigonometric and Walsh series and uniform distribution mod 1, *J. London Math. Soc.* (2) **50** (1994), 454–464. *MR* **96e**:11099

[5] GAPOSKIN, V., Lacunary series and independent functions, *Uspehi Mat. Nauk* **21** (1966), vyp. 6, 3–82 (in Russian). *MR* **34** #6374

[6] GAPOSKIN, V., The central limit theorem for certain weakly dependent sequences, *Teor. Verojatnost. Primenen.* **15** (1970), 666–684 (in Russian). *MR* **43** #8106

[7] KAC, M., On the distribution of values of sums of the type $\sum f(2^k t)$, *Ann. of Math.* (2) **47** (1946), 33–49. *MR* **7**, 436f

[8] KAC, M., Probability methods in some problems of analysis and number theory, *Bull. Amer. Math. Soc.* **55** (1949), 641–665. *MR* **11**, 161b

[9] PHILIPP, W., Limit theorems for lacunary series and uniform distribution mod 1, *Acta Arith.* **26** (1974/75), 241–251. *MR* **52** #325

[10] PHILIPP, W., Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory, *Trans. Amer. Math. Soc.* **345** (1994), 705–727. *MR* **95a**:11067

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

[12] TAKAHASHI, S., An asymptotic property of a gap sequence, *Proc. Japan. Acad.* **38** (1962), 101–104. *MR* **25** #4279

[13] TAKAHASHI, S., A gap sequence with gaps bigger than Hadamard's, *Tôhoku Math. J.* **13** (1961), 105–111. *MR* **23** #A3946

[14] TAKAHASHI, S., The law of the iterated logarithm for a gap sequence with infinite gaps, *Tôhoku Math. J.* (2) **15** (1963), 281–288. *MR* **28** #384

[15] ZYGMUND, A., *Trigonometric series*, 2nd ed., Vol. I, University Press, Cambridge, 1959. *MR* **21** #6498

(Received November 4, 1997)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY
berkes@math-inst.hu

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
725 S. WRIGHT STREET
CHAMPAIGN, IL 61820
U.S.A.
philipp@stat.uiuc.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 15–23

DECOMPOSITIONS IN DISCRETE SEMIGROUPS

W. E. CLARK, W. C. HOLLAND and G. J. SZÉKELY

Dedicated to the memory of Alfréd Rényi

Abstract

In this paper we prove that under some finiteness conditions in a (not necessarily commutative and not necessarily cancellative) semigroup every non-unit is a product of weakly irreducible elements. In commutative, finitely generated semigroups every infinitely divisible element is idempotent. Without commutativity this is not true. An interesting open problem is to find necessary and sufficient conditions for this implication.

1. Introduction

The most well-known semigroups in which every non-unit is a product of irreducible elements are the multiplicative semigroup of positive integers and the multiplicative semigroup of non-zero polynomials. Their common generalization is the class of Gaussian semigroups. These are commutative, cancellative semigroups with identity element 1 in which every element (other than units, that is, divisors of 1) is a product of irreducible elements and this decomposition is essentially unique (for necessary and sufficient conditions see e.g. Kurosh [11]). For some more recent factorization results see Anderson et al. [1], Halter-Koch [7] and several references there. In these papers it was supposed that the semigroup is commutative and cancellative. Here we consider similar types of decompositions for not necessarily commutative and not necessarily cancellative semigroups. The classical definition of irreducibility is as follows. In a semigroup S with an identity, an element s of S is called irreducible if it is not a unit and $s=ab$ (a,b from S) implies a or b is a unit. There is no hope that in every semigroup every non-unit turns out to be a product of irreducibles.

EXAMPLE 1.1. Let S be the semigroup of all subsets of an infinite set with the union operation. The singletons are the only irreducibles in S and no infinite set is a "product" (= finite union) of singletons.

In order to avoid this kind of problem we either need some kind of "finiteness" condition or we have to introduce a topology. In topological semigroups

1991 *Mathematics Subject Classification.* Primary 20M99.
Key words and phrases. Weak irreducibility, divisibility, idempotents, finitely generated semigroups.

The third author was one of Rényi's last students.

we can, of course, consider infinite products, too, and in this case "finiteness" can be replaced by some kind of "compactness". An important example of this type is due to Khintchin [9]. In the convolution semigroup of probability distributions (on the real line) endowed with the topology of weak convergence, Khintchin proved that every element is decomposable into a product analogous to the above mentioned decompositions. In this special (commutative) semigroup Khintchin introduced a new type of elements that have no irreducible divisors at all (examples include the normal and Poisson distributions). If we call these elements *anti-irreducible*, then Khintchin's theorem can be formulated as follows: every probability distribution on the real line is a convolution product of finitely or at most countably many irreducible factors and an anti-irreducible one. Khintchin proved that the divisors of every element in this semigroup form a compact set provided that they are suitably "shifted". This is a kind of compactness that can replace "finiteness" conditions. Khintchin proved that all anti-irreducible elements are infinitely divisible but the converse is not true (examples include exponential distributions: they are infinitely divisible but not anti-irredicible). This idea of Khintchin can be generalized to almost arbitrary topological semigroups. We refer to Ruzsa and Székely [19], [20], [21] where an extensive literature is included.

In this paper we restrict ourselves to discrete semigroups and want to discuss purely algebraic and not topological types of problems. To see that finiteness itself is not enough to settle all problems in this context let us consider the following example.

EXAMPLE 1.2. Let S be the commutative semigroup $S = \{1, 0, e, c\}$ where 1 is the identity element, 0 is the zero element, e is idempotent, $c^2 = 0$, $ec = c$. In this semigroup there are no irreducibles at all, so no element is a product of irreducibles.

To overcome the problem of this example, in Section 2 we are going to introduce a weaker notion of irreducibility, and we shall see that two elements of our semigroup above (e and c) will turn out to be weakly irreducible, and 0 is a square of a weakly irreducible element.

From the point of view of decompositions, irreducibility is one extreme, the other one is infinite divisibility. We shall prove that if S is commutative and finitely generated then the only infinitely divisible elements are the idempotents, but if we drop the condition of commutativity then this is not true anymore.

2. Main results

Let us fix some notation and notions. We assume that all semigroups have an identity element which we denote by 1.

DEFINITION 2.1. For two elements a and b of a semigroup S, we say that a *divides* b, and write $a|b$, if for some $x, y \in S$, $b = xay$. Further, a and b are called *associates* (denoted $a \sim b$) if they divide each other; that is, $a = xby$ and $b = uav$ for some, x, y, u, v from S. (In other words a and b are associates if they generate the same two-sided ideal.)

DEFINITION 2.2. An s from S is called *weakly irreducible* if s is not a unit and $s = ab$ implies that either a or b is an associate of s.

In the following, irreducibility always means weak irreducibility except if we emphasize that we think of the usual (strong) irreducibility. It is clear that for commutative cancellative semigroups, including Gaussian semigroups, the notion of weak and strong irreducibility coincide.

If we apply the above introduced notion of irreducibility then in our Example 1.2, e and c become irreducible and $0 = c^2$; thus every non-unit is irreducible or a product of irreducibles. (Interestingly, e is infinitely divisible and also (weakly!) irreducible.) This example is in fact typical as our first result shows.

THEOREM 2.1. *Let S be a finitely generated (not necessarily commutative) semigroup. Then every element in S is either a unit or a product of irreducible elements.*

Although this theorem does not cover the case of the multiplicative semigroup of integers, the following trivial corollary does.

COROLLARY 2.1. *If S is an arbitrary semigroup and the set of divisors of an s in S is contained in a finitely generated subsemigroup then s is a unit or a product of irreducible elements.*

THEOREM 2.2. (i) *If S is a commutative finitely generated semigroup then every infinitely divisible element is idempotent.*

(ii) *If S is a finite (commutative or non-commutative) semigroup then every infinitely divisible element is idempotent.*

(iii) *There exist non-commutative finitely generated semigroups where not all infinitely divisible elements are idempotent.*

REMARK 2.1. Parts of this theorem, especially part (ii), are folklore but for completeness we include two short proofs of part (ii), one of them a nice application of Ramsey's theorem from graph theory.

Finitely generated semigroups are not the only ones having some kind of "finiteness" condition. Another well-known type is the residually finite semigroups, that is, those semigroups which are subdirect products of finite semigroups (a very nice paper on these semigroups is Schein [22]). It is clear that the decomposition theorem cannot hold for all residually finite semigroups. Take, e.g., a countably infinite product of finite semigroups where each factor is the union semigroup of all subsets of a finite (non-empty) set. The analogue of Theorem 2.2 (ii), however, holds for residually finite semigroups.

THEOREM 2.3. *If S is a residually finite semigroup then all infinitely divisible elements of S are idempotent.*

It is not hard to prove (by induction with respect to the number of generating elements) that every finitely generated semigroup in which each element is idempotent, is finite. This result is in fact a consequence of a Burnside-type theorem for semigroups (see Restivo and Reutenauer [18]). Since by (iii) above, in non-commutative semigroups infinitely divisible elements are not necessarily idempotents it is interesting to ask if there are non-finite but finitely generated divisible semigroups (where all elements are infinitely divisible). This is in fact true. Moreover, very surprisingly, there exist infinite but finitely generated divisible *groups*. See the paper by Guba [5] and also the book by Ol'shanskii [14]. Now we arrive at an interesting conjecture.

CONJECTURE 2.1 (Dual Burnside Problem). *A finitely generated divisible semigroup is finite if and only if it is permutable (there exists an n such that for the product of any n elements there exists a nontrivial (nonidentity) permutation of these elements whose product is the same).*

DEFINITION 2.3. In a semigroup S an element s is called *anti-irreducible* if (i) s is not irreducible, and (ii) $s = bac$ where a is an irreducible element from S implies that either b or c is an associate of s. (In other words s is anti-irreducible if it is not irreducible and is not divisible "effectively" by any irreducible element where "effective" means that the "other factors" are not associates of s.)

DEFINITION 2.4. For elements $a, b \in S$, we write $a \sqsubseteq b$ if there are factorizations

$$b = \prod_{i=1}^{n} b_i$$

and

$$a = \prod_{j=1}^{k} b_{i_j},$$

where

$$\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\} \quad (i_r \neq i_s \text{ if } r \neq s).$$

THEOREM 2.4. *Let S be a finitely generated (not necessarily commutative) semigroup. If $s \in S$ is anti-irreducible, then $s^2 \sqsubseteq s$.*

COROLLARY 2.2. *If S is a commutative finitely generated semigroup, then every anti-irreducible element of S is an associate of an idempotent.*

PROBLEM 2.1. Is there a non-commutative finitely generated semigroup where an anti-irreducible element is not an associate of an infinitely divisible one? Recall that in the non-commutative finitely generated case we have not characterized the infinitely divisible elements.

In finitely generated semigroups Theorem 2.1 shows that for the decomposition we do not need anti-irreducible elements. Thus their characterization is obviously less important. There exist, however, important discrete semigroups (not finitely generated) where anti-irreducible elements play a crucial role in decompositions. Take, e.g., the multiplicative semigroup S_k of $k \times k$ stochastic matrices (these are $k \times k$ matrices with nonnegative entries such that all row-sums are equal to 1). Stochastic matrices play a fundamental role in the theory of Markov chains. The infinitely divisible elements in this semigroup were characterized in Székely–Móri–Göndőcs–Michaletzky [24]. Here we only mention the following decomposition theorem. The proof will be published elsewhere.

THEOREM 2.5. (i) *In S_k the only units are the permutation matrices.*

(ii) *Every element of S_k is either a unit or can be decomposed into a finite product of irreducibles and an anti-irreducible element.*

(iii) *For each $M \in S_k$, the number of irreducible factors in the shortest such decomposition of M is bounded above by a constant times $k^2 2^k \log k$.*

(iv) *A necessary condition for a matrix to be strongly irreducible is that every row in it contains at least one 0. (This condition is not sufficient. The characterization of irreducible and anti-irreducible stochastic matrices is an open problem.)*

3. Proofs

PROOF OF THEOREM 2.1. The theorem is a consequence of the following lemma.

LEMMA 3.1. *If $x_1, \dots, x_t$ is any minimal set of generators for a semigroup S, then each x_i is irreducible or a unit.*

PROOF. Suppose $x_1 = s_1 s_2$. Then by the minimality, without loss of generality we may suppose that s_1 does not belong to the subsemigroup generated by $x_2, \dots, x_t$. Hence, $s_1 = u x_1 v$ for some u, v. Thus, since $x_1 = 1 \cdot s_1 s_2$, s_1 is an associate of x_1. This shows that x_1 is irreducible if it is not a unit.

Now to finish the proof of Theorem 2.1, note that if a is irreducible and u is a unit, then ua is also irreducible. Hence, any product of units and irreducibles is either a product of units alone, and thus, a unit, or is a product of irreducibles.

PROOF OF THEOREM 2.2(i). Mal'cev [12] proved that every finitely generated commutative semigroup is residually finite. Therefore, this theorem follows from Theorem 2.3. We note that a more direct proof is also possible using Dickson's Lemma (see [17, p. 62]).

FIRST PROOF OF THEOREM 2.2(ii). Let $a \in S$ be infinitely divisible. For each $s \in S$, let $X_s = \{n \in N \mid s^n = a\}$. Noting that the finite collection of sets $\{X_s \mid s \in S\}$ covers the set $\mathbb{N}$ of positive integers, we color the edges of the complete graph on $\mathbb{N}$ by coloring the edge between integers p and q with any color s such that $|q-p| \in X_s$. By Ramsey's Theorem (see, e.g., Graham, Rothschild and Spencer [4]), there must be a monochrome triangle, that is, for some positive integers $p < q < r$, and some $s \in S$, we have $q-p, r-q, r-p \in X_s$. This implies that $a^2 = s^{r-q}s^{q-p} = s^{r-p} = a$, and so a is an idempotent.

SECOND PROOF OF THEOREM 2.2(ii). Because S is finite, for each $s \in S$ there exists a positive integer $n(s)$ such that $s^{n(s)}$ is an idempotent. If N is the least common multiple of $\{n(s) \mid s \in S\}$, then for all $x \in S$, x^N is an idempotent. If $a \in S$ is infinitely divisible, then since $a = x^N$ for some x, a is idempotent.

PROOF OF THEOREM 2.2(iii). Since every countable semigroup is embeddable into a semigroup generated by two elements (see e.g. Evans [3] or Hall [6]), then the additive semigroup of positive rational numbers is contained in a two-generator semigroup. Clearly, each rational number in this semigroup is infinitely divisible but not idempotent.

REMARK 3.1. In the preceeding proof, the two generating elements are irreducible, according to Lemma 3.1. Thus, every countable semigroup can be embedded in a semigroup in which each element is a finite product of irreducible elements. In fact, one can easily see much more: every semigroup can be embedded in a semigroup in which each element is irreducible. This follows from Clifford and Preston [2, §8.5]: Any semigroup is embeddable into a simple semigroup with identity. In simple semigroups there are no non-trivial two-sided ideals, thus every non-unit is irreducible. It is also true that every semigroup is embeddable into a divisible one (see, e.g., Shutov [23]; for commutative semigroups see Tamura [25]). For groups this is a result of Neumann [13] (see also [10]).

PROOF OF THEOREM 2.3. If $s \in S$ is infinitely divisible and $s \neq s^2$ then there exists a homomorphism h from S into a finite semigroup (one of its subdirect factors) such that h separates s and s^2, that is, $h(s)$ is different from $h(s^2)$. On the other hand $h(s)$ is clearly infinitely divisible in this finite semigroup and thus by Theorem 2.2 (ii) $h(s) = (h(s))^2 = h(s^2)$. This contradiction proves our claim.

PROOF OF THEOREM 2.4. If s is a unit, then $s = ssss^{-1}s^{-1}$ gives an appropriate factorization. If s is not a unit, then by Theorem 2.1, $s = x_1 \cdots x_m = 1 \cdot x_1 \cdots x_m$ where each x_i is irreducible. Since x_i is irreducible and s is anti-irreducible, it must be that $s \sim x_1 \cdots x_{i-1}$ or $s \sim x_{i+1} \cdots x_m$;

that is, $x_1 \cdots x_{i-1} = p_i s q_i$ or $x_{i+1} \cdots x_m = p_i s q_i$ for some $p_i, q_i \in S$. In the first case, we have a factorization

$$\begin{aligned} s &= x_1 \cdots x_m \\ &= p_i s q_i x_i x_{i+1} \cdots x_m \\ &= p_i (x_1 \cdots x_{i-1} x_i x_{i+1} \cdots x_m) q_i x_i x_{i+1} \cdots x_m \\ &= p_i ((p_i s q_i) x_i x_{i+1} \cdots x_m) q_i x_i x_{i+1} \cdots x_m \end{aligned}$$

in which s appears and x_i appears twice. The second case is similar. If $j \neq i$, we may replace the last s in the displayed equation with a similar expression involving x_j. The result is a factorization in which s appears and each of x_i, x_j appears twice. Continuing in this way, we obtain a factorization of s in which each of the irreducible factors x_i appears twice. Therefore, $s^2 \sqsubseteq s$.

PROOF OF COROLLARY 2.2. If $s^2 \sqsubseteq s$ then $s = s^2 y$ for some $y \in S$. Then s is an associate of the idempotent sy.

4. Atoms

In decompositions, instead of irreducible factors we might want to use any prescribed kind of factors. We shall call these *atoms*.

DEFINITION 4.1. Take an arbitrary subset A of a semigroup S and call the elements of A *atoms*. An s in S is an *anti-atom* if (i) s is not an atom, and (ii) s is not effectively divisible by any atom. (For the notion of "effectively divisible" see Definition 2.3.)

THEOREM 4.1. *In a finite semigroup S every element s is a product of atoms and anti-atoms.*

PROOF. The relation "divides" is a quasi-order on S, and modulo the relation $\sim$, we get the associated partial order $<$. If $s \in S$ and s has no atom as effective divisor, then s is either an atom or an anti-atom, and we are done. In the other case, $s = x a_1 y$ for some atom a_1, and say $s \not\sim x$. Then $x < s$. We factor x by an anti-atom a_2 if possible, and continue in this way. The process must stop, since S cannot have an infinite descending chain in the partial order. Finally, $s = \prod x_i a_i y_i$ where the a_i's are atoms and the x_i, y_i's are anti-atoms (or 1).

As an example, the set of idempotents can be chosen to be the set of atoms and thus we get that every element is a product of idempotents and anti-idempotents. In this way we can get a more systematic description of all semigroups of given (small) order. (For a complete list of all semigroups of order 2, 3 and 4 see, e.g., Petrich [16]; the total number of these semigroups is 4, 18, and 126, respectively.)

If the set of atoms A is the set of irreducibles together with the units, then (as we mentioned before) every semigroup can be embedded into a semigroup where every element is an atom = irreducible or unit. In case atoms are the idempotents then every semigroup can be embedded in a semigroup in which every member is a product of two atoms = two idempotents (see, e.g., Higgins [8] and Pastijn [15]). If atoms = divisible elements then by the above mentioned result of Sutov [23] every semigroup can be embedded into a semigroup where every element is an atom = divisible.

Let us close this paper with a general problem.

PROBLEM 4.1. For what kinds of atoms is it true that every semigroup can be embedded into a semigroup S where every element is a product of atoms?

ACKNOWLEDGEMENT. The authors are indebted to I. Z. Ruzsa for many helpful suggestions. The work of the third author was supported by a BGSU FRC Grant and by Hungarian NSF Grant OTKA T-016237.

REFERENCES

[1] ANDERSON, D. D., ANDERSON, D. F. and ZAFRULLAH, M., Factorization in integral domains, *J. Pure Appl. Algebra* **69** (1990), 1–19. *MR* **92b**:13028

[2] CLIFFORD, A. H and PRESTON, G. B., *The algebraic theory of semigroups*, Vol I, II, Mathematical Surveys, No. 7, American Mathematical Society, Providence, RI, 1961, 1967. *MR* **24** #A2627, **36** #1558

[3] EVANS, T., Embedding theorems for multiplicative systems and projective geometries, *Proc. Amer. Math. Soc.* **3** (1952), 614–620. *MR* **14**, 347h

[4] GRAHAM, R. L., ROTHSCHILD, B. L. and SPENCER, J. H., *Ramsey theory*, Wiley-Interscience Series in Discrete Mathematics, Wiley, New York, 1980. *MR* **82b**:05001

[5] GUBA, V. S., A finitely generated complete group, *Izv. Akad. Nauk SSSR Ser. Mat.* **50** (1986), 883–924 (in Russian). *MR* **88e**:20034

[6] HALL, M., The word problem for semigroups with two generators, *J. Symbolic Logic* **14** (1947), 115–118. *MR* **11**, 1f

[7] HALTER-KOCH, F., Finiteness theorems for factorizations, *Semigroup Forum* **44** (1992), 112–117. *MR* **92k**:20121

[8] HIGGINS, P. M., Embedding semigroups in semibands of depth two, *Semigroup Forum* **51** (1994), 395–396. *MR* **96f**:20096

[9] KHINTCHINE, A., Contribution à l'arithmétique des lois de distribution, *Bull. Univ. État Moscou, Sér. Int., Sect. A: Math. et Mécan.* **1** (1937), fasc. 1, 6–17. *Zbl* **18**, 316

[10] KUROSH, A. G., *The theory of groups*, Chelsea, New York, 1960. *MR* **22** #727

[11] KUROSH, A. G., *Lectures in general algebra*, Pergamon Press, Oxford–Edinburgh–New York, 1965. *MR* **31** #3483

[12] MAL'CEV, A. I., On homomorphisms into finite groups, *Lecture Notes of Ivanovsk. Ped. Inst.* **18** (1958), 49–60 (in Russian).

[13] NEUMANN, B. H., An essay on free products of groups with amalgamations, *Philos. Trans. Roy. Soc. London Ser. A* **246** (1954), 503–554. *MR* **16**, 10d

[14] OL'SHANSKII, A. YU., *Geometry of defining relations in groups*, Nauka, Moscow, 1989 (in Russian). *MR* **91i**:20035 (English translation: Kluwer, 1991).

[15] PASTIJN, F., Embedding semigroups in semibands, *Semigroup Forum* **14** (1977), 247–263. *MR* **56** #12153

[16] PETRICH, M., *Introduction to semigroups*, Merrill Research and Lecture Series, Charles E. Merrill Publishing Co., Columbus, OH, 1973. *MR* **52** #14016

[17] RÉDEI, L., *The theory of finitely generated commutative semigroups*, Pergamon Press, Oxford–Edinburgh–New York, 1965. *MR* **32** #5761

[18] RESTIVO, A. and REUTENAUER, C., On the Burnside problem for semigroups, *J. Algebra* **89** (1984), 102–104. *MR* **85k**:20188

[19] RUZSA, I. Z. and SZÉKELY, G. J., Theory of decomposition in semigroups, *Adv. in Math.* **56** (1985), 9–27. *MR* **86h**:43001

[20] RUZSA, I. Z. and SZÉKELY, G. J., A note on our paper "Theory of decomposition in semigroups", *Adv. in Math.* **60** (1986), 235–236. *MR* **87h**:43002

[21] RUZSA, I. Z. and SZÉKELY, G. J., *Algebraic probability theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, Chichester–New York, 1988. *MR* **89j**:60006

[22] SCHEIN, B. M., Homomorphism and subdirect decompositions of semigroups, *Pacific J. Math.* **17** (1966), 529–547. *MR* **33** #5768

[23] SUTOV, E. G., Embedding of semigroups into simple and complete semigroups, *Mat. Sb. (N. S.)* **62 (104)** (1963), 496–511 (in Russian). *MR* **30** #2100

[24] GÖNDŐCS, F., MICHALETZKY, G., MÓRI, T. F. and SZÉKELY, G. J., A characterization of infinitely divisible Markov chains with finite state space, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **27** (1984), 137–141. *MR* **87m**:60145

[25] TAMURA, T., Minimal commutative divisible semigroups, *Bull. Amer. Math. Soc.* **69** (1963), 713–716. *MR* **27** #1521

(Received August 25, 1997)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH FLORIDA
TAMPA, FL 33620-5700
U.S.A.

eclark@math.usf.edu

DEPARTMENT OF MATHEMATICS
BOWLING GREEN STATE UNIVERSITY
BOWLING GREEN, OH 43403
U.S.A.

chollan@bgnet.bgsu.edu

DEPARTMENT OF MATHEMATICS
BOWLING GREEN STATE UNIVERSITY
BOWLING GREEN, OH 43403
U.S.A.

gabors@bgnet.bgsu.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 25–39

A STRONG INVARIANCE PRINCIPLE FOR THE LOCAL TIME DIFFERENCE OF A SIMPLE SYMMETRIC PLANAR RANDOM WALK

E. CSÁKI[1], A. FÖLDES[2] and P. RÉVÉSZ

Dedicated to the memory of Alfréd Rényi

Abstract

Let $\xi(\mathbf{a}, n)$ be the local time at $\mathbf{a}$ of the simple symmetric random walk on the plane. Our main result says, that the difference $\xi(\mathbf{a}, n) - \xi(\mathbf{0}, n)$ can be strongly approximated by $\sigma_{\mathbf{a}} W(\xi^{(1)}(\mathbf{0}, n))$ where $\xi(\mathbf{0}, n)$ and $\xi^{(1)}(\mathbf{0}, n)$ have the same distribution and the latter is independent from $W(\cdot)$.

1. Introduction and main results

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.-s with

$$\mathbf{P}(X_1 = (0,1)) = \mathbf{P}(X_1 = (0,-1)) = \mathbf{P}(X_1 = (1,0)) = \mathbf{P}(X_1 = (-1,0)) = \frac{1}{4}$$

and let $S_0 = \mathbf{0}$, $S_n = X_1 + X_2 + \cdots + X_n$ $(n = 1, 2, \ldots)$ be a random walk on $\mathbf{Z}^2$ $(\mathbf{0} = (0,0))$. Its local time is defined by

$$\xi(\mathbf{a}, n) = \#\{k;\ 0 < k \leqq n,\ S_k = \mathbf{a}\},$$

where $\mathbf{a} = (a_1, a_2)$ is a lattice point on the plane. The aim of the paper is to prove the following result.

THEOREM 1.1. *There is a probability space with*

(i) *a simple symmetric random walk process S_n with its two parameter local time $\xi(\mathbf{a}, n)$,*

(ii) *a standard Wiener process $\{W(t), t \geqq 0\}$*

(iii) *and a process*

$$\{\xi^{(1)}(\mathbf{0}, n),\ n = 0, 1, 2, \ldots\} \stackrel{\mathcal{D}}{=} \{\xi(\mathbf{0}, n),\ n = 0, 1, 2, \ldots\}$$

1991 *Mathematics Subject Classification.* Primary 60J55; Secondary 60F15, 60F17, 60J15.

Key words and phrases. Planar random walk, local time, invariance principle.

[1]Research supported by the Hungarian National Foundation for Scientific Research Grant No. T 016384 and T 019346.

[2]Research supported by a PSC CUNY Grant No. 6-67383.

such that for an arbitrary but fixed $\mathbf{a}$

(1.1) $$\xi(\mathbf{a},n)-\xi(\mathbf{0},n)=\sigma_{\mathbf{a}}W(\xi^{(1)}(\mathbf{0},n))+O(\log n)^{\frac{2}{5}} \qquad a.s.$$

(1.2) $$\xi(\mathbf{0},n)=\xi^{(1)}(\mathbf{0},n)+O(\log n)^{\frac{4}{5}} \qquad a.s.$$

as $n\to\infty$, *where the processes* $\xi^{(1)}(\mathbf{0},n)$ *and* $\{W(t),\ t\geqq 0\}$ *are independent and* $\sigma_{\mathbf{a}}$ *is a constant depending on* $\mathbf{a}$.

This result has a long history. Denoting by $\xi_1(a,n)$ the local time of a simple symmetric random walk on the line, we have

THEOREM A (Dobrushin [8]). *For any fixed integer* $a\neq 0$

(1.3) $$\frac{\xi_1(a,n)-\xi_1(0,n)}{(4|a|-2)^{1/2}n^{1/4}}\xrightarrow{\mathcal{D}} U\sqrt{|V|}$$

as $n\to\infty$, *where* U *and* V *are two independent standard normal variables and* $\xrightarrow{\mathcal{D}}$ *denotes convergence in distribution.*

As stated here, (1.3) is only a special case of Dobrushin's theorem. It has several generalizations mostly for Brownian local time in one dimension; see Borodin [2], Kasahara [12], Papanicolaou et al. [18], Yor [21], Csörgő and Révész [7], Csáki and Földes [5]. The corresponding one dimensional result (in fact much more) was proved, and generalized for additive functionals in Csáki et al. [3], [4]. A weak convergence version of our present theorem was proved by Kesten [14] and Kasahara [13].

2. Preliminary results

Our theorem heavily relies on some basic results concerning the local time of the simple symmetric planar random walk.

THEOREM B (Erdős and Taylor [10]).

(2.1) $$\lim_{n\to\infty}\mathbf{P}(\xi(\mathbf{0},n)<x\log n)=1-e^{-\pi x}$$

uniformly for $0\leqq x\leqq(\log n)^{3/4}$, *and*

(2.2) $$\limsup_{n\to\infty}\frac{\pi\xi(\mathbf{0},n)}{\log n\log_3 n}=1 \qquad a.s..$$

Introduce

(2.3) $$\begin{aligned}&\rho_0=0,\\ &\rho_k=\inf\{n;\ n>\rho_{k-1},\ S_n=\mathbf{0}\},\quad k=1,2,\dots\end{aligned}$$

the consecutive return times of the planar random walk to the origin. The portion of the random walk between ρ_{k-1} and ρ_k is called the k-th excursion.

THEOREM C (Dvoretzky and Erdős [9], Erdős and Taylor [10]).

$$\mathbf{P}(\rho_1 > n) = \mathbf{P}(\xi(\mathbf{0}, n) = 0) = \frac{\pi}{\log n} + O((\log n)^{-2}). \tag{2.4}$$

For any lattice point $\mathbf{a}$ on the plane put

$$\xi_i(\mathbf{a}) = \xi(\mathbf{a}, \rho_i) - \xi(\mathbf{a}, \rho_{i-1}) \qquad i = 1, 2, \ldots. \tag{2.5}$$

Then $\{\xi_i(\mathbf{a})\}_{i=1}^{\infty}$ is a sequence of i.i.d. r.v.-s and

THEOREM D (Auer [1]).

$$\begin{aligned} &\mathbf{P}(\xi_i(\mathbf{a}) = 0) = q(\mathbf{a}) \\ &\mathbf{P}(\xi_i(\mathbf{a}) = l + 1) = p^2(\mathbf{a})q^l(\mathbf{a}) \qquad l = 0, 1, 2 \ldots, \end{aligned} \tag{2.6}$$

where $p(\mathbf{a}) = \mathbf{P}(S_{\cdot}$ *reaches* $\mathbf{a} \in \mathbf{Z}^2$ *before returning to the origin*$)$, $q(\mathbf{a}) = 1 - p(\mathbf{a})$. *Furthermore*

$$\mathbf{E}(\xi_i(\mathbf{a})) = 1 \quad \textit{and} \quad \sigma_{\mathbf{a}}^2 = \mathbf{Var}\, \xi_i(\mathbf{a}) = \frac{2(1 - p(\mathbf{a}))}{p(\mathbf{a})}. \tag{2.7}$$

REMARK. We can infer from Spitzer [20], Chapter 3, that $p(\mathbf{a}) = 1/(2y(\mathbf{a}))$, where $y(\mathbf{a})$ is the potential kernel of the random walk. For the simple symmetric case, with $\mathbf{a} = (a_1, a_2)$,

$$y(\mathbf{a}) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1 - \cos(a_1\theta_1 + a_2\theta_2)}{1 - \frac{1}{2}(\cos\theta_1 + \cos\theta_2)}\, d\theta_1\, d\theta_2. \tag{2.8}$$

LEMMA 2.1. *Let* $a_n = \exp((\log n)^K)$ *and* $b_n = \exp((\log n)^\gamma)$, *where* $K > 0$, $\gamma > 0$ *are arbitrary. Then for any* $1 < \eta$

$$\sup_{a \leqq a_n} (\xi(\mathbf{0}, a + b_n) - \xi(\mathbf{0}, a)) = O((\log n)^{\gamma\eta}) \qquad a.s.. \tag{2.9}$$

PROOF. Observe first that

$$\sup_{a \leqq a_n} (\xi(\mathbf{0}, a + b_n) - \xi(\mathbf{0}, a)) \leqq \sup_{i;\, \rho_i \leqq a_n} (\xi(\mathbf{0}, \rho_i + b_n) - \xi(\mathbf{0}, \rho_i)), \tag{2.10}$$

and the number of terms in the supremum on the right-hand side is $\xi(\mathbf{0}, a_n)$. Moreover

$$\mathbf{P}\Big(\sup_{i;\, \rho_i \leqq a_{n_k}} (\xi(\mathbf{0}, \rho_i + b_{n_k}) - \xi(\mathbf{0}, \rho_i)) > (\log n_{k-1})^{\gamma\eta}\Big)$$

$$\leqq \mathbf{P}\Big(\sup_{i;\rho_i \leqq a_{n_k}} (\xi(\mathbf{0}, \rho_i + b_{n_k}) - \xi(\mathbf{0}, \rho_i)) > (\log n_{k-1})^{\gamma\eta},\ \xi(\mathbf{0}, a_{n_k}) < (\log a_{n_k})^{3/2}\Big)$$

$$+\mathbf{P}\Big(\xi(\mathbf{0}, a_{n_k}) \geqq (\log a_{n_k})^{3/2}\Big),$$

where $n_k = e^k$. Using now Theorem B twice, we get

$$\mathbf{P}\left(\sup_{i:\, \rho_i \leqq a_{n_k}} (\xi(\mathbf{0}, \rho_i + b_{n_k}) - \xi(\mathbf{0}, \rho_i)) > (\log n_{k-1})^{\gamma\eta}\right)$$

$$\leqq k^{\frac{3}{2}K}\mathbf{P}\left(\xi(\mathbf{0}, b_{n_k}) > (k-1)^{\gamma\eta}\right) + C\exp(-\pi k^{\frac{K}{2}})$$

$$\leqq Ck^{\frac{3K}{2}}\exp\left(-\frac{\pi}{2}k^{\gamma(\eta-1)}\right) + C\exp(-\pi k^{\frac{K}{2}}),$$

where C is a constant (the value of which might change here and throughout the paper from line to line). Since the right-hand side of the above inequality is summable, we get our statement from Borel–Cantelli lemma combined with the usual monotonicity argument. □

3. Proof of Theorem 1.1

Now assume that $\{S_k^{(1)}\}_{k=0}^{\infty}$ and $\{S_k^{(2)}\}_{k=0}^{\infty}$ are two independent simple symmetric planar random walks, and denote their respective local times at $\mathbf{a}$ by $\xi^{(1)}(\mathbf{a}, n)$ and $\xi^{(2)}(\mathbf{a}, n)$, respectively and define $\rho_i^{(j)}$ $(j=1,2)$ by (2.3) with S replaced by $S^{(j)}$ and similarly $\xi_i^{(j)}$, $j=1,2$ by (2.5) with ξ replaced by $\xi^{(j)}$ and ρ_i replaced by $\rho_i^{(j)}$. Put

$$Z_i^{(j)} = \rho_i^{(j)} - \rho_{i-1}^{(j)}, \qquad j=1,2, \qquad i=1,2,\dots. \tag{3.1}$$

We will define a new simple symmetric planar walk $S_0, S_1, \dots$ to be constructed in blocks as follows. Let $A_k = 2^k$ and $r_k = A_k - A_{k-1} = 2^{k-1}$, $k = 1, 2\dots.$

The excursions with indices $A_{k-1} \leqq i < A_k$ form the k-th block. We build up our new walk as follows. Consider the excursions of the first and second walks in the k-th block and the lengths $Z_i^{(j)}$, $j=1,2$ of the consecutive excursions in the two walks. We label the excursions $Z_i^{(j)}$ large if

$$Z_i^{(j)} \geqq \exp(r_k^{\beta}), \qquad A_{k-1} \leqq i < A_k, \qquad j=1,2, \tag{3.2}$$

where the value of β will be selected later on. Denote by $\nu_k^{(j)}$, $\mu_k^{(j)}$ the number of large and small excursions, respectively in the k-th block. We create the k-th block of our new walk as follows. If $\mu_k^{(1)} \leqq \mu_k^{(2)}$ then we construct $\{S.\}$ by replacing all the $\mu_k^{(1)}$ small excursions from $\{S^{(1)}\}$ (within the block) by the first $\mu_k^{(1)}$ small excursions (first after $\rho_{A_{k-1}}$) from $\{S^{(2)}\}$ and leaving the $\nu_k^{(1)}$ large excursions from $\{S^{(1)}\}$ unaltered. Keep also the order of the consecutive large and small excursions as in $\{S^{(1)}\}$. If $\mu_k^{(2)} < \mu_k^{(1)}$ then construct $\{S_\cdot\}$ by replacing the first (after $\rho_{A_{k-1}}$) $\mu_k^{(2)}$ small excursions from $\{S^{(1)}\}$ by the small excursions from $\{S^{(2)}\}$ and leaving the $\nu_k^{(1)}$ large excursions and the other $\mu_k^{(1)} - \mu_k^{(2)}$ small excursions from $\{S^{(1)}\}$ unaltered. Performing this procedure for each block, we obtain a simple symmetric planar random walk $S_1, S_2, \ldots$. Now $\xi(\cdot,\cdot)$, ρ, $\xi_i(\cdot)$, and Z_i without superscript are the related quantities of this resulting random walk.

Introduce the notations

$$\hat{Z}_i^{(j)}(l,\beta) = Z_i^{(j)} \mathbf{1}\left(Z_i^{(j)} \leqq \exp\left(r_l^\beta\right)\right)$$
$$l = 1, 2, \ldots, \quad A_{l-1} \leqq i < A_l, \; j = 1, 2 \tag{3.3}$$

and

$$M_l(\mathbf{a}) = \max\left(\max_{A_{l-1}\leqq i<A_l}\left(\xi_i^{(1)}(\mathbf{a})\right), \max_{A_{l-1}\leqq i<A_l}\left(\xi_i^{(2)}(\mathbf{a})\right)\right), \tag{3.4}$$

and observe that for $A_{k-1} \leqq N < A_k$

$$\sup_{l\leqq N}|\rho_l - \rho_l^{(1)}| \leqq \sup_{l\leqq A_k}|\rho_l - \rho_l^{(1)}| \leqq \sum_{j=1}^{2}\sum_{l=1}^{k}\sum_{i=A_{l-1}+1}^{A_l}\hat{Z}_i^{(j)}(l,\beta) \tag{3.5}$$

and

$$\sup_{l\leqq N}|\xi(\mathbf{a},\rho_l) - \xi^{(2)}(\mathbf{a},\rho_l^{(2)})|$$
$$\leqq \sup_{l\leqq A_k}|\xi(\mathbf{a},\rho_l) - \xi^{(2)}(\mathbf{a},\rho_l^{(2)})| \leqq \sum_{l=1}^{k}2(\nu_l^{(1)} + \nu_l^{(2)})M_l(a). \tag{3.6}$$

We will prove a number of lemmas now.

LEMMA 3.1. *For any $\beta > 0$ and any integer $l \geqq 1$*

$$\mathbf{E}\left(\hat{Z}_i^{(j)}(l,\beta)\right) < C\exp(r_l^\beta)r_l^{-\beta}, \quad j = 1, 2, \quad A_{l-1} \leqq i < A_l. \tag{3.7}$$

PROOF. The proof is based on Theorem C and the following simple identity. For any $L \geqq 3$

$$\sum_{k=2}^{L}\frac{1}{\log k} \leqq C\frac{L}{\log L}. \tag{3.8}$$

By Theorem C and (3.8)

$$\mathbf{E}\left(\hat{Z}_i^{(j)}(l,\beta)\right) \leqq 2 + \sum_{k=3}^{[\exp(r_l^\beta)]} \mathbf{P}(Z_i^{(j)} \geqq k) \leqq 2 + \sum_{k=2}^{[\exp(r_l^\beta)]+1} \mathbf{P}(Z_i^{(j)} > k) \leqq$$

$$(3.9) \qquad 2 + \sum_{k=2}^{[\exp(r_l^\beta)]+1} \left(\frac{\pi}{\log k} + O\left(\frac{1}{(\log k)^2}\right)\right) \leqq C \sum_{k=2}^{[\exp(r_l^\beta)]+1} \frac{1}{\log k} \leqq C \exp(r_l^\beta) r_l^{-\beta}.$$

□

LEMMA 3.2. *For any* $\beta > 0$

$$(3.10) \qquad \sup_{l \leqq N} |\rho_l - \rho_l^{(1)}| \leqq \exp\left((4N)^\beta\right) \qquad a.s.$$

if N *is big enough.*

PROOF. Observe that by Markov's inequality and (3.9)

$$(3.11) \qquad \mathbf{P}\left(\sum_{i=A_{l-1}+1}^{A_l} \hat{Z}_i^{(j)}(l,\beta) > r_l^\gamma \exp(r_l^\beta)\right) \leqq C 2^{(l-1)(1-\gamma-\beta)}.$$

Select now $\gamma > 1$ to conclude that

$$\sum_{l+1}^{\infty} 2^{(l-1)(1-\gamma-\beta)} < \infty,$$

implying by the Borel–Cantelli lemma, that for $l \geqq l_0(\omega)$

$$(3.12) \qquad \sum_{i=A_{l-1}+1}^{A_l} \hat{Z}_i^{(j)}(l,\beta) \leqq r_l^\gamma \exp(r_l^\beta) \quad \text{a.s.}.$$

Hence by (3.5), for k big enough,

$$(3.13) \qquad \sup_{l \leqq A_k} |\rho_l - \rho_l^{(1)}| \leqq 2C \sum_{l=1}^{k} 2^{l\gamma} \exp(2^{l\beta}) \leqq \exp(2^{(k+1)\beta}) \quad \text{a.s.}.$$

Observe that for $2^{k-1} \leqq N < 2^k$ we have

$$(3.14) \qquad \sup_{l \leqq N} |\rho_l - \rho_l^{(1)}| \leqq \exp(2^{(k+1)\beta}) \leqq \exp\left((4N)^\beta\right) \quad \text{a.s.}$$

for N big enough, as $2^{k+1} \leqq 4N$. □

LEMMA 3.3. *For any fixed $\beta > 0$*

$$\mathbf{P}\left(\nu_l^{(1)} + \nu_l^{(2)} > 17 r_l^{1-\beta}\right) \leqq \exp\left(-r_l^{1-\beta}\right) \tag{3.15}$$

if l is large enough.

PROOF. Clearly $\nu_l^{(1)}$ and $\nu_l^{(2)}$ are both binomial with parameters r_l and $p_l = \mathbf{P}(\rho_1 > \exp(r_l^\beta))$ and independent, hence their sum is binomial with parameters $2r_l$ and p_l. Now observe that

$$\begin{aligned} &\mathbf{P}(\nu_l^{(1)} + \nu_l^{(2)} > A) \leqq \exp(-A)\mathbf{E}(\exp(\nu_l^{(1)} + \nu_l^{(2)})) \\ &= \exp(-A)(1 + p_l(e-1))^{2r_l} \leqq \exp(-A)\exp(4p_l r_l) \\ &\leqq \exp(-A + 5\pi r_l^{1-\beta}) \leqq \exp(-A + 16 r_l^{1-\beta}), \end{aligned} \tag{3.16}$$

where the last but one step follows from Theorem C. Select now $A = 17 r_l^{1-\beta}$ to get the lemma. □

LEMMA 3.4. *For any $0 < \beta < 1$, $\tau > 0$, such that $\beta + \tau > 1$ we have*

$$\sup_{i<N} |\xi(\mathbf{a}, \rho_i) - \xi^{(2)}(\mathbf{a}, \rho_i^{(2)})| = O(N^\tau) \qquad a.s.. \tag{3.17}$$

PROOF. By Lemma 3.3 and (2.6)

$$\mathbf{P}(2(\nu_l^{(1)} + \nu_l^{(2)})M_l(\mathbf{a}) > r_l^\tau) \leqq \exp(-r_l^{1-\beta}) + C(\mathbf{a}) r_l \exp(-d(\mathbf{a}) r_l^{\tau+\beta-1}) \tag{3.18}$$

where $C(\mathbf{a})$, $d(\mathbf{a})$ are positive constants depending only on $\mathbf{a}$. Now (3.18) implies by Borel–Cantelli lemma that

$$2(\nu_l^{(1)} + \nu_l^{(2)})M_l(\mathbf{a}) = O(1)2^{l\tau} \qquad \text{a.s.}, \tag{3.19}$$

hence by (3.6)

$$\sup_{i<A_k} |\xi(\mathbf{a}, \rho_i) - \xi^{(2)}(\mathbf{a}, \rho_i)| = \sum_{i=1}^{k} O(1)2^{i\tau} = O(1)2^{k\tau} \qquad \text{a.s.} \tag{3.20}$$

from which (3.17) follows. □

LEMMA 3.5. *On a rich enough probability space there exists a Wiener process $\{W(s),\ s \geqq 0\}$ such that for any fixed $\mathbf{a}$, β, τ satisfying $0 < \beta < 1$, $\tau > 0$, $\beta + \tau > 1$, we have*

$$|\xi(\mathbf{a}, \rho_N) - N - \sigma_{\mathbf{a}} W(N)| = O(N^\tau) \qquad a.s.. \tag{3.21}$$

PROOF. Apply the celebrated Komlós–Major–Tusnády [16] theorem for

$$\xi^{(2)}(\mathbf{a}, \rho_N^{(2)}) - N$$

(which is now considered as a sum of N i.i.d. random variables with common expectation 1). The existence of the moment generating function of the summands follows from Theorem D. Hence we get

$$(3.22)\qquad |\xi^{(2)}(\mathbf{a},\rho_N^{(2)})-N-\sigma_{\mathbf{a}}W(N)|=O(\log N)\qquad \text{a.s..}$$

Now apply Lemma 3.4 to get (3.21). □

A simple consequence of our Lemma 3.5 and Theorem B is that for any $\epsilon>0$, under the condition of Lemma 3.5

$$(3.23)\qquad \xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})-\xi(\mathbf{0},n)=\sigma_{\mathbf{a}}W(\xi(\mathbf{0},n))+O(\log n)^{\tau(1+\epsilon)}\qquad \text{a.s..}$$

However, to get the first statement of our theorem, we need to replace $\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})$ and $W(\xi(\mathbf{0},n))$ with $\xi(\mathbf{a},n)$ and $W(\xi^{(1)}(\mathbf{0},n))$, respectively. To be able to perform these replacements we need two further lemmas.

LEMMA 3.6. *For any* $\alpha>0$

$$(3.24)\qquad |\xi(\mathbf{a},n)-\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})|=O(\log n)^{\alpha}\qquad \textit{a.s.}$$

as $n\to\infty$.

PROOF. Consider the local time differences

$$\xi(\mathbf{a},\rho_{k+1})-\xi(\mathbf{a},\rho_k),\quad k=0,1,2,\dots.$$

They are i.i.d. random variables and from Theorem D, for an arbitrary $\alpha'>0$,

$$\sum_{k=1}^{\infty}\mathbf{P}(\xi(\mathbf{a},\rho_{k+1})-\xi(\mathbf{a},\rho_k)>k^{\alpha'})<\infty$$

implying that

$$(3.25)\qquad \xi(\mathbf{a},\rho_{k+1})-\xi(\mathbf{a},\rho_k)=O(k^{\alpha'})\qquad \text{a.s..}$$

Observe now that

$$(3.26)\qquad \xi(\mathbf{a},n)-\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})\leqq\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)+1})-\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})$$

and, by Theorem B, $\xi(\mathbf{0},n)\leqq(\log n)^{1+\epsilon}$ for any $\epsilon>0$ and n big enough. Consequently

$$(3.27)\qquad |\xi(\mathbf{a},n)-\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})|=O(\log n)^{\alpha'(1+\epsilon)}\qquad \text{a.s..}$$

Select now $\alpha=\alpha'(1+\epsilon)$ to get the lemma. □

LEMMA 3.7. *For* $\beta>0$ *and* $\epsilon>0$

$$(3.28)\qquad |\xi^{(1)}(\mathbf{0},n)-\xi(\mathbf{0},n)|=O(\log n)^{\beta(1+\epsilon)}\qquad \textit{a.s..}$$

PROOF. First observe that

$$\xi^{(1)}(\mathbf{0},n)=\xi^{(1)}(\mathbf{0},\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)})$$

and as $\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)} \leqq n$,

$$\begin{aligned}\xi^{(1)}(\mathbf{0},n)-\xi(\mathbf{0},n) &\leqq \xi^{(1)}(\mathbf{0},\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)})-\xi(\mathbf{0},\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)})\\ &=\xi(\mathbf{0},\rho_{\xi^{(1)}(\mathbf{0},n)})-\xi(\mathbf{0},\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)}).\end{aligned} \tag{3.29}$$

Now observe that, by Theorem B, $\xi^{(1)}(\mathbf{0},n)<(\log n)^{1+\delta}$ for any $\delta>0$, if n is big enough. Consequently, by Lemma 3.2,

$$\begin{aligned}|\rho_{\xi^{(1)}(\mathbf{0},n)}-\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)}| &\leqq \sup_{i\leqq(\log n)^{(1+\delta)}} |\rho_i-\rho^{(1)}_i|\\ &\leqq \exp(\log n)^{(1+\delta')\beta} \quad \text{a.s.}\end{aligned} \tag{3.30}$$

with any $\delta'>\delta>0$. Now apply Lemma 2.1 with $b_n=\exp(\log n)^{(1+\delta')\beta}$ and $a_n=n$ (being $\rho^{(1)}_{\xi^{(1)}(\mathbf{0},n)}\leqq n$). Thus we get by (3.30) that

$$\xi^{(1)}(\mathbf{0},n)-\xi(\mathbf{0},n)=O((\log n)^{(1+\epsilon)\beta}) \qquad \text{a.s.}$$

where $\epsilon>\delta'>0$. Repeating the argument for $\xi(\mathbf{0},n)-\xi^{(1)}(\mathbf{0},n)$ we get the lemma. □

PROOF OF THE THEOREM. Based on the above lemmas we have now three local time processes, constructed in such a way that the $\xi^{(1)}(\cdot,n)$ and $\xi^{(2)}(\cdot,n)$ processes are independent. $W(\cdot)$ was constructed to correspond to the process $\xi^{(2)}(\cdot,n)$, hence it is independent from the process $\xi^{(1)}(\cdot,n)$. Moreover, by Lemma 3.6, (3.23) and Lemma 3.7

$$\begin{aligned}\xi(\mathbf{a},n)-\xi(\mathbf{0},n) &=\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})-\xi(\mathbf{0},n)+\xi(\mathbf{a},n)-\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})\\ &=\xi(\mathbf{a},\rho_{\xi(\mathbf{0},n)})-\xi(\mathbf{0},n)+O(\log n)^{\alpha}\\ &=\sigma_{\mathbf{a}}W(\xi(\mathbf{0},n))+O(\log n)^{\tau(1+\epsilon)}+O(\log n)^{\alpha}\\ &=\sigma_{\mathbf{a}}W(\xi^{(1)}(\mathbf{0},n))+\sigma_{\mathbf{a}}(W(\xi(\mathbf{0},n))-W(\xi^{(1)}(\mathbf{0},n)))\\ &\quad +O(\log n)^{\tau(1+\epsilon)}+O(\log n)^{\alpha}\\ &=\sigma_{\mathbf{a}}W(\xi^{(1)}(\mathbf{0},n))+O(\log n)^{\frac{\beta(1+\epsilon')}{2}}\\ &\quad +O(\log n)^{\tau(1+\epsilon)}+O(\log n)^{\alpha} \quad \text{a.s.}\end{aligned} \tag{3.31}$$

for any $\epsilon'>\epsilon$. In the last line we used Theorem 1.2.1 of Csörgő and Révész [6] on the maximal increment of the Wiener process. Taking into account that the only conditions throughout the construction on β and τ were $0<\beta<1$, $\tau>0$, such that $\beta+\tau>1$ (and that $\alpha>0$, $\epsilon'>\epsilon>0$ are arbitrary small), we can select $\epsilon>0$ small enough to have $\frac{6}{5}\frac{1}{1+\epsilon'}>1$, and select $\tau(1+\epsilon')=\frac{2}{5}$ and $\beta(1+\epsilon')=\frac{4}{5}$ to get

$$\xi(\mathbf{a},n)-\xi(\mathbf{0},n)=\sigma_{\mathbf{a}}W(\xi^{(1)}(\mathbf{0},n))+O(\log n)^{\frac{2}{5}} \quad \text{a.s.,} \tag{3.32}$$

and from (3.28)

$$\xi^{(1)}(\mathbf{0},n)-\xi(\mathbf{0},n)=O(\log n)^{\frac{4}{5}} \quad \text{a.s.} \tag{3.33}$$

that, in turn, also proves our theorem. □

4. Applications

In this section we apply our Theorem 1.1 to obtain some limit theorems for $\xi(\mathbf{a},n)-\xi(\mathbf{0},n)$. It follows from (1.1) that the limit distribution of

$$\frac{\xi(\mathbf{a},n)-\xi(\mathbf{0},n)}{\sigma_{\mathbf{a}}\sqrt{\log n}}$$

should be the same as that of

$$\frac{W(\xi^{(1)}(n))}{\sqrt{\log n}},$$

where $\xi^{(1)}(n)=\xi^{(1)}(\mathbf{0},n)$.

Obviously, from Theorem B,

$$\frac{W(\xi^{(1)}(n))}{\sqrt{\log n}}=\frac{W(\xi^{(1)}(n))}{\sqrt{\xi^{(1)}(n)}}\sqrt{\frac{\xi^{(1)}(n)}{\log n}}\xrightarrow{\mathcal{D}}U\sqrt{Z}$$

as $n\to\infty$, where U is a standard normal r.v. and Z is an exponential r.v. with parameter π, and U and Z are independent. One can obtain similarly

$$\frac{|W(\xi^{(1)}(n))|}{\sqrt{\log n}}\xrightarrow{\mathcal{D}}|U|\sqrt{Z},$$

$$\frac{\sup_{k\le n}W(\xi^{(1)}(k))}{\sqrt{\log n}}\xrightarrow{\mathcal{D}}|U|\sqrt{Z},$$

and

$$\frac{\sup_{k\le n}|W(\xi^{(1)}(k))|}{\sqrt{\log n}}\xrightarrow{\mathcal{D}}T\sqrt{Z}$$

as $n \to \infty$, where T has the distribution of $\sup_{s \leq 1}|W(s)|$ and is independent of Z. It can be seen furthermore that the distribution of $U\sqrt{Z}$ is two-sided exponential with parameter $\sqrt{2\pi}$, i.e., its density function is

$$g(x) = \sqrt{\frac{\pi}{2}} e^{-|x|\sqrt{2\pi}} \qquad -\infty < x < \infty. \tag{4.1}$$

The distribution of $|U|\sqrt{Z}$ is exponential with parameter $\sqrt{2\pi}$.

Furthermore by using the formula (cf., e.g., Révész [19])

$$\begin{aligned} \mathbf{P}(T \leqq x) &= \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \exp\left(-\frac{\pi^2(2k+1)^2}{8x^2}\right) \\ &= \sum_{k=-\infty}^{\infty} (-1)^k \left(\Phi((2k+1)x) - \Phi((2k-1)x)\right), \end{aligned} \tag{4.2}$$

straightforward calculations give

$$\begin{aligned} H(x) = \mathbf{P}(T\sqrt{Z} \leqq x) &= \frac{32x^2}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^3\pi + 8(2k+1)x^2} \\ &= 1 - \frac{1}{\cosh(x\sqrt{2\pi})}. \end{aligned} \tag{4.3}$$

Hence we have the following limit distributions:

THEOREM 4.1.

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\xi(\mathbf{a}, n) - \xi(\mathbf{0}, n)}{\sigma_{\mathbf{a}}\sqrt{\log n}} \leqq x\right) = \int_{-\infty}^{x} g(u)\, du, \tag{4.4}$$

where $g(x)$ is given by (4.1),

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{|\xi(\mathbf{a}, n) - \xi(\mathbf{0}, n)|}{\sigma_{\mathbf{a}}\sqrt{\log n}} \leqq x\right) = 1 - \exp(-x\sqrt{2\pi}) \tag{4.5}$$

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\sup_{1\leqq k\leqq n} |\xi(\mathbf{a}, k) - \xi(\mathbf{0}, k)|}{\sigma_{\mathbf{a}}\sqrt{\log n}} \leqq x\right) = H(x). \tag{4.6}$$

We note that (4.4) was given by Kesten [14] and Kasahara [13].

By applying Theorem 1.1 we can also obtain strong limit theorems for $\xi(\mathbf{a},n)-\xi(\mathbf{0},n)$ via establishing the corresponding results for $W(\xi^{(1)}(n))$. Standard methods give the following laws of the iterated logarithm

$$\limsup_{n\to\infty}\frac{W(\xi^{(1)}(n))}{\sqrt{\log n\log_3 n}}=\frac{1}{\sqrt{2\pi}}\qquad \text{a.s.}$$

and this implies the LIL of Marcus and Rosen [17]:

$$\limsup_{n\to\infty}\frac{\xi(\mathbf{a},n)-\xi(\mathbf{0},n)}{\sqrt{\log n\log_3 n}}=\frac{\sigma_{\mathbf{a}}}{\sqrt{2\pi}}\qquad \text{a.s..}$$

Concerning liminf results, we first note that one can get the following asymptotics from (4.5) and (4.6):

$$\mathbf{P}(|U|\sqrt{Z}\leqq x)\sim x\sqrt{2\pi},\qquad x\to 0 \tag{4.7}$$

and

$$\mathbf{P}(T\sqrt{Z}\leqq x)\sim x^2\pi,\qquad x\to 0. \tag{4.8}$$

These results show that one cannot expect proper Chung-type LIL for $\sup_{k\leqq n}(\xi(\mathbf{a},k)-\xi(\mathbf{0},k))$ and $\sup_{k\leqq n}|\xi(\mathbf{a},k)-\xi(\mathbf{0},k)|$. The following integral tests, however, hold true.

THEOREM 4.2. (i) *Let $\alpha(x)$ be a nonincreasing function such that $\alpha(x)\sqrt{\log x}$ is nondecreasing. Then*

$$\mathbf{P}\left(\sup_{1\leqq k\leqq n}(\xi(\mathbf{a},k)-\xi(\mathbf{0},k))\leqq\alpha(n)\sqrt{\log n}\ \text{ i.o.}\right)=0\quad\text{or}\quad 1 \tag{4.9}$$

according as

$$\int_2^\infty\frac{\alpha(x)\,dx}{x\log x}<\infty\quad\text{or}\quad=\infty. \tag{4.10}$$

(ii) *Let $\beta(x)$ be a nonincreasing function such that $\beta(x)\sqrt{\log x}$ is nondecreasing. Then*

$$\mathbf{P}\left(\sup_{1\leqq k\leqq n}|\xi(\mathbf{a},k)-\xi(\mathbf{0},k)|\leqq\beta(n)\sqrt{\log n}\ \text{ i.o.}\right)=0\quad\text{or}\quad 1 \tag{4.11}$$

according as

$$\int_2^\infty\frac{\beta^2(x)\,dx}{x\log x}<\infty\quad\text{or}\quad=\infty. \tag{4.12}$$

PROOF. It follows from Theorem 1.1, that it suffices to prove similar results for $J(n) := \sup_{s \leqq \xi^{(1)}(n)} W(s)$ and $K(n) := \sup_{s \leqq \xi^{(1)}(n)} |W(s)|$. The convergent parts can be seen by considering the subsequence $n_k = \exp(e^k)$ and defining the events

$$C_k = \left\{ J(n_{k-1}) \leqq \alpha(n_k) \sqrt{\log n_k} \right\}$$

and

$$D_k = \left\{ K(n_{k-1}) \leqq \beta(n_k) \sqrt{\log n_k} \right\}.$$

It follows from (4.7) and (4.8) that $\mathbf{P}(C_k) \leqq c\alpha(n_k)$ and $\mathbf{P}(D_k) \leqq c\beta^2(n_k)$. But $\sum_k \alpha(n_k)$ and the integral in (4.10) and also $\sum_k \beta^2(n_k)$ and the integral in (4.12) are easily seen to be equiconvergent, hence the Borel–Cantelli lemma and the usual monotonicity arguments complete the proof of the convergent parts.

To show the divergent part we apply the following Lemma (cf. Klass [15]):

LEMMA 4.1. *Let $\{A_n\}_{n \geq 1}$ be an arbitrary sequence of events such that $\mathbf{P}(A_n$ i.o.$) = 1$ and let $\{B_n\}_{n \geq 1}$ be another sequence of events that is independent of $\{A_n\}_{n \geq 1}$ such that $\liminf_{n \to \infty} \mathbf{P}(B_n) \geqq p > 0$. Then we have $\mathbf{P}(A_n B_n$ i.o.$) \geqq p$.*

To show the divergent part of (i), let

$$A_n = \left\{ J(n) \leqq \sqrt{\xi^{(1)}(n)} \alpha(e^{\xi^{(1)}(n)}) \right\}$$

and

$$B_n = \{\xi^{(1)}(n) \leqq \log n\}.$$

The divergence of the integral in (4.10) implies $\int_1^\infty u^{-1} \alpha(e^u)\, du = \infty$ and hence by a theorem of Hirsch [11] (see also Révész [19]) we have

$$\mathbf{P}\left(\sup_{s \leqq k} W(s) \leqq \sqrt{k} \alpha(e^k) \ \text{ i.o.} \right) = 1.$$

Since $\xi^{(1)}(n)$ increases by 1, we also have $\mathbf{P}(A_n$ i.o.$) = 1$. Obviously $\lim_{n \to \infty} \mathbf{P}(B_n) > 0$ by Theorem B, hence Lemma 4.1 combined with 0-1 law for S_n proves the divergent part of (i).

To show the divergent part of (ii), let

$$\tilde{A}_n = \{\xi^{(1)}(n) \leqq \beta^2(n) \log n\}$$

and

$$\bar{B}_n = \left\{ \sup_{s \leqq \beta^{-2}(n) \log n} |W(s)| \leqq \beta(n) \sqrt{\log n} \right\}.$$

Erdős and Taylor [10] proved that $\mathbf{P}(\tilde{A}_n \text{ i.o.}) = 1$ and, clearly, $\lim_{n\to\infty} \mathbf{P}(\bar{B}_n) > 0$. Hence Lemma 4.1 combined with 0-1 law for S_n implies the divergent part of (ii). This completes the proof of our Theorem 4.2. □

REFERENCES

[1] Auer, P., The circle homogeneously covered by random walk on Z^2, *Statist. Probab. Lett.* **9** (1990), 403–407. *MR* **91g**:60080

[2] Borodin, A. N., On the character of convergence to Brownian local time. II, *Probab. Theory Related Fields* **72** (1986), 251–277. *MR* **87i**:60087

[3] Csáki, E., Csörgő, M., Földes, A. and Révész, P., Brownian local time approximated by a Wiener sheet, *Ann. Probab.* **17** (1989), 516–537. *MR* **90b**:60102

[4] Csáki, E., Csörgő, M., Földes, A. and Révész, P., Strong approximation of additive functionals, *J. Theoret. Probab.* **5** (1992), 679–706. *MR* **93k**:60073

[5] Csáki, E. and Földes, A., On the local time process standardized by the local time at zero, *Acta Math. Acad. Sci. Hungar.* **52** (1988), 175–186. *MR* **90c**:60049

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

[7] Csörgő, M. and Révész, P., On the stability of the local time of a symmetric random walk, *Acta Sci. Math. (Szeged)* **48** (1985), 85–96. *MR* **87c**:60056

[8] Dobrusin, R. L., Two limit theorems for the simplest random walk on a line, *Uspehi Mat. Nauk (N.S.)* **10** (1955), 139–146 (in Russian). *MR* **17**, 166b

[9] 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**, 852b

[10] Erdős, P. and Taylor, S. J., Some problems concerning the structure of random walk paths, *Acta Math. Acad. Sci. Hungar.* **11** (1960), 137–162. *MR* **22** #12599

[11] Hirsch, W. M., A strong law for the maximum cumulative sum of independent random variables, *Comm. Pure Appl. Math.* **18** (1965), 109–127. *MR* **31** #4069

[12] Kasahara, Y., Limit theorems for Lévy processes and Poisson point processes and their applications to Brownian excursions, *J. Math. Kyoto Univ.* **24** (1984), 521–538. *MR* **86i**:60091

[13] Kasahara, Y., A limit theorem for sums of random number of i.i.d. random variables and its application to occupation times of Markov chains, *J. Math. Soc. Japan* **37** (1985), 197–205. *MR* **86m**:60061

[14] Kesten, H., Occupation times for Markov and semi-Markov chains, *Trans. Amer. Math. Soc.* **103** (1962), 82–112. *MR* **25** #1569

[15] Klass, M., Toward a universal law of the iterated logarithm. I, *Z. Wahrsch. Verw. Gebiete* **36** (1976), 165–178. *MR* **54** #3822

[16] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent r.v.'s and the sample df. I, *Z. Wahrsch. Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[17] Marcus, M. B. and Rosen, J., Laws of the iterated logarithm for the local times of recurrent random walks on Z^2 and of Lévy processes and random walks in the domain of attraction of Cauchy random variables, *Ann. Inst. H. Poincaré Probab. Statist.* **30** (1994), 467–499. *MR* **95k**:60176

[18] PAPANICOLAOU, G. C., STROOCK, D. W. and VARADHAN, S. R. S., Martingale approach to some limit theorems, *Papers from the Duke Turbulence Conference* (Duke Univ., Durham, NC, 1976), Paper No. 6, Duke Univ. Math. Ser. III. Statistical Mechanics and Dynamical System, Duke Univ., Durham, NC, 1977. *MR* **57** #1669

[19] RÉVÉSZ, P., *Random walk in random and non-random environments*, World Scientific Publishing Co., Teaneck, NJ, 1990. *MR* **92c**:60096

[20] SPITZER, F., *Principles of random walk*, The University Series in Higher Mathematics, D. Van Nostrand, Princeton, 1964. *MR* **30** #1521

[21] YOR, M., Le drap Brownien comme limite en loi de temps locaux linéaires, *Séminaire de probabilités XVII*, Lecture Notes in Math. **986**, Springer, Berlin, 1983, 89–105. *MR* **86d**:60091

(Received May 29, 1997)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

csaki@math-inst.hu

DEPARTMENT OF MATHEMATICS
CITY UNIVERSITY OF NEW YORK
2800 VICTORY BLVD.
STATEN ISLAND, NY 10314
U.S.A.

afoldes@email.gc.cuny.edu

INSTITUT FÜR STATISTIK UND
WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
WIEDNER HAUPTSTRASSE 8–10/107
A–1040 WIEN
AUSTRIA

revesz@ci.tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 41–50

ON A PROBLEM ABOUT I-PROJECTIONS[1]

I. CSISZÁR and L. FINESSO

Dedicated to the memory of Alfréd Rényi

Abstract

The minimizer P^* of the I-divergence $I(P||Q)$ for P in a set Π defined by linear constraints is known to be mutually absolutely continuous with Q $(P^* \equiv Q)$ providing a $\hat{P}$ in Π exists with $\hat{P} \equiv Q$ and $I(\hat{P}||Q) < \infty$. We ask when the existence of $\overline{P}$ and P, both in Π, with $\overline{P} \equiv Q$ and $I(P||Q) < \infty$ is already sufficient for $P^* \equiv Q$. We give a positive answer for measures on a product space when Π is determined by prescribing the two marginals.

1. Introduction

For probability measures (p.m.'s) on a measurable space $(X, \mathcal{X})$, the I-divergence of P from Q (or relative entropy or Kullback–Leibler distance) is

$$I(P||Q) = \begin{cases} \int \varphi\left(\frac{dP}{dQ}\right) dQ & \text{if } P \ll Q \\ \infty & \text{otherwise,} \end{cases} \tag{1}$$

where

$$\varphi(t) = \begin{cases} t \log t & \text{if } t > 0 \\ 0 & \text{if } t = 0. \end{cases} \tag{2}$$

For a convex set Π of p.m.'s with

$$\inf_{P \in \Pi} I(P||Q) < \infty, \tag{3}$$

if the minimum of $I(P||Q)$ subject to $P \in \Pi$ is attained, the minimizer P^* is unique. It is called the I-projection of Q onto Π. A sufficient condition

1991 *Mathematics Subject Classification*. Primary 94A17, 62B10, 62H99.
Key words and phrases. I-divergence, iterative scaling, specified marginals, Schrödinger's problem.

[1] This work was supported by the CNR–MTA Cooperation Agreement under Grant no. 132.07.02.

for its existence is the closedness of Π in variation distance. An important property of the I-projection is

$$P \ll P^* \ll Q \quad \text{for all} \quad P \in \Pi \quad \text{with} \quad I(P||Q) < \infty. \tag{4}$$

For the above results cf. [3].

On account of (4), a (necessary and) sufficient condition for $P^* \equiv Q$ (where $\equiv$ denotes mutual absolute continuity) is the existence of some $\tilde{P} \in \Pi$ with $\tilde{P} \equiv Q$, $I(\tilde{P}||Q) < \infty$.

We will consider the problem whether the hypothesis $I(\tilde{P}||Q) < \infty$ can be dropped in the last condition, i.e. whether the following is true.

Assertion A. The existence of $\tilde{P} \in \Pi$ with $\tilde{P} \equiv Q$ implies $P^* \equiv Q$.

Notice that subject to (3), Assertion A is equivalent to the apparently weaker

Assertion B. The existence of $\tilde{P} \in \Pi$ with $\tilde{P} \equiv Q$ implies that to any $P \in \Pi$ with $P \not\equiv Q$, $I(P||Q) < \infty$, there exists $P' \in \Pi$ with $P' \not\ll P$, $I(P'||Q) < \infty$.

Indeed, were Assertion A false, applying Assertion B to $P = P^*$ would yield a $P' \in \Pi$ with $P' \not\ll P^*$, $I(P'||Q) < \infty$, contradicting (4).

For Π determined by a finite number of linear constraints, i.e.

$$\Pi = \left\{ P: \int f_i dP = c_i, \quad i = 1, 2, \dots, k \right\}$$

for given measurable functions f_i on X and constants c_i, $i = 1, 2, \dots, k$, Assertion A is true, cf. [4].

For Π determined by an infinite number of linear constraints, however, Assertion A may be false. For a counterexample let $X = N$ and Q and $\overline{P}$ arbitrary strictly positive p.m.'s with $D(\overline{P}||Q) = \infty$. Let Π be the set of p.m.'s P satisfying $p_0 + p_i(1 - \overline{p_0})/\overline{p_i} = 1$, $i = 1, 2, \dots$. Then the p.m. δ_0 concentrated at 0 is in Π and has $D(\delta_0||Q) < \infty$, but $D(P||Q) = \infty$ for all other $P \in \Pi$.

In this paper, we will concentrate on the case when Π is the set of those p.m.'s on a product space that have specified marginals. Let $\Pi(P_1, P_2)$ denote the (convex) set of those p.m.'s on $(X_1, \mathcal{X}_1) \times (X_2, \mathcal{X}_2)$ whose marginals are, respectively, P_1 and P_2. Since this set is closed in variation distance, the I-projection of Q onto $\Pi(P_1, P_2)$ exists for any p.m. Q on $(X_1, \mathcal{X}_1) \times (X_2, \mathcal{X}_2)$ satisfying the finiteness condition (3), i.e.,

$$\inf_{P \in \Pi(P_1, P_2)} I(P||Q) < \infty. \tag{5}$$

We will assume in the sequel that

$$Q \ll Q_1 \times Q_2, \tag{6}$$

where Q_1 and Q_2 denote the marginals of Q. Then the I-projection P^* of Q onto $\Pi(P_1, P_2)$ satisfies

$$\frac{dP^*}{dQ} = a(x_1)b(x_2) \qquad P^*\text{-a.s.,} \tag{7}$$

where a and b are strictly positive measurable functions on X_1 and X_2, respectively. The latter has been proved in [9], filling a gap in a proof of [3] (the proof in [3] assumed that the set of functions of form $f(x_1)+g(x_2)$, $f \in L_1(P_1)$, $g \in L_1(P_2)$ is closed in $L_1(P^*)$, which is not always true).

One reason for our special interest in the case $P^* \equiv Q$ is that then (7) becomes

$$\frac{dP^*}{dQ} = a(x_1)b(x_2) \qquad Q\text{-a.s.,} \tag{8}$$

where the identity holds in the same sense as the Radon–Nikodym derivative is determined.

The problem of finding positive measurable functions $a(x_1)$, $b(x_2)$ such that

$$\frac{dP}{dQ} = a(x_1)b(x_2) \quad \text{defines a p.m.} \quad P \in \Pi(P_1, P_2) \tag{9}$$

has first been raised, apparently, by Schrödinger [11] who arrived at it treating a problem about diffusions, of conceptual interest in physics. Since (8) provides a solution to this problem, we may assert – subject to hypotheses (5), (6) – that a solution to Schrödinger's problem (9) always exists if a $\bar{P} \in \Pi(P_1, P_2)$ with $\bar{P} \equiv Q$, $I(\bar{P}||Q) < \infty$ exists at all; if Assertion A holds for $\Pi(P_1, P_2)$ and Q, the hypothesis $I(\bar{P}||Q) < \infty$ can be dropped in the last condition. Relevant references about Schrödinger's problem include [1], [6], [9]. As shown in [6], solutions to (9) with $I(P||Q) < \infty$ necessarily satisfy $P = P^*$, though solutions with $I(P||Q) = \infty$ perhaps may exist.

REMARK. Assumption (6) is crucial for (8). An example where $P^* \equiv Q \not\ll Q_1 \times Q_2$ and dP^*/dQ is not representable as the product of measurable functions, appears in [6]. Still, such representation, with perhaps non-measurable $a(x_1)$ and $b(x_2)$, is always possible, cf. [1], [6].

2. The countable case

In this section we deal with p.m.'s on countable sets. The probability mass function (pmf) of a p.m. denoted by a capital letter will be denoted by the same lowercase letter. The support of a p.m. P on a countable set X is $S_P = \{x : p(x) > 0\}$. Notice that $P \ll Q$ is equivalent to $S_P \subset S_Q$.

In the countable case, to prove Assertion A for Π and Q satisfying (3), it suffices to prove the following assertion not involving Q.

ASSERTION C. To any P and $\tilde{P}$ in Π with $S_P \subset S_{\tilde{P}}$, $S_P \neq S_{\tilde{P}}$, there exists $P' \in \Pi$ with $S_{P'} \not\subset S_P$ such that the pmf of P' differs from that of P on a finite subset of $S_{\tilde{P}}$ only.

Indeed, if $S_{\tilde{P}} = S_Q$, the last condition makes sure that $I(P||Q) < \infty$ implies $I(P'||Q) < \infty$. Thus Assertion C implies Assertion B which, as noticed in the Introduction, is equivalent to Assertion A.

THEOREM 1. *If X_1 and X_2 are (finite or) countable sets, Assertion C always holds for* $\Pi(P_1, P_2)$, *hence so does Assertion A subject to the finiteness condition* (5).

REMARK. Hypothesis (6) is trivially satisfied in this case.

PROOF. Given P and $\tilde{P}$ in $\Pi = \Pi(P_1, P_2)$ as in Assertion C, notice first that a P' required there can be trivially given if for some $(\bar{x}_1, \bar{x}_2) \in S_{\tilde{P}} \setminus S_P$ there exists $\bar{x}_3 \in X_1$ and $\bar{x}_4 \in X_2$ such that $(\bar{x}_3, \bar{x}_2)$ and $(\bar{x}_1, \bar{x}_4)$ are in S_P and $(\bar{x}_3, \bar{x}_4)$ is in $S_{\tilde{P}}$. Indeed, changing the pmf of P at the points $(\bar{x}_1, \bar{x}_2), (\bar{x}_3, \bar{x}_2), (\bar{x}_3, \bar{x}_4), (\bar{x}_1, \bar{x}_4)$ only, by alternatingly adding and subtracting some $\epsilon > 0$, the resulting P' has the same marginals as P and meets all requirements in Assertion C.

This simple idea always works if $S_{\tilde{P}} = X_1 \times X_2$, but not necessarily otherwise. We will show, however, that to any $(\bar{x}_1, \bar{x}_2) \in S_{\tilde{P}} \setminus S_P$ one can always find $\bar{x}_3, \dots, \bar{x}_{2n-1}$ in X_1 and $\bar{x}_4, \dots, \bar{x}_{2n}$ in X_2 (for some $n \geqq 2$) such that

$$(\bar{x}_{2i-1}, \bar{x}_{2i}) \in S_{\tilde{P}}, \quad (\bar{x}_{2i+1}, \bar{x}_{2i}) \in S_P \quad i = 1, 2, \dots, n \tag{10}$$

with the convention $\bar{x}_{2n+1} = \bar{x}_1$. This still suffices to obtain a P' as in Assertion C, by changing the pmf of P at the $2n$ points in (10) only, by alternatingly adding and subtracting some $\epsilon > 0$.

Consider a Markov chain with state space $X_1 \cup X_2$ (assuming w.l.o.g. that $X_1 \cap X_2 = \emptyset$) and transition probabilities

$$p(y|x) = \begin{cases} \tilde{p}(x, y)/p_1(x) & \text{if } x \in X_1, y \in X_2, \ (x, y) \in S_{\tilde{P}} \\ p(y, x)/p_2(x) & \text{if } x \in X_2, y \in X_1, \ (y, x) \in S_P \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

This Markov chain has an invariant p.m. with pmf $p(x) = \frac{1}{2}p_1(x)$ or $\frac{1}{2}p_2(x)$ according as $x \in X_1$ or $x \in X_2$, thus $\bar{x}_1$ with $p_1(\bar{x}_1) > 0$ cannot be transient ([5], p. 395).

Since $\bar{x}_2$ can be reached from $\bar{x}_1$ (in one step, by the assumption $(\bar{x}_1, \bar{x}_2) \in S_{\tilde{P}}$), it follows that $\bar{x}_1$ can also be reached from $\bar{x}_2$. On account of (11), the latter is equivalent to the existence of $\bar{x}_3, \dots, \bar{x}_{2n}$ satisfying (10). This completes the proof of Theorem 1.

3. The general case

In this section we consider $\Pi(P_1, P_2)$ and Q as in Section 1 in the case of arbitrary $(X_1, \mathcal{X}_1)$, $(X_2, \mathcal{X}_2)$. Recall the hypothesis $Q \ll Q_1 \times Q_2$, cf. (6), and write (for any $P \ll Q_1 \times Q_2$)

$$\frac{dP}{d(Q_1 \times Q_2)} = p(x_1, x_2), \qquad \frac{dQ}{d(Q_1 \times Q_2)} = q(x_1, x_2). \tag{12}$$

For the densities p and q define

$$S_P = \{(x_1, x_2) : p(x_1, x_2) > 0\}, \qquad S_Q = \{(x_1, x_2) : q(x_1, x_2) > 0\}. \tag{13}$$

While we believe that Assertion A holds in this setting without any additional hypotheses, at present we are able to prove this under the following condition:

$$S_Q = \bigcup_{j \in J} (A_j \times B_j), \qquad A_j \in \mathcal{X}_1, B_j \in \mathcal{X}_2, \tag{14}$$

where J is a (finite or) countable index set. Since $q(x_1, x_2)$ is not a uniquely defined function on $X_1 \times X_2$, (14) is required for a suitable version of $q(x_1, x_2)$.

Notice that (14) is automatically satisfied if X_1 or X_2 is (finite or) countable, or if $Q \equiv Q_1 \times Q_2$. It is also satisfied if X_1 and X_2 are separable metric spaces endowed with the Borel σ-algebras, and $q(x_1, x_2)$ is continuous or lower-semicontinuous. The hypothesis (14) appeared previously in [1] where (8) was proved under that hypothesis, assuming (6) and the existence of $\bar{P} \in \Pi(P_1, P_2)$ with $\bar{P} \equiv Q$, $I(\bar{P}||Q) < \infty$.

THEOREM 2. *For $\Pi(P_1, P_2)$ and Q satisfying hypotheses* (5), (6) *and* (14), *Assertion A is true.*

COROLLARY 1. *Under hypotheses* (5), (6) *and* (14), *there exists a solution to Schrödinger's problem* (9) *whenever a $\bar{P} \in \Pi(P_1, P_2)$ with $\bar{P} \equiv Q$ exists.*

For the proof we need the following

LEMMA 1. *Given any $P \in \Pi(P_1, P_2)$ with $P \not\equiv Q$, $I(P||Q) < \infty$, there exists $P' \in \Pi(P_1, P_2)$ with $P' \not\ll P$, $I(P'||Q) < \infty$ if random variables $Y_1, \dots, Y_n$ with the following properties exist:*

(i) *Y_i takes values in X_1 or X_2 according as i is odd or even;*

(ii) *the joint distribution of (Y_{2i-1}, Y_{2i}) is absolutely continuous with respect to P, $i = 1, \dots, n$, the joint distribution of (Y_{2i+1}, Y_{2i}) is absolutely continuous with respect to Q, $i = 1, \dots, n-1$, and the joint distribution of (Y_1, Y_{2n}) is absolutely continuous with respect to $Q_1 \times Q_2$.*

(iii) $\Pr\{(Y_1, Y_{2n}) \in S_Q, (Y_3, Y_2) \in S_Q \setminus S_P\} > 0$.

PROOF. Write $Y_{2n+1} = Y_1$ for convenience, and set

$$C_{K,i} = \{(x_1, x_2) : p_{Y_{2i-1}, Y_{2i}}(x_1, x_2) < Kp(x_1, x_2)\} \qquad i = 1, \dots, n \tag{15}$$

$$D_{K,i} = \{(x_1, x_2) : p_{Y_{2i+1}, Y_{2i}}(x_1, x_2) < Kq(x_1, x_2)\} \qquad i = 1, \dots, n, \tag{16}$$

where $p_{Y_{2i-1}, Y_{2i}}$ and $p_{Y_{2i+1}, Y_{2i}}$ denote the joint densities of the indicated random variables with respect to $Q_1 \times Q_2$. Let $\mathcal{E}_K$ denote the event that

$$\begin{gathered} (Y_{2i-1}, Y_{2i}) \in C_{K,i}, \quad (Y_{2i+1}, Y_{2i}) \in D_{K,i}, \\ i = 1, \dots, n, \quad (Y_3, Y_2) \in S_Q \setminus S_P \end{gathered} \tag{17}$$

simultaneously hold. By assumptions (ii), (iii),

$$\lim_{k \to \infty} \Pr\{\mathcal{E}_K\} = \Pr\{(Y_1, Y_{2n}) \in S_Q, (Y_3, Y_2) \in S_Q \setminus S_P\} > 0. \tag{18}$$

Fix some K with $\Pr\{\mathcal{E}_K\} > 0$. Then (15), (16), (17) imply that the conditional densities of (Y_{2i-1}, Y_{2i}), respectively of (Y_{2i+1}, Y_{2i}), on the condition $\mathcal{E}_K$, are upper bounded by $Kp(x_1, x_2)/\Pr\{\mathcal{E}_K\}$ and $Kq(x_1, x_2)/\Pr\{\mathcal{E}_K\}$, respectively. It follows that

$$\begin{gathered} p'(x_1, x_2) = p(x_1, x_2) \\ + \epsilon \left(\sum_{i=1}^{n} p_{Y_{2i+1}, Y_{2i} | \mathcal{E}_K}(x_1, x_2) - \sum_{i=1}^{n} p_{Y_{2i-1}, Y_{2i} | \mathcal{E}_K}(x_1, x_2) \right) \end{gathered} \tag{19}$$

is the density of a p.m. P' if $0 < \epsilon < \Pr\{\mathcal{E}_K\}/Kn$, and that $I(P'||Q) < \infty$. It is obvious from (19) that the marginals of P' are equal to those of P, hence $P' \in \Pi(P_1, P_2)$. Finally, it follows from (19) and the definition (17) of $\mathcal{E}_K$ that

$$P'(S_Q \setminus S_P) \geqq \epsilon \Pr\{(Y_3, Y_2) \in S_Q \setminus S_P \mid \mathcal{E}_K\} = \epsilon > 0, \tag{20}$$

thus $P' \not\ll P$.

PROOF OF THEOREM 2. As noted in Section 1 it suffices to prove Assertion B. Now, fix $P \in \Pi(P_1, P_2)$ with $P \not\equiv Q$, $I(P||Q) < \infty$ and $\bar{P} \in \Pi(P_1, P_2)$ with $\bar{P} \equiv Q$. Let $Y_1, Y_2, \dots$ be a Markov chain satisfying Condition (i) of the Lemma such that the joint distribution of (Y_{2i-1}, Y_{2i}) is P and the joint distribution of (Y_{2i+1}, Y_{2i}) is $\bar{P}$, $i = 1, 2, \dots$. Then Condition (ii) of the Lemma is also satisfied, for each n, and it suffices to show that Condition (iii) is satisfied for some n. As the joint distribution of (Y_3, Y_2) is $\bar{P} \equiv Q$, and $P \ll Q$, $P \not\equiv Q$, we have $\Pr\{(Y_3, Y_2) \in S_Q \setminus S_P\} > 0$. Hence, using hypothesis

(14) and the fact that the joint distribution P of (Y_1, Y_2) satisfies $P \ll Q$, it follows that for some measurable rectangle $A \times B \subset S_Q$

(21) $$\Pr\{(Y_1, Y_2) \in A \times B,\ \ (Y_3, Y_2) \in S_Q \setminus S_P\} > 0.$$

Denote by $\mathcal{E}_n$ the event that

(22) $$(Y_{2n-1}, Y_{2n}) \in A \times B,\ \ (Y_{2n+1}, Y_{2n}) \in S_Q \setminus S_P.$$

Since the pairs $(Y_1, Y_2), (Y_3, Y_4), \dots$ form a stationary Markov process, $\mathcal{E}_1, \mathcal{E}_2, \dots$ is a stationary sequence of events of positive probability, cf. (21). Hence, by the Recurrence Theorem ([8], p. 27), there exists $n > 2$ (actually, infinitely many such n's) such that

(23) $$\Pr\{\mathcal{E}_1 \cap \mathcal{E}_n\} > 0.$$

If both $\mathcal{E}_1$ and $\mathcal{E}_n$ obtain then $(Y_1, Y_2) \in A \times B$ and $(Y_{2n-1}, Y_{2n}) \in A \times B$, cf. (22), hence also $(Y_1, Y_{2n}) \in A \times B$. Thus (23) implies that

$$\Pr\{(Y_1, Y_{2n}) \in S_Q,\ \ (Y_3, Y_2) \in S_Q \setminus S_P\} \geqq$$
$$\Pr\{(Y_1, Y_{2n}) \in A \times B,\ \ (Y_3, Y_2) \in S_Q \setminus S_P\} \geqq \Pr\{\mathcal{E}_1 \cap \mathcal{E}_n\} > 0.$$

This means that Condition (iii) of the Lemma is also satisfied (for a suitable $n > 2$). This completes the proof of Theorem 2. The Corollary is immediate from the discussion in the Introduction.

4. Extensions and open problems

Some simple extensions of our results are as follows:

(i) If no $\bar{P} \in \Pi$ satisfies $\bar{P} \equiv Q$, it is still of interest whether in equation (4) the condition $I(P||Q) < \infty$ can be relaxed to $P \ll Q$, i.e., whether the I-projection P^* of Q onto Π dominates every $P \in \Pi$ dominated by Q. In the countable case the latter is obviously implied by Assertion C, thus the corresponding extension of Theorem 1 is immediate. Similarly, an obvious modification of the proof of Theorem 2 gives that if $\bar{P} \in \Pi(P_1, P_2)$ dominates every $P \in \Pi(P_1, P_2)$ dominated by Q (the proof of the existence of such $\bar{P}$, unique up to mutual absolute continuity, is standard) and hypothesis (14) holds for $\bar{P}$ rather than Q, then $P^* \equiv \bar{P}$.

(ii) Corollary 1 about Schrödinger's problem can be extended to some cases when the finiteness condition (5) is not satisfied. Indeed, let $Q^{(1)}$ be a p.m. on $X_1 \times X_2$ with X_1-marginal P_1, obtained from Q by "scaling", i.e. having $Q_1 \times Q_2$-density $q^{(1)}(x_1, x_2) = q(x_1, x_2) \dfrac{dP_1}{dQ_1}(x_1)$. A $P \in \Pi(P_1, P_2)$

with $I(P||Q^{(1)})<\infty$ may exist even if none with $I(P||Q)<\infty$ does; this can happen if $I(P_1||Q_1)=\infty$, since clearly

$$I(P||Q)=I(P||Q^{(1)})+I(P_1||Q_1) \quad \text{for all} \quad P\in\Pi(P_1,P_2). \tag{24}$$

Applying Theorem 2 to $Q^{(1)}$ in the role of Q, it follows that the finiteness hypothesis (5) in the Corollary can be relaxed to

$$\inf_{P\in\Pi(P_1,P_2)} I(P||Q^{(1)})<\infty. \tag{25}$$

Notice that $I(P||Q^{(1)})$ can be interpreted as conditional I-divergence: taking the conditional distributions on X_2 given an $x_1\in X_1$ induced by P and Q, respectively, the average with respect to P_1 of their I-divergence is equal to $I(P||Q^{(1)})$.

The p.m. $Q^{(1)}$ can be further "scaled" to get $Q^{(2)}$ with X_2-marginal P_2 and $Q_1\times Q_2$-density $q^{(2)}(x_1,x_2)=q^{(1)}(x_1,x_2)\dfrac{dP_2}{dQ_2^{(1)}}(x_2)$ where $Q_2^{(1)}$ denotes the X_2-marginal of $Q^{(1)}$. This permits to relax (25) replacing $Q^{(1)}$ by $Q^{(2)}$.

These $Q^{(1)}$, $Q^{(2)}$ are the first two elements of the sequence of p.m.'s obtained by "iterative scaling", a procedure that in case of finite X_1 and X_2 is known [3] to converge to the I-projection P^* of Q onto $\Pi(P_1,P_2)$, when (5) is satisfied. In the present context, the hypothesis (5) in the Corollary of Theorem 2 can be relaxed beyond (25), replacing $Q^{(1)}$ there by any $Q^{(n)}$ of the iterative scaling procedure.

REMARK. A proof that the iterative scaling procedure converges to the I-projection P^* in the general case, subject only to (5), is still elusive. For substantial partial results cf. [1], [10].

(iii) Instead of I-divergence, we could have considered φ-divergences [2], [7] as well, letting φ in (1) be a continuous and strictly convex function on $[0,\infty)$ other than (2). The key property (4) and Theorems 1 and 2 remain valid whenever φ satisfies:

$$\lim_{t\to 0}\varphi'(t)=-\infty, \qquad \lim_{t\to\infty}\varphi'(t)=\infty. \tag{26}$$

The general problem raised in this paper appears unexpectedly difficult. Already for the three-dimensional analogue of $\Pi(P_1,P_2)$ in the countable case, viz. for the set of p.m.'s on $X_1\times X_2\times X_3$ with given one-dimensional marginals (or with given two-dimensional marginals), where X_1, X_2, X_3 are countable, we did not succeed in deciding whether Assertion A is true. It also remains open in what cases, if any, is Assertion A true, but Assertion C false.

A bold conjecture might be that Assertion C always holds for sets of p.m.'s on a countable set X defined by prescribing the probabilities of subsets

$A_1, A_2, \dots$ of X such that each $x \in X$ is contained in at most k distinct A_j's, for some constant k. Proving this conjecture (if true) apparently requires methods different from those in this paper. For the case $k=2$, however, the positive answer is an easy consequence of Theorem 1.

THEOREM 3. *Let $\mathcal{A}=(A_1, A_2, \dots)$ be a countable family of subsets of a countable set X such that each $x \in X$ belongs to at most two distinct A_j's. Then Assertion C holds for*

$$\Pi = \{P : P(A_j) = p_j, \quad j = 1, 2, \dots\}, \tag{27}$$

where p_j, $j=1,2,\dots$ are given positive numbers such that $\Pi \neq \emptyset$.

PROOF. Denote by A_0 the subset of X not covered twice by $\mathcal{A}$, i.e. the complement of the union of all pairwise intersections of sets in $\mathcal{A}$. Then the (non-empty ones among the) sets $X \setminus \bigcup_{j=1}^{\infty} A_j$ and $A_i \cap A_j$, $i \neq j$, i and j in $N = \{0, 1, \dots\}$. represent a partition of X. Given P and $\bar{P}$ as in Assertion C, a P' as required there can be trivially given if some of these sets with positive P-probability contains an x with $p(x)=0$. Henceforth we assume that this is not the case. Associate with the given P and $\bar{P}$ on X p.m.'s on N^2 letting

$$p(i,j) = \begin{cases} \frac{1}{2}P(A_i \cap A_j) & \text{if } i \neq j \\ P(X \setminus \bigcup_{j=1}^{\infty} A_j) & \text{if } i = j = 0 \\ 0 & \text{if } i = j \neq 0 \end{cases} \tag{28}$$

and similarly defining $\bar{p}(i,j)$. Clearly, $p(i,j)$ and $\bar{p}(i,j)$ are symmetric pmf's on N^2, both having marginals $P_1 = P_2$ with pmf $p_1(i)$ satisfying

$$p_1(i) = \frac{1}{2}P(A_i) = \frac{1}{2}\bar{P}(A_i) = \frac{1}{2}p_i, \qquad i = 1, 2, \dots \tag{29}$$

cf. (27). By our assumptions on P and $\bar{P}$, the associated p.m.'s on N^2 defined above satisfy the hypotheses of Assertion C, i.e. $\{(i,j) : p(i,j) > 0\}$ is a proper subset of $\{(i,j) : \bar{p}(i,j) > 0\}$. Hence by Theorem 1, there exists a p.m. on N^2 with both marginals equal to P_1 (uniquely defined by (29)) whose pmf $p'(i,j)$ differs from $p(i,j)$ on a finite subset of $\{(i,j) : \bar{p}(i,j) > 0\}$, and $p'(i,j)$ is positive for some $(i,j) \in N^2$ with $p(i,j) = 0$. Though Theorem 1 does not guarantee it, the symmetry of this p.m. on N^2 can always be assumed, else $p'(i,j)$ could be replaced by $p''(i,j) = \frac{1}{2}(p'(i,j) + p'(j,i))$, retaining the above properties. Then it is a simple matter to give a p.m. P' on X that satisfies the analogue of (28) with primes, and differs from P but on a finite subset of $S_{\bar{P}}$. The fact that the marginals of $p'(i,j)$ are equal to P_1, given by (29), means that $P'(A_i) = p_i$, $i = 1, 2, \dots$ hence P' belongs to Π defined by (27). Since clearly $S_{P'}$ is not a subset of S_P, the proof is complete.

REFERENCES

[1] Borwein, J. M., Lewis, A. S. and Nussbaum, R. D., Entropy minimization, DAD problems, and doubly stochastic kernels, *J. Funct. Anal.* **123** (1994), 264–307. *MR* **96b**:47067

[2] Csiszár, I., Information-type measures of difference of probability distributions and indirect observations, *Studia Sci. Math. Hungar.* **2** (1967), 299–318. *MR* **36** #2428

[3] Csiszár, I., I-divergence geometry of probability distributions and minimization problems, *Ann. Probab.* **3** (1975), 146–158. *MR* **51** #2050

[4] Csiszár, I., Sanov property, generalized I-projection and a conditional limit theorem, *Ann. Probab.* **12** (1984), 768–793. *MR* **86h**:60065

[5] Feller, W., *An introduction to probability theory and its applications*, vol. 1, Third ed., Wiley, New York, 1968. *MR* **37** #3604

[6] Föllmer, H. and Gantert, N., Entropy minimization and Schrödinger processes in infinite dimensions, *Ann. Probab.* **25** (1997), 901–926.

[7] Liese, F. and Vajda, I., *Convex statistical distances*, Teubner-Texte zur Mathematik, 95, Teubner, Leipzig, 1987. *MR* **89g**:62006

[8] Loève, M., *Probability theory*, Second ed., The University Series in Higher Mathematics, Van Nostrand, Princeton, 1960. *MR* **23** #A670

[9] Rüschendorf, L. and Thomsen, W., Note on the Schrödinger equation and I-projections, *Statist. Probab. Lett.* **17** (1993), 369–375. *MR* **94j**:60126

[10] Rüschendorf, L., Convergence of the iterative proportional fitting procedure, *Ann. Statist.* **23** (1995), 1160–1174. *MR* **97c**:62119

[11] Schrödinger, E., Über die Umkehrung der Naturgesetze, *Sitzungsber. Preuss. Akad. Wiss. Berlin Phys. Math. Kl.* **8/9** (1931), 144–153.

(Received September 11, 1997)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

csiszar@math-inst.hu

INSTITUTE OF SYSTEMS SCIENCE AND
BIOMEDICAL ENGINEERING
LADSEB–CNR
CORSO STATI UNITI, 4
I–35127 PADOVA
ITALY

finesso@ladseb.pd.cnr.it

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 51–69

STANDARDIZED SEQUENTIAL EMPIRICAL PROCESSES

M. CSÖRGŐ and R. NORVAIŠA

To the memory of Alfréd Rényi

Abstract

We study the asymptotic extreme value behaviour of standardized sequential empirical processes that are tied down at the points $(0,0)$ and $(1,1)$. These types of empirical processes arise naturally in the context of change-point analysis. We relate their asymptotic behaviour to the extreme value distribution of a two parameter Ornstein–Uhlenbeck process over a sequence of expanding subsets of the unit square $[0,1]^2$.

1. Introduction and the main result

Let $\{Y_i : i \geqq 1\}$ be a sequence of independent identically distributed real valued random variables with a continuous distribution function F. Then $\{F(Y_i) : i \geqq 1\}$ are independent random variables that are uniformly distributed over the interval $[0,1]$. Consequently, from the point of view of theorem proving for the empirical distribution function on the real line, without loss of generality, we let $\{U_i\colon i \geqq 1\}$ be a sequence of independent uniformly distributed over $[0,1]$ random variables and let F_n be the empirical distribution function based on a random sample $U_1, \dots, U_n$, i.e.,

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[0,x]}(U_i) \quad \text{for } x \in [0,1]. \tag{1.0}$$

The (uniform) empirical process α_n that is defined by

$$\alpha_n(x) := n^{1/2}(F_n(x) - x), \quad x \in [0,1], \tag{1.1}$$

has played a fundamental role in the development of probability theory (cf., e.g., M. Csörgő [5] for a short historical review).

1991 *Mathematics Subject Classification.* Primary 60G70; Secondary 62G30.

Key words and phrases. Standardized sequential empirical processes, extremal processes, multiparameter Ornstein–Uhlenbech processes, multiparameter Gaussian processes, change-point analysis.

Research supported by an NSERC Canada Grant of M. Csörgő and by an NSERC Canada Collaborative Grant at Carleton University, Ottawa, Canada.

0081–6906/98/\$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

The study of suprema of weighted empirical processes was initiated by Rényi [26], who investigated the asymptotic behaviour of statistics like

$$\sup_{a \leqq x \leqq b} \alpha_n(x)/x,\ 0<a<b\leqq 1 \quad \text{and} \quad \sup_{a \leqq x \leqq b} \alpha_n(x)/(1-x),\ 0\leqq a<b<1,$$

as well as that of their two sided versions (cf. also M. Csörgő [4]). For results on replacing the constants a and b in these statistics by constants $0<a(n)\leqq a$ and $b \leqq b(n)<1$ satisfying, as $n\to\infty$,

$$na(n)\to\infty,\ n(1-b(n))\to\infty \ \text{ even when } \ a(n)\to 0,\ b(n)\to 1,$$

we refer to Csáki [3], Mason [23], Section 4.5 of M. Csörgő, S. Csörgő, Horváth and Mason [6], M. Csörgő, Shao and Szyszkowicz [13], M. Csörgő and Horváth [8], and to Chapter 5 of M. Csörgő and Horváth [10].

There is a huge literature that is devoted to the study of the asymptotic behaviour in sup norm of weighted empirical processes

$$\{\alpha_n(x)/q(x):0<x<1\}$$

with positive weight functions q on $(0,1)$ that are to satisfy certain integrability conditions in addition to necessarily being so that

$$q(x)/x^{1/2}\to\infty \ \text{ as } \ x\downarrow 0, \ \text{ and } \ q(x)/(1-x)^{1/2}\to\infty \ \text{ as } x\uparrow 1. \tag{1.2}$$

More precisely, in terms of the following class of functions

$$\begin{aligned} Q_{0,1}:=\Big\{q:\ &\inf_{\delta\leqq x\leqq 1-\delta} q(x)>0 \text{ for all } 0<\delta<1/2,\\ &q \text{ is non-decreasing in a neighbourhood of zero}\\ &\text{and non-increasing in a neighbourhood of one.}\Big\}\end{aligned} \tag{1.3}$$

and the integral

$$I(q,c):=\int_0^1 \frac{1}{t(1-t)}\exp(-cq^2(t)/(t(1-t))dt, \quad c>0, \tag{1.4}$$

we have (cf. Chibisov [2], O'Reilly [24], and M. Csörgő, S. Csörgő, Horváth and Mason [6])

THEOREM A. *If $q\in Q_{0,1}$, then the following two statements are equivalent:*

(i) *There is a sequence of Brownian bridges $\{B_n(x):0\leqq x\leqq 1\}$ such that, as $n\to\infty$,*

$$\sup_{0<x<1} |\alpha_n(x)-B_n(x)|/q(x)=o_P(1),$$

(ii) *$I(q,c)<\infty$ for all $c>0$.*

We have also (cf. M. Csörgő, S. Csörgő, Horváth and Mason [6])

THEOREM B. *If $q \in Q_{0,1}$, then the following two statements are equivalent:*

(i) $\sup_{0<x<1} |\alpha_n(x)|q(x) \xrightarrow{\mathcal{D}} \sup_{0<x<1} |B(x)|/q(x)$,

as $n \to \infty$, where $\{B(x): 0 < x < 1\}$ is a Brownian bridge,

(ii) $I(q,c) < \infty$ *for some* $c > 0$.

REMARK 1.1. We note that if $q \in Q_{0,1}$ is a *càdlàg* function then, by Theorem A, as $n \to \infty$, we have

$$\alpha_n(\cdot)/q(\cdot) \xrightarrow{\mathcal{D}} B(\cdot)/q(\cdot) \quad \text{in } D[0,1]$$

if and only if $I(q,c) < \infty$ for all $c > 0$. The latter, characterization of weighted weak convergence does not, however, imply the convergence in distribution statement (i) of Theorem B. We note also in passing that, omitting the absolute value signs, the corresponding version of Theorem B holds true.

For further results and their applications along these lines we refer to M. Csörgő, S. Csörgő, Horváth and Mason [6], Shorack and Wellner [27], and M. Csörgő and Horváth [8], [9]. For relating these types of results to Rényi [26] and Csáki [3] type statistics, we refer to Section 4.5 of M. Csörgő, S. Csörgő, Horváth and Mason [6] and to Section 4 of M. Csörgő, Shao and Szyszkowicz [13].

The results of Theorems A and B do not, of course, hold true with $q(x) := (x(1-x))^{1/2}$ (cf., e.g., Corollaries 2.2 and 3.2 in Chapter 4 of Csörgő and Horváth [9]).

Let, for each $x > e^e$,

$$a(x) := (2 \log x)^{1/2}, \; b(x) := 2 \log x + 2^{-1} \log\log x - 2^{-1} \log \pi,$$

$$a_n := a(\log n) \quad \text{and} \quad b_n := b(\log n).$$

We have (cf. Eicker [19], Jaeschke [20])

THEOREM C. *For any $-\infty < t < \infty$*

$$\lim_{n\to\infty} P\Big\{a_n \sup_{0<x<1} \alpha_n(x)/(x(1-x))^{1/2} - b_n \leqq t\Big\} = \exp(-e^{-t}), \tag{1.5}$$

$$\lim_{n\to\infty} P\Big\{a_n \sup_{0<x<1} |\alpha_n(x)|/(x(1-x))^{1/2} - b_n \leqq t\Big\} = \exp(-2e^{-t}). \tag{1.6}$$

For a discussion of these results and that of their relationship to the one-time parameter Ornstein–Uhlenbeck process and the Darling–Erdős [16] theorem, we refer to M. Csörgő and Révész [12], to Section 4.4 of M. Csörgő, S. Csörgő, Horváth and Mason [6], and to Section 5.1 of M. Csörgő and Horváth [9] that also contains many more further results along these lines.

In order to introduce the problem of standardized sequential empirical processes that we are to deal with in this exposition, and for the sake of summarizing some recent results, we first assume that we have multivariate observations taking values in $\mathbf{R}^d$, $d \geqq 1$. Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ be independent random vectors in $\mathbf{R}^d$ with right continuously defined distribution functions $F_{(1)}(\mathbf{x}), \ldots, F_{(n)}(\mathbf{x})$. Suppose we wish to test the 'no-change in distribution' null hypothesis

$$H_0: \qquad F_{(1)}(\mathbf{x}) = \cdots = F_{(n)}(\mathbf{x}) \quad \text{for all } \ \mathbf{x} \in \mathbf{R}^d$$

against the 'one change in distribution' alternative

$$\begin{aligned} H_A: \quad & \text{there is an integer } k^*, \ 1 \leqq k^* < n, \ \text{such that} \\ & F_{(1)}(\mathbf{x}) = \cdots = F_{(k^*)}(\mathbf{x}), \ F_{(k^*+1)}(\mathbf{x}) = \cdots = F_{(n)}(\mathbf{x}) \ \text{for all } \ \mathbf{x} \in \mathbf{R}^d \\ & \text{and } F_{(k^*)}(\mathbf{x}_0) \neq F_{(k^*+1)}(\mathbf{x}_0) \ \text{ for some } \mathbf{x}_0 \in \mathbf{R}^d. \end{aligned}$$

Since k^* of H_A is usually unknown, in order to test H_0 versus H_A, it appears reasonable to consider (cf., e.g., Darkhovsky [15], Picard [25], Deshayes and Picard [17], [18], Szyszkowicz [28], [29], [30], M. Csörgő and Szyszkowicz [14], M. Csörgő, Horváth and Szyszkowicz [11], and M. Csörgő and Horváth [10]) the sequence of statistics

$$(1.7) \qquad \begin{aligned} & n^{1/2} \max_{1 \leqq k < n} \sup_{\mathbf{x}} \left| \frac{1}{k} \sum_{i=1}^{k} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) - \frac{1}{n-k} \sum_{i=k+1}^{n} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) \right| \\ & = \max_{1 \leqq k < n} \sup_{\mathbf{x}} \frac{\left| \sum_{1 \leqq i \leqq k} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) - \frac{k}{n} \sum_{1 \leqq i \leqq n} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) \right|}{n^{1/2}\left(\frac{k}{n}\left(1 - \frac{k}{n}\right)\right)}, \end{aligned}$$

where $\leqq$ in $\mathbf{R}^d$ is meant to compare vectors componentwise, and to reject H_0 in favour of H_A, if this sequence of statistics were to produce 'far too many large values', as $n \to \infty$. This nonparametric heuristic reasoning can, for example, be argued via the left-hand side of (1.7) that compares the empirical distribution functions of the first k observations to those of the remaining $n-k$ observations uniformly in $\mathbf{x} \in \mathbf{R}^d$, and uniformly in k as well, over the possible values $1 \leqq k \leqq n-1$ of the unknown random variable k^*. Via the right-hand side of (1.7) it can, however, be easily seen that this sequence of statistics $\xrightarrow{\mathcal{D}} \infty$, as $n \to \infty$, even if H_0 were to be true. Consequently, in order to secure a nondegenerate limiting behaviour under H_0 for statistics that are based on comparing empirical distributions *à la* (1.7), we must introduce some appropriate renormalization.

We note in passing that a rationalization of introducing the sequence of statistics in (1.7) could be simply based also on trying to answer the following simple question. Given a sample of independent random vectors $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ in $\mathbf{R}^d$, is it reasonable to assume that they constitute a random sample? That is to say, for the sake of introducing the sequence of statistics in (1.7) it is not necessary to use the language of change-point analysis. Instead, we could have just simply asked whether a given set of independent data could possibly be viewed as being homogeneous in distribution.

Back to (1.7), on multiplying the respective arguments of the $\max\limits_{1 \le k < n} \sup\limits_{\mathbf{x}}$ functional operations on both sides by $(\frac{k}{n}(1-\frac{k}{n}))^{1/2}$, we obtain the sequence of statistics

(1.8)
$$V_n := \max_{1 \le k < n} n^{1/2}\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2} \sup_{\mathbf{x}} \left| \frac{1}{k}\sum_{i=1}^{k} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) - \frac{1}{n-k}\sum_{i=k+1}^{n} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)\right|$$
$$= \max_{1 \le k < n} n^{-1/2}\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} \sup_{\mathbf{x}} \left| \sum_{i=1}^{k} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i) - \frac{k}{n}\sum_{i=k+1}^{n} \mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)\right|.$$

Again, even if H_0 were true, as $n \to \infty$, this new sequence of statistics $V_n \xrightarrow{\mathcal{D}} \infty$, though we are now somewhat nearer to saying something more sensible than this. Namely, if for any fixed $\mathbf{x}_0 \in \mathbf{R}^d$, we let

$$V_n(\mathbf{x}_0)$$
$$:= \max_{1 \le k < n} n^{1/2}\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2} \left| \frac{1}{k}\sum_{i=1}^{k} \mathbf{1}_{(-\infty,\mathbf{x}_0]}(\mathbf{Y}_i) - \frac{1}{n-k}\sum_{i=k+1}^{n} \mathbf{1}_{(-\infty,\mathbf{x}_0]}(\mathbf{Y}_i)\right|$$
$$= \max_{1 \le k < n} n^{-1/2}\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{-1/2} \left| \sum_{i=1}^{k} \mathbf{1}_{(-\infty,\mathbf{x}_0]}(\mathbf{Y}_i) - \frac{k}{n}\sum_{i=k+1}^{n} \mathbf{1}_{(-\infty,\mathbf{x}_0]}(\mathbf{Y}_i)\right|,$$

then we have (cf. Corollary 2.2 of M. Csörgő and Horváth [7], or Theorem A.4.2 in M. Csörgő and Horváth [9])

(1.9)
$$\lim_{n\to\infty} P\{a_n V_n(\mathbf{x}_0) - b_n \le t\} = \exp(-2e^{-t})$$

for any $-\infty < t < \infty$, where a_n and b_n are as in Theorem C. The key to the proofs of (1.5), (1.6) and (1.9), and to the asymptotic equivalence of the respective left-hand sides of (1.6) and (1.9) is their relationship to the asymptotic behaviour of the one-time parameter Ornstein-Uhlenbeck process *à la* Darling and Erdős [16] (cf. the references quoted right after Theorem C). However, *the problem of the asymptotic extreme value behaviour of V_n of* (1.8) *with appropriate sequences of norming constants, say $\tilde{a}_n$ and $\tilde{b}_n$, that is to say the problem of finding $\tilde{a}_n$ and $\tilde{b}_n$ such that, as $n \to \infty$,*

(1.10)
$$\tilde{a}_n V_n - \tilde{b}_n \xrightarrow{\mathcal{D}} \text{a nondegenerate random variable,}$$

remains open. One of the aims of this exposition is to throw some light on the nature of the difficulties one encounters when trying to deal with this interesting problem of extreme value asymptotics that would be of interest to resolve from the point of view of change-point analysis.

Before continuing with studying the asymptotic extreme value behaviour of V_n via seeking appropriate norming constants $\bar{a}_n$ and $\bar{b}_n$, we summarize some recent results for these statistics that deal with renormalizing them so that they should have limits based on appropriate Gaussian processes. Let (cf. (1.7))

$$w_n(\mathbf{x},t):=\begin{cases} 0, & 0\leqq t<\frac{1}{n},\\ \dfrac{\frac{1}{[nt]}\sum_{i=1}^{[nt]}\mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)-\frac{1}{n-[nt]}\sum_{i=[nt]+1}^{n}\mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)}{n^{-1/2}\left(\frac{[nt]}{n}\left(1-\frac{[nt]}{n}\right)\right)^{-1}}, & \frac{1}{n}\leqq t\leqq\frac{n-1}{n},\\ 0, & \frac{n-1}{n}<t\leqq 1,\end{cases}$$

(1.11)

$$=\begin{cases} 0, & 0\leqq t<\frac{1}{n},\\ n^{-1/2}\left(\sum_{i=1}^{[nt]}\mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)-\frac{[nt]}{n}\sum_{i=1}^{n}\mathbf{1}_{(-\infty,\mathbf{x}]}(\mathbf{Y}_i)\right), & \frac{1}{n}\leqq t\leqq\frac{n-1}{n},\\ 0, & \frac{n-1}{n}<t\leqq 1,\end{cases}$$

where, as usual, $[a]$ is the integer part of a. On denoting the common distribution function under H_0 by F, it can be easily seen via weak convergence of multivariate empirical processes that $w_n(\mathbf{x},t)$ converges weakly to $\{\Gamma_F(\mathbf{x},t):\mathbf{x}\in\mathbf{R}^d,\ 0\leqq t\leqq 1\}$, a Gaussian process with

$$E\Gamma_F(\mathbf{x},t)=0,\quad E\Gamma_F(\mathbf{x},t)\Gamma_F(\mathbf{y},s)=\{F(\mathbf{x}\wedge\mathbf{y})-F(\mathbf{x})F(\mathbf{y})\}(t\wedge s-ts),$$

where $\mathbf{x}\wedge\mathbf{y}=(\min(x_1,y_1),\dots,\min(x_d,y_d))$, $\mathbf{x}=(x_1,\dots,x_d)$, $\mathbf{y}=(y_1,\dots,y)\in\mathbf{R}^d$. We have also (cf. M. Csörgő, Horváth and Szyszkowicz [11])

THEOREM D. *Assume that H_0 holds true and that $q\in Q_{0,1}$. Then*

(i) *there is a sequence of Gaussian processes $\{\Gamma_n(\mathbf{x},t):\mathbf{x}\in\mathbf{R}^d,\ 0\leqq t\leqq 1\}$ such that for each $n\geqq 1$*

$$\{\Gamma_n(\mathbf{x},t):\mathbf{x}\in\mathbf{R}^d,0\leqq t\leqq 1\}\overset{\mathcal{D}}{=}\{\Gamma_F(\mathbf{x},t):\mathbf{x}\in\mathbf{R}^d,0\leqq t\leqq 1\}$$

and, as $n\to\infty$,

$$\sup_{0<t<1}\sup_{\mathbf{x}}|w_n(\mathbf{x},t)-\Gamma_n(\mathbf{x},t)|=o_P(1)$$

if and only if $I(q,c)<\infty$ *for all* $c>0$, *where* $I(q,c)$ *is defined in* (1.4). *Moreover,*

(ii) $I(q,c)<\infty$ *for some* $c>0$ *if and only if, as* $n\to\infty$,

$$\max_{1\leqq k<n}\sup_{\mathbf{x}} w_n(\mathbf{x},k/n)/q(k/n)\xrightarrow{\mathcal{D}}\sup_{0<t<1}\sup_{\mathbf{x}}|\Gamma_F(\mathbf{x},t)|/q(t).$$

We note in passing that if we were to fix the value of $\mathbf{x}\in\mathbf{R}^d$, then the statements (i) and (ii) of Theorem D, formulated accordingly without $\sup_{\mathbf{x}}$ would follow immediately from Corollary 2.1 of M. Csörgő and Horváth [7]. It is that we can take $\sup_{\mathbf{x}}$ as well and still have (i) and (ii) of Theorem D in their respective present forms is the essence of this theorem.

Now for the sake of focusing in on the problem of studying the asymptotic extreme value behaviour of the sequence of statistics V_n of (1.8), and for the sake of simplifying notation, presentation and calculations, *we go back to assuming, and from now on throughout this exposition, that* $\{Y_i:i\geqq 1\}$ *is a sequence of independent identically distributed real valued random variables with a continuous distribution function* F. Hence, without loss of generality, from now on we base our investigations on the uniform empirical distribution function F_n as it is defined in (1.0).

For each $n\geqq 1$, we define a two-parameter stochastic process $\mathbb{D}_n=$ $=\{\mathbb{D}_n(x,t)\colon (x,t)\in[0,1]^2\}$ by

$$\mathbb{D}_n(x,t):=\begin{cases}[nt](F_{[nt]}(x)-F_n(x)), & \text{for } 0\leqq t\leqq 1/2,\\ ([nt]+1)(F_{[nt]+1}(x)-F_n(x)), & \text{for } 1/2<t<1,\\ 0, & \text{for } t=1.\end{cases}\tag{1.12}$$

Then, for each random sample $U_1,\dots,U_n$ such that $0<U_{1:n}\leqq\dots\leqq U_{n:n}<1$, the order statistics, we have

$$\sup_{(x,t)\in(0,1)^2}\frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}}=\sup_{1/n\leqq t\leqq 1-1/n}\ \sup_{U_{n:1}\leqq x\leqq U_{n:n}}\frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}}$$

(1.13)

$$=\max_{1\leqq k\leqq n-1} n^{-1/2}\left(\frac{k}{n}(1-\frac{k}{n})\right)^{-1/2}\sup_{0<x<1}\frac{\left|\sum_{i=1}^k \mathbf{1}_{[0,x]}(U_i)-\frac{k}{n}\sum_{i=1}^n \mathbf{1}_{[0,x]}(U_i)\right|}{\sqrt{x(1-x)}}$$

$$=\max_{1\leqq k\leqq n-1}\sqrt{n}\sqrt{\frac{k}{n}(1-\frac{k}{n})}\sup_{0<x<1}\frac{\left|\frac{1}{k}\sum_{i=1}^k \mathbf{1}_{[0,x]}(U_i)-\frac{1}{n-k}\sum_{i=k+1}^n \mathbf{1}_{[0,x]}(U_i)\right|}{\sqrt{x(1-x)}}.$$

The right-hand side is a weighted suprema of the difference between the empirical distribution functions based on $U_1,\dots,U_k$ and $U_{k+1},\dots,U_n$, respectively. Comparing (1.8) and (1.13) in this context, we call attention to the fact that in the latter we use complete standardization of the underlying two parameter stochastic process. This, in turn, will enable us to relate the *thus modified* original problem of (1.10) to the extreme value distribution of a two parameter Ornstein–Uhlenbeck process over a sequence of expanding subsets of $[0,1]^2$ (cf. (3.1) and (3.2)).

For a set $\mathbf{A}\subset[0,1]^2$, an integer $n\geqq 1$ and the stochastic process $\mathbb{D}_n$ defined by (1.12), let

$$D_n(A):=\sup_{(x,t)\in\mathbf{A}}\frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}}. \tag{1.14}$$

Here we find the asymptotic behaviour of $D_n(\mathbf{A}_n)$ as $n\to\infty$ for a certain expanding sequence $\{\mathbf{A}_n\colon n\geqq 1\}$ of subsets of $[0,1]^2$. More specifically, for each $T>e^e$, let

$$a(T):=1/\sqrt{4\ln T}\quad\text{and}$$
$$b(T):=\sqrt{4\ln T}+[\frac{3}{2}\ln(4\ln T)-\frac{1}{2}\ln(2\pi)]/\sqrt{4\ln T}.$$

The numbers $a(T)$ and $b(T)$ are special cases of the numbers $a_d(T)$ and $b_d(T)$ with $d=2$, respectively, defined in Section 2. For each $r\geqq 0$, let

$$Q(r)=\{(x,t)\in(0,1)^2\colon |\ln\frac{x}{1-x}|+|\ln\frac{t}{1-t}|\leqq r\}\quad\text{for } r\geqq 0.$$

The following statement is the main result of this paper.

THEOREM 1.1. *Let $\{D_n\colon n\geqq 1\}$ be a sequence defined by* (1.14). *Then, for each real number z,*

$$\lim_{n\to\infty}\Pr\Big(\Big\{D_n\Big(Q\Big(\ln\frac{n}{2\ln\ln n}\Big)\Big)<a(\ln n)z+b(\ln n)\Big\}\Big)=\exp\{-e^{-z}\}.$$

The proof follows from Lemmas 3.1 and 3.2 below.

2. Multiparameter Gaussian processes

We start with the formulation of an extremal type theorem for a multiparameter analogue of the Ornstein–Uhlenbeck process. Let λ be a positive real number, d be a positive integer and let $\mathbf{D}^d$ be any subset of $\mathbf{R}^d$. A real-valued Gaussian process $\mathbb{Z}=\{\mathbb{Z}(\mathbf{u})\colon \mathbf{u}\in\mathbf{D}^d\}$ with mean zero and covariance function

$$E\mathbb{Z}(\mathbf{u})\mathbb{Z}(\mathbf{v})=\exp\{-\lambda\sum_{i=1}^{d}|u_i-v_i|\},\quad\text{for }\mathbf{u}=(u_i),\ \mathbf{v}=(v_i)\in\mathbf{D}^d,$$

is called a *d-parameter Ornstein–Uhlenbeck process with coefficient* λ. For each $T>e^e$, let $a_d(T):=1/\sqrt{2d\ln T}$ and let

$$b_d(T):=\sqrt{2d\ln T}+[(d-1/2)\ln(2d\ln T)-(1/2)\ln(2\pi)]/\sqrt{2d\ln T}.$$

By Corollaries 1, 2 and 3 of Bickel and Rosenblatt [1], we have:

THEOREM 2.1. *Let $\mathbb{Z}$ be a d-parameter Ornstein–Uhlenbeck process with coefficient λ and let* $\mathbf{B}$ *be a bounded, Jordan measurable subset of* $\mathbf{R}^d$ *with* $\operatorname{vol}(\mathbf{B})=1$. *Then, for each real number z,*

$$\lim_{T\to\infty} \Pr(\{\sup_{\mathbf{u}/T\in\mathbf{B}} \mathbb{Z}(\mathbf{u}) < a_d(T)z + b_d(T)\}) = \exp\{-\lambda^d e^{-z}\}$$

and

$$\lim_{T\to\infty} \Pr(\{\sup_{\mathbf{u}/T\in\mathbf{B}} |\mathbb{Z}(\mathbf{u})| < a_d(T)z + b_d(T)\}) = \exp\{-2\lambda^d e^{-z}\}.$$

We apply this result to a certain class of multiparameter Gaussian processes to be defined next. Let $\mathbf{X}^d := X_1 \times \cdots \times X_d$, where each X_i is a right-open or right-closed interval of non-negative real numbers. Let h be a non-negative real-valued function on $\mathbf{X}^d$, and let k_i, $i=1,\ldots,d$, be real-valued functions on X_i such that

$$f_i(x_i) := \int_0^{x_i} k_i^2(t)\,dt < \infty \quad \text{for all } x_i \in X_i. \tag{2.1}$$

Define a mean-zero Gaussian process $\mathbb{G}=\{\mathbb{G}(\mathbf{x})\colon \mathbf{x}\in\mathbf{X}^d\}$ by

$$\mathbb{G}(\mathbf{x}) := h(\mathbf{x}) \int_0^{x_1}\ldots\int_0^{x_d} k_1(t_1)\cdots k_d(t_d)\,d\mathbb{W}(t_1,\ldots,t_d) \tag{2.2}$$

for $\mathbf{x}=(x_i)\in\mathbf{X}^d$, where $\mathbb{W}$ is a Brownian sheet (cf. Example 2.3 below for definition) and the integral is stochastic. Using the functions in (2.1), define the transformation

$$L_f\colon \mathbf{X}^d \mapsto \mathbf{D}^d := \bigotimes_{i=1}^d \{\ln f_i(x)\colon x\in X_i\}$$

by $L_f(\mathbf{x}) := (\ln f_i(x_i)) \in \mathbf{R}^d$ for each $\mathbf{x}=(x_i)\in\mathbf{X}^d$. Since each f_i possess a well-defined inverse function f_i^{-1}, we can define the inverse transformation L_f^{-1} by

$$L_f^{-1}(\mathbf{u}) := (f_i^{-1}(e^{u_i})) \in \mathbf{X}^d \quad \text{for } \mathbf{u}=(u_i)\in\mathbf{D}^d.$$

The next statement relates Gaussian processes $\mathbb{G}$ to the Ornstein–Uhlenbeck process $\mathbb{Z}$.

LEMMA 2.2. *Let $\mathbb{G}=\{\mathbb{G}(\mathbf{x})\colon \mathbf{x}\in\mathbf{X}^d\}$ be a mean-zero Gaussian process given by* (2.2). *Then the stochastic process $\mathbb{Z}=\{\mathbb{Z}(\mathbf{u})\colon \mathbf{u}\in\mathbf{D}^d\}$ defined by*

$$\begin{aligned}\mathbb{Z}(\mathbf{u}) := {}&\exp\Big\{-\frac{1}{2}\sum_{i=1}^d u_i\Big\}\times\\ &\times \int_0^{f_1^{-1}(e^{u_1})}\cdots\int_0^{f_d^{-1}(e^{u_d})} k_1(t_1)\cdots k_d(t_d)\,d\mathbb{W}(t_1,\ldots,t_d)\end{aligned} \tag{2.3}$$

is a d-parameter Ornstein–Uhlenbeck process with coefficient $\lambda = 1/2$, *and*

$$\mathbb{G}(\mathbf{x}) = \sqrt{E\mathbb{G}^2(\mathbf{x})}\,\mathbb{Z}(L_f(\mathbf{x})), \quad \textit{for all } \mathbf{x} \in \mathbf{X}^d.$$

PROOF. Due to assumption (2.1), the right-hand side of (2.3) is a well defined stochastic integral. Therefore $\mathbb{Z}$ is a Gaussian process. Now the claim follows simply by computing the covariance function of $\mathbb{Z}$. □

Here is a list of a few examples of Gaussian processes which possess integral representation (2.2).

EXAMPLE 2.3. A Brownian sheet $\mathbb{W} = \{\mathbb{W}(\mathbf{t})\colon \mathbf{t} \in [0,\infty)^d\}$ is a centered Gaussian process with covariance

$$E\mathbb{W}(\mathbf{s})\mathbb{W}(\mathbf{t}) = \prod_{i=1}^{d} s_i \wedge t_i \quad \text{for } \mathbf{s} = (s_i),\ \mathbf{t} = (t_i) \in [0,\infty)^d.$$

Taking $k_1 = \cdots = k_d \equiv 1$ and $h \equiv 1$ in (2.2) we conclude the same covariance for $\mathbb{G}$ and that (2.1) holds with $f_1(t) = \cdots = f_d(t) = t$ for $t \in [0,\infty)$.

EXAMPLE 2.4. A centered Gaussian process $\mathbb{H} = \{\mathbb{H}(\mathbf{x})\colon \mathbf{x} \in [0,1]^d\}$ with covariance

$$E\mathbb{H}(\mathbf{x})\mathbb{H}(\mathbf{y}) = \prod_{i=1}^{d}(x_i \wedge y_i - x_i y_i) \quad \text{for } \mathbf{x} = (x_i),\ \mathbf{y} = (y_i) \in [0,1]^d.$$

In representation (2.2) taking

$$h(\mathbf{x}) = \prod_{i=1}^{d}(1 - x_i) \quad \text{for } \mathbf{x} = (x_i) \in [0,1)^d \tag{2.4}$$

and, for each $i = 1,\ldots,d$,

$$k_i(t) = 1/(1-t) \quad \text{for } t \in [0,1), \tag{2.5}$$

we obtain the same covariance for $\mathbb{G}$. In this case (2.1) holds for $\mathbf{x} \in [0,1)^d$ with $f_1(x) = \cdots = f_d(x) = x/(1-x)$ for $x \in (0,1)$ and $L_f^{-1}(\mathbf{u}) = (e^{u_i}/(1+e^{u_i}))$ for $\mathbf{u} = (u_i) \in \mathbf{R}^d$.

We note that, unless $d = 1$, the Gaussian process $\mathbb{H}$ does not coincide with the Brownian bridge process $\mathbb{B}(\mathbf{x}) := \mathbb{W}(\mathbf{x}) - \mathrm{vol}([0,\mathbf{x}])\mathbb{W}(\mathbf{1})$ for $\mathbf{x} \in [0,1]^d$, also called a pinned Wiener sheet, which has covariance function

$$E\mathbb{B}(\mathbf{x})\mathbb{B}(\mathbf{y}) = \prod_{i=1}^{d} x_i \wedge y_i - \prod_{i=1}^{d} x_i y_i \quad \text{for } \mathbf{x} = (x_i),\ \mathbf{y} = (y_i) \in [0,1]^d.$$

EXAMPLE 2.5. A centered Gaussian process

$$\mathbb{K}_d = \{\mathbb{K}_d(\mathbf{x}, t) \colon (\mathbf{x}, t) \in [0,1]^{d-1} \times [0, \infty)\}$$

with covariance

$$E\mathbb{K}_d(\mathbf{x}, s)\mathbb{K}_d(\mathbf{y}, t) = (s \wedge t) \prod_{i=1}^{d-1} (x_i \wedge y_i - x_i y_i)$$

for $\mathbf{x} = (x_i), \mathbf{y} = (y_i) \in [0,1]^{d-1}$ and $s, t \in [0, \infty)$. In the representation (2.2) taking h as in (2.4) with $d-1$ factors instead of d, k_i, $i = 1, \dots, d-1$, as in (2.5) and $k_d(t) = 1$ for all $t \in [0, \infty)$, we get the same covariance for $\mathbb{G}$. Here (2.1) holds with $f_1(x) = \cdots = f_{d-1}(x) = x/(1-x)$ for $x \in [0,1)$ and $f_d(t) = t$ for $t \in [0, \infty)$.

Using integral representation (2.2) we apply Theorem 2.1 to Gaussian process $\mathbb{H}$ of Example 2.4. For a set $\mathbf{A} \subset [0,1]^d$, let

$$H(\mathbf{A}) := \sup_{\mathbf{x} \in \mathbf{A}} \frac{\mathbb{H}(\mathbf{x})}{\sqrt{E\mathbb{H}^2(\mathbf{x})}} \quad \text{and} \quad |H|(\mathbf{A}) := \sup_{\mathbf{x} \in \mathbf{A}} \frac{|\mathbb{H}(\mathbf{x})|}{\sqrt{E\mathbb{H}^2(\mathbf{x})}}.$$

Let $\mathbf{B}_d(r) := \{\mathbf{x} \in \mathbf{R}^d \colon \sum_{i=1}^d |x_i| \leqq r\}$ be a ball with radius r in the Banach space ℓ_1^d. Then

$$\mathrm{vol}(\mathbf{B}_d(r)) = r^d \mathrm{vol}\Big\{\mathbf{x} \in \mathbf{R}^d \colon \sum_{i=1}^d |x_i| \leqq 1\Big\} = (2r)^d/d!.$$

Put $c_d := (d!)^{1/d}/2$ so that $\mathrm{vol}(\mathbf{B}_d(c_d)) = 1$ and

$$\mathbf{Q}_d(r) := L_f^{-1}(\mathbf{B}_d(r)) := \{\mathbf{x} \in \mathbf{X}^d \colon L_f(\mathbf{x}) \in \mathbf{B}_d(r)\}, \tag{2.6}$$

where $L_f(\mathbf{x}) = (\ln(x_i/(1-x_i))) \in \mathbf{R}^d$. The next statement follows from Theorem 2.1 and Lemma 2.2.

COROLLARY 2.6. *Let $\mathbb{H}$ be a Gaussian process defined in Example 2.4. Then, for all $z \in \mathbf{R}^1$,*

$$\lim_{T \to \infty} \Pr(\{H(Q_d(c_d T)) < a_d(T)z + b_d(T)\}) = \exp\{-2^{-d} e^{-z}\} \tag{2.7}$$

and

$$\lim_{T \to \infty} \Pr(\{|H|(Q_d(c_d T)) < a_d(T)z + b_d(T)\}) = \exp\{-2^{1-d} e^{-z}\}. \tag{2.8}$$

To elucidate the last statement we restate it for $d=1$. In this case we identify Gaussian process $\mathbb{H}$ with a Brownian bridge process $\mathbb{B}$. For each positive real number r, let

$$B(r) := \sup_{r \leqq x \leqq 1-r} \frac{\mathbb{B}(x)}{\sqrt{x(1-x)}}.$$

COROLLARY 2.7. *Let $\mathbb{B}$ be a Brownian bridge process. Then, for all $z \in \mathbf{R}^1$, any $c \in \mathbf{R}^1$ and $\beta \in (0,1)$, we have*

$$\lim_{n\to\infty} \Pr(\{B((\ln n)^c/n) < a_1(\ln n)z + b_1(\ln n)\}) = \exp\{-2e^{-z}\} \tag{2.9}$$

and

$$\lim_{n\to\infty} \Pr(\{B(n^{-\beta}) < a_1(\ln n)z + b_1(\ln n)\}) = \exp\{-2\beta^2 e^{-z}\}. \tag{2.10}$$

Likewise, (2.9) *and* (2.10) *hold with squared limits if $\mathbb{B}(x)$ is replaced by $|\mathbb{B}(x)|$ in the definition of $B(\cdot)$.*

PROOF follows from Corollary 2.6 and Khintchine's convergence theorem. Indeed, since $c_1=1/2$, relation $x \in Q_1(c_1 T)$ means that $|\ln(x/(1-x))| \leqq T/2$. We note that, for any $0 < a \leqq 1/2$, $a \leqq x \leqq 1-a$ if and only if $|\ln(x/(1-x))| \leqq \ln((1-a)/a)$. Let $a := a_n := (\ln n)^c/n$ and let $T := T_n := 2\ln(n/(\ln n)^c - 1)$. Therefore $x \in Q_1(c_1 T_n)$ if and only if $a_n \leqq x \leqq 1 - a_n$. Moreover, as $n \to \infty$, $T_n \sim 2\ln n$, $a_1(2\ln n) \sim a_1(\ln n)$ and $[b_1(\ln n) - b_1(2\ln n)]/a_1(2\ln n) \sim -2\ln 2$. Therefore, by Khintchine's convergence theorem (cf. Theorem 1.2.3 in Leadbetter, Lindgren and Rootzén [22]), (2.9) follows from (2.7). Similar calculations show that (2.10) and corresponding relations for the absolute value of $\mathbb{B}$ follow from (2.7) and (2.8). □

Corollary 2.7 relates to a classical result of Darling and Erdős [16] (cf. also Section 1.9 of Csörgő and Révész [12], and Chapter 5 of Csörgő and Horváth [9]).

3. Sequential empirical processes

Let $\{U_i : i \geqq 1\}$ be a sequence of independent uniformly distributed over $[0,1]$ random variables and let F_n be the empirical distribution function based on a sample $U_1, \ldots, U_n$, i.e.,

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[0,x]}(U_i) \quad \text{for } x \in [0,1].$$

For each integer $n \geqq 1$, define as in (1.12)

$$\mathbb{D}_n(x,t) := \begin{cases} [nt](F_{[nt]}(x) - F_n(x)), & \text{for } 0 \leqq t \leqq 1/2, \\ ([nt]+1)(F_{[nt]+1}(x) - F_n(x)), & \text{for } 1/2 < t < 1, \\ 0, & \text{for } t = 1. \end{cases}$$

Then, for each sample $U_1, \dots, U_n$ such that $0 < U_{1:n} \leqq \dots \leqq U_{n:n} < 1$ we have

$$\sup_{(x,t)\in[0,1]^2} \frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}} = \sup_{1/n \leqq t \leqq 1-1/n} \sup_{U_{n:1} \leqq x \leqq U_{n:n}} \frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}}$$

$$= \max_{1 \leqq k \leqq n-1} n^{-1/2} \left(\frac{k}{n}(1-\frac{k}{n})\right)^{-1/2} \sup_{0<x<1} \frac{|\sum_{i=1}^{k} \mathbf{1}_{[0,x]}(U_i) - \frac{k}{n}\sum_{i=1}^{n} \mathbf{1}_{[0,x]}(U_i)|}{\sqrt{x(1-x)}}$$

$$= \max_{1 \leqq k \leqq n-1} \sqrt{n}\sqrt{\frac{k}{n}(1-\frac{k}{n})} \sup_{0<x<1} \frac{|\frac{1}{k}\sum_{i=1}^{k} \mathbf{1}_{[0,x]}(U_i) - \frac{1}{n-k}\sum_{i=k+1}^{n} \mathbf{1}_{[0,x]}(U_i)|}{\sqrt{x(1-x)}}.$$

The right-hand side is a weighted suprema of the difference between the empirical processes based on $U_1, \dots, U_k$ and $U_{k+1}, \dots, U_n$, respectively.

Recalling notation (2.6) for the case $d = 2$, i.e.,

$$Q_2(r) = \{(x,t) \in (0,1)^2 \colon |\ln \frac{x}{1-x}| + |\ln \frac{t}{1-t}| \leqq r\} \quad \text{for } r \geqq 0,$$

we have:

LEMMA 3.1. *For each* $z \in \mathbf{R}^1$,

$$\lim_{n\to\infty} \Pr\left(\left\{\sup_{(x,t)\in Q_2(\mu_n)} \frac{|\mathbb{D}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}} < a_2(\ln n)z + b_2(\ln n)\right\}\right) = \exp\{-e^{-z}\},$$

where $\mu_n = \ln(n/(\ln n)^5)$.

PROOF. By (2.8) with $d = 2$, using Khintchine's convergence theorem as in the proof of Corollary 2.7, we get that

$$\text{(3.1)} \quad \lim_{n\to\infty} \Pr\left(\left\{\sup_{(x,t)\in Q_2(\mu_n)} \frac{|\mathbb{H}(x,t)|}{\sqrt{t(1-t)x(1-x)}} < a_2(\ln n)z + b_2(\ln n)\right\}\right) = \exp\{-e^{-z}\},$$

where $\mathbb{H}$ is a Gaussian process defined in Example 2.4 with $d = 2$. Therefore, it suffices to construct a sequence $\{U_i \colon i \geqq 1\}$ of independent uniform random variables and a sequence $\{\mathbb{H}_n \colon n \geqq 1\}$ of mean-zero Gaussian processes, each with the covariance function

$$E\mathbb{H}_n(x,t)\mathbb{H}_n(y,s) = (x \wedge y - xy)(s \wedge t - st) \quad \text{for } x, y, t, s \in [0,1]$$

such that

$$(3.2)\quad \lim_{n\to\infty} \sqrt{\ln_2 n} \sup_{(x,t)\in Q_2(\mu_n)} \frac{|\mathbb{D}_n(x,t) - \sqrt{n}\mathbb{H}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}} = 0 \quad \text{in probability.}$$

By Theorem 4 of Komlós, Major and Tusnády [21], there exist a sequence $\{U_i : i \geqq 1\}$ of independent uniform random variables and a sequence $\{\mathbb{B}_i : i \geqq 1\}$ of independent Brownian bridge processes such that

$$(3.3)\quad \overline{\lim_{n\to\infty}} (\ln n)^{-2} \max_{1\leqq k\leqq n} \sup_{0\leqq x\leqq 1} |k[F_k(x) - x] - \sum_{i=1}^{k} \mathbb{B}_i(x)| \leqq C \quad \text{almost surely}$$

for some finite constant C. Then, as in Csörgő and Révész ([12], pp. 22, 58, 59), one can construct on an enlarged probability space, a Kiefer process $\mathbb{K} = \{\mathbb{K}(x,t) : (x,t) \in [0,1] \times [0,\infty)\}$ such that $\mathbb{K}(x,n) = \sum_{i=1}^{n} \mathbb{B}_i(x)$ for all $x \in [0,1]$ and $n \geqq 1$. For each integer $n \geqq 1$, define a mean-zero Gaussian process $\mathbb{H}_n = \{\mathbb{H}_n(x,t) : (x,t) \in [0,1]^2\}$ by

$$\mathbb{H}_n(x,t) := n^{-1/2}(\mathbb{K}(x,nt) - t\mathbb{K}(x,n)).$$

Then, adding and subtracting a Gaussian process that is defined by replacing each indicator $\mathbf{1}_{[0,x]}(U_i)$ in (3.1) by the Brownian bridge processes $\mathbb{B}_i(x)$, we get

$$\sup_{(x,t)\in[0,1]^2} |\mathbb{D}_n(x,t) - \sqrt{n}\mathbb{H}_n(x,t)| \leqq \max_{1\leqq k\leqq n} \|k[F_k - F] - \mathbb{K}(\cdot,k)\|_\infty$$

$$(3.4)\qquad + \|n[F_n - F] - \mathbb{K}(\cdot,n)\|_\infty$$

$$+ \sup_{0\leqq t\leqq n} \sup_{0\leqq s\leqq 1} \|\mathbb{K}(\cdot,t+s) - \mathbb{K}(\cdot,t)\|_\infty + \|\mathbb{K}(\cdot,n)\|_\infty / n =: A_n,$$

where $F(x) = x$ for all $x \in [0,1]$. By Corollary 1.12.4 of Csörgő and Révész [12], we have

$$(3.5)\quad \overline{\lim_{n\to\infty}} (\ln n)^{-1/2} \sup_{0\leqq t\leqq n} \sup_{0\leqq s\leqq 1} \|\mathbb{K}(\cdot,t+s) - \mathbb{K}(\cdot,t)\|_\infty \leqq 2^{3/2}$$

almost surely. Moreover, for a Kiefer process $\mathbb{K}$, we have (cf. Corollary 1.15.1 of Csörgő and Révész, [12])

$$(3.6)\quad \overline{\lim_{n\to\infty}} \frac{\|\mathbb{K}(\cdot,n)\|_\infty}{\sqrt{n\ln_2 n}} = 1/\sqrt{2} \quad \text{almost surely.}$$

Considering the suprema over $Q_2(\mu_n)$ separately over each corner of the square $[0,1]^2$ it follows that, for each $n \geqq 1$,

$$\sup_{(x,t)\in Q_2(\mu_n)} \frac{1}{\sqrt{nt(1-t)x(1-x)}} \leqq \frac{4}{(\ln n)^{5/2}}.$$

Then applying (3.3), (3.5) and (3.6) to estimate (3.4), we get

$$\varlimsup_{n\to\infty} \sqrt{\ln_2 n} \sup_{(x,t)\in Q_2(\mu_n)} \frac{|\mathbb{D}_n(x,t) - \sqrt{n}\mathbb{H}_n(x,t)|}{\sqrt{nt(1-t)x(1-x)}} \leq \varlimsup_{n\to\infty} \frac{(\ln_2 n)^{1/2}}{(\ln n)^{5/2}} A_n = 0.$$

Therefore (3.2) holds and the proof of Lemma 3.1 is complete. □

LEMMA 3.2. *If $\kappa > 0$ then, for all γ, sufficiently close to 1,*

$$\lim_{n\to\infty} \Pr\left(\left\{ \sup_{(x,t)\in Q_n} \frac{\left|\sum_{i=1}^{[nt]} \left(\mathbf{1}_{[0,x]}(U_i) - x\right)\right|}{\sqrt{ntx}} > 2\sqrt{\gamma \ln_2 n}\right\}\right) = 0, \tag{3.7}$$

where $Q_n := \{(x,t)\in[0,1]^2\colon 2\ln_2 n \leqq nxt \leqq (\ln n)^\kappa\}$.

PROOF. The proof is obtained by breaking Q_n into three subsets and estimating the corresponding probabilities separately using exponential inequalities for empirical processes.

Define a function ψ on $(0,\infty)$ by

$$\psi(u) := \frac{2}{u^2}\int_0^u \ln(1+v)\,dv \quad \text{for } u > 0. \tag{3.8}$$

Since $2\psi(\sqrt{2}) > 1$ then, for any γ sufficiently close to 1, we have

$$\mu(\gamma) := 2\gamma^3\psi\left(\sqrt{2}\right) > 1. \tag{3.9}$$

Let $\kappa > 0$ and let $\gamma \in [1/4, 1)$ be such that $\mu(\gamma) > 1$. For each $n \geqq 3$, define $k_n := \min\{k \in \mathbb{N}\colon (\ln n)^\kappa/k \leqq \eta\}$, where $\eta := 1 - \sqrt{\gamma} \leqq 1/2$. Then, denoting $Y_i(x) := (\mathbf{1}_{[0,x]}(U_i) - x)/\sqrt{x}$ for $x \in (0,1)$, we have

$$\begin{aligned}
&\Pr\left(\left\{ \sup_{(x,t)\in Q_n} \frac{\left|\sum_{i=1}^{[nt]} \left(\mathbf{1}_{[0,x]}(U_i) - x\right)\right|}{\sqrt{ntx}} > 2\sqrt{\gamma \ln_2 n}\right\}\right)\\
&\leqq \Pr\left(\left\{ \max_{k_n\leqq k\leqq n} \sup_{x\in[2\ln_2 n/k,(\ln n)^\kappa/k]} \left|\sum_{i=1}^k Y_i(x)\right|/\sqrt{k} > 2\sqrt{\gamma\ln_2 n}\right\}\right)\\
&\quad + \Pr\left(\left\{ \max_{1\leqq k\leqq k_n} \sup_{x\in[2\ln_2 n/k,\eta]} \left|\sum_{i=1}^k Y_i(x)\right|/\sqrt{k} > 2\sqrt{\gamma\ln_2 n}\right\}\right)\\
&\quad + \Pr\left(\left\{ \max_{1\leqq k\leqq k_n} \sup_{x\in[\eta,(\ln n)^\kappa/k]} \left|\sum_{i=1}^k Y_i(x)\right|/\sqrt{k} > 2\sqrt{\gamma\ln_2 n}\right\}\right)
\end{aligned} \tag{3.10}$$

$$=: P_{n,1} + P_{n,2} + P_{n,3},$$

where a suprema over an empty set is defined to be zero.

To estimate $P_{n,1}$, for each $n \geqq 3$ and $l \geqq 1$, let

$$Q(l) := Q_n(l) := \{x \in [0,1]\colon 2\gamma^l \ln_2 n \leqq x \leqq \gamma^{l-1}(\ln n)^\kappa\}$$

and let m_n and M_n denote respectively the smallest and the largest among all integers l such that

$$\Delta(n,l) := \{k \in \mathbb{N}\colon k_n \leqq k \leqq n\} \cap [\gamma^{-l+1}, \gamma^{-l}] \neq \emptyset. \tag{3.11}$$

Then we have

$$P_{n,1} \leqq \sum_{l=m_n}^{M_n} \Pr\Big(\Big\{\max_{k\in\Delta(n,l)} \Big\|\sum_{i=1}^k Y_i\Big\|_{Q(l)} > 2\gamma\sqrt{\gamma^{-l}\ln_2 n}\Big\}\Big). \tag{3.12}$$

Using exponential inequality (3) on p. 446 of Shorack and Wellner [27] we get, for each $k \in \Delta(n,l)$,

$$\Pr\Big(\Big\{\Big\|\sum_{i=k+1}^{[\gamma^{-l}]\wedge n} Y_i\Big\|_{Q(l)} > 2\gamma(1-\gamma)\sqrt{([\gamma^{-l}]\wedge n)\ln_2 n}\Big\}\Big)$$

$$\leqq 2\max_{k\in\Delta(n,l)} \Pr\Big(\Big\{\Big\|\sum_{i=1}^k Y_i\Big\|_{Q(l)} > \gamma(1-\gamma)\sqrt{k\ln_2 n}\Big\}\Big) \tag{3.13}$$

$$\leqq \frac{12}{\eta}\ln\Big((\ln n)^\kappa/\gamma\Big)(\ln n)^{-c},$$

where $c := \gamma^3(1-\gamma)^2\psi(\sqrt{\gamma}(1-\gamma)/2\sqrt{2})/2 > 0$. There exists an integer N such that the right-hand side of (3.13) is $\leqq 1/2$. Then using Octaviani inequality and exponential inequality (3) on p. 446 of Shorack and Wellner [27], again, we get, for each $n \geqq N$ and all $m_n \leqq l \leqq M_n$,

$$\Pr\Big(\Big\{\max_{k\in\Delta(n,l)} \Big\|\sum_{i=1}^k Y_i\Big\|_{Q(l)} > 2\gamma\sqrt{\gamma^{-l}\ln_2 n}\Big\}\Big)$$

$$\leqq \frac{\Pr\Big(\Big\{\Big\|\sum_{i=1}^{[\gamma^{-l}]\wedge n} Y_i\Big\|_{Q(l)} > 2\gamma^2\sqrt{([\gamma^{-l}]\wedge n)\ln_2 n}\Big\}\Big)}{1 - \max_{k\in\Delta(n,l)} \Pr\Big(\Big\{\Big\|\sum_{i=k+1}^{[\gamma^{-l}]\wedge n} Y_i\Big\|_{Q(l)} > 2\gamma(1-\gamma)\sqrt{([\gamma^{-l}]\wedge n)\ln_2 n}\Big\}\Big)}$$

$$\leqq 2\Pr\left(\left\{\left\|\sum_{i=1}^{[\gamma^{-l}]\wedge n} Y_i\right\|_{Q(l)} > 2\gamma^2\sqrt{([\gamma^{-l}]\wedge n)\ln_2 n}\right\}\right)$$

$$\leqq \frac{12}{\eta}\ln\Big((\ln n)^\kappa/\gamma\Big)\exp\Big\{-\mu(\gamma)\ln_2 n\Big\},$$

where $\mu(\gamma) > 1$ is defined by (3.9). Since $M_n \leqq 1 - \ln n/\ln\gamma$, inserting the last estimate into (3.12) and summing over l we get

$$\text{(3.14)}\qquad \overline{\lim_{n\to\infty}}\, P_{n,1} \leqq \frac{12}{\eta}\,\overline{\lim_{n\to\infty}}\Big(1-\ln n/\ln\gamma\Big)\ln\Big((\ln n)^\kappa/\gamma\Big)(\ln n)^{-\mu(\gamma)} = 0.$$

To show (3.14) we were anxious to get the estimate of order $(\ln n)^{-1-\epsilon}$ for some $\epsilon > 0$ of each probability in (3.12) because the number of such probabilities is $\leqq$ const $\times \ln n$. To estimate $P_{n,2}$ and $P_{n,3}$ our task is easier because the inside maximum is taken over $\{1,\ldots,k_n\}$ rather than over $\{k_n,\ldots,n\}$ and $k_n < (\ln n)^\kappa/\eta$. So, after division of the set $\{1,\ldots,k_n\}$ into blocks $\{2^{l-1},\ldots,2^l\}$, we are facing a sum whose cardinality $\leqq \ln k_n/\ln 2$ $\leqq$ const $\times \ln_2 n$. Therefore, it suffices to get an estimate of order $(\ln n)^{-c}$ for some $c > 0$ of the corresponding probabilities. To this aim, as above, using Octaviani inequality in conjunction with exponential inequality (3) on p. 446 of Shorack and Wellner [27], we get

$$\text{(3.15)}\qquad \overline{\lim_{n\to\infty}}\, P_{n,2} \leqq \frac{12\kappa}{\eta}\,\overline{\lim_{n\to\infty}}\,(\ln_2 n)^2(\ln n)^{-c} = 0,$$

where $c = \gamma^2\psi(\sqrt{\gamma}/2)/4 > 0$. Moreover, using again Octaviani inequality in conjunction with Dvoretzky–Kiefer–Wolfowitz exponential inequality (cf. p. 354 in Shorack and Wellner, [27]) it follows that

$$\text{(3.16)}\qquad \overline{\lim_{n\to\infty}}\, P_{n,3} \leqq \frac{116\kappa}{\ln 2}\,\overline{\lim_{n\to\infty}}\,(\ln_2 n - \ln\eta)(\ln n)^{-\eta\gamma} = 0.$$

Now going back to (3.10), we use (3.14), (3.15) and (3.16) to conclude that (3.7) holds. □

REFERENCES

[1] BICKEL, P. and ROSENBLATT, M., Two-dimensional random fields, *Multivariate Analysis*, III (Proc. Third Internat. Sympos., Wright State Univ., Dayton, Ohio, 1972), ed. P. R. Krishnaiah, Academic Press, New York, 1973, 3–15. *MR* **50** #1327

[2] CHIBISOV, D., Some theorems on the limiting behaviour of an empirical distribution function, *Trudy Mat. Inst. Steklov.* **71** (1964), 104–112 (in Russian). *MR* **30** #2543. See also: *Selected Transl. Math. Statist. Probab.* **6** (1964), 147–156.

[3] CSÁKI, E., On tests based on empirical distribution functions, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **23** (1974), 239–327 (1977) (in Hungarian). *MR* **57**

#4415. English translation: *Selected Translations Math. Statist. Probab.* **15**. American Mathematical Society, Providence, RI, 1981, 229–317.

[4] Csörgő, M., A new proof of some results of Rényi and the asymptotic distribution of the range of his Kolmogorov–Smirnov type random variables, *Canad. J. Math.* **19** (1967), 550–558. *MR* **35** #3712

[5] Csörgő, M., Empirical processes with applications to statistics, by Galen R. Shorack and John A. Wellner, Wiley, 1986. Review in *Bull. Amer. Math. Soc.* (New Series) **17** (1987), 189–200.

[6] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[7] Csörgő, M. and Horváth, L., Nonparametric methods for changepoint problems, *Handbook of Statistics*, Vol. 7, Ed. by P. R. Krishnaiah and C. R. Rao, North-Holland, Amsterdam–New York–Oxford, 1988, 403–425. See *MR* **90i**:62125

[8] Csörgő, M. and Horváth, L., Rényi-type empirical processes, *J. Multivariate Anal.* **41** (1992), 338–358. *MR* **93g**:60069

[9] Csörgő, M. and Horváth, L., *Weighted approximations in probability and statistics*, Wiley Series in Probability and Mathematical Statistics. Probability and Mathematical Statistics, John Wiley & Sons, Chichester, 1993. *MR* **94c**:60060

[10] Csörgő, M. and Horváth, L., *Limit theorems in change-point analysis*, John Wiley & Sons, Chichester, 1997.

[11] Csörgő, M., Horváth, L. and Szyszkowicz, B., Integral tests for suprema of Kiefer processes with application, *Statistics & Decisions* **15** (1997), 365–377.

[12] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Akadémiai Kiadó, Budapest and Academic Press, New York, 1981. *MR* **84d**:60050

[13] 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

[14] Csörgő, M. and Szyszkowicz, B., Applications of multi-time parameter processes to change-point analysis, *Probability theory and Mathematical Statistics VSP/TEV*, 1994, 159–222.

[15] Darhovskii, B. S., A non-parametric method for the a posteriori detection of the "disorder" time of a sequence of independent random variables, *Teor. Verojatnost. i Primenen.* **21** (1976), 180–184 (in Russian). *MR* **54** #3952. See also *Theory Probab. Appl.* **21** (1976), 178–183.

[16] Darling, D. A. and Erdős, P., A limit theorem for the maximum of normalized sums of independent random variables, *Duke Math. J.* **23** (1956), 143–155. *MR* **17**, 635c

[17] Deshayes, J. and Picard, D., Ruptures de modèles en statistique, Thèses d'État, Université Paris Sud, 1983.

[18] Deshayes, J. and Picard, D., Off-line statistical analysis of change-point models using nonparametric and likelihood methods, Lect. Notes in Control and Inform. Sci. **77** (1986), 103–168.

[19] Eicker, F., The asymptotic distribution of the suprema of standardized empirical processes, *Ann. Statist.* **7** (1979), 116–138. *MR* **80g**:62010

[20] Jaeschke, D., The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals, *Ann. Statist.* **7** (1979), 108–115. *MR* **80g**:62009

[21] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's and the sample DF. I, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **32** (1975), 111–131. *MR* **51** #11605b

[22] Leadbetter, M. R., Lindgren, G. and Rootzén, H., *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York–Berlin, 1983. *MR* **84h**:60050

[23] MASON, D. M., The asymptotic distribution of generalized Rényi statistics, *Acta Sci. Math. (Szeged)* **48** (1985), 315–323. *MR* **87e**:60042

[24] O'REILLY, N., On the weak convergence of empirical processes in sup-norm metrics, *Ann. Probability* **2** (1974), 642–651. *MR* **52** #4367

[25] PICARD, D., Testing and estimating change-points in time series, *Adv. in Appl. Probab.* **17** (1985), 841–867. *MR* **87a**:62122

[26] RÉNYI, A., On the theory of order statistics, *Acta. Math. Acad. Sci. Hungar.* **4** (1953), 191–231. *MR* **15**, 885e

[27] 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 & Sons, New York, 1986. *MR* **88e**:60002

[28] SZYSZKOWICZ, B., Empirical type processes and contiguity, *C. R. Math. Rep. Acad. Sci. Canada* **13** (1991), 161–166. *MR* **92i**:60063

[29] SZYSZKOWICZ, B., Weak convergence of stochastic processes in weighted metrics and their applications to contiguous changepoint analysis, Ph. D. Thesis, Carleton University, Ottawa, 1992.

[30] SZYSZKOWICZ, B., Weak convergence of weighted empirical type processes under contiguous and changepoint alternatives, *Stochastic Process. Appl.* **50** (1994), 281–313. *MR* **95e**:62057

(Received November 12, 1997)

DEPARTMENT OF MATHEMATICS AND STATISTICS
CARLETON UNIVERSITY
1125 COLONEL BY DRIVE
OTTAWA, ONTARIO
K1S 5B6
CANADA

mcsorgo@math.carleton.ca

DEPARTMENT OF MATHEMATICAL SCIENCES
INSTITUTE OF MATHEMATICS AND INFORMATICS
AKADEMIJOS 4
LT–2600 VILNIUS
LITHUANIA

norvaisa@ktl.mii.lt

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 71–87

RÉNYI CONFIDENCE BANDS [1]

S. CSÖRGŐ

Dedicated to the memory of Alfréd Rényi

Abstract

Rényi's asymptotic confidence bands for distribution or survival functions, the width of which at each point is proportional to the natural estimator of the function to be estimated, are shown to extend far out to small and large order statistics, respectively. Certain combinations of these bands are also proposed.

1. The bands

Let $X_1, \ldots, X_n$ be a sample of size $n \in \mathbf{N} := \{1, 2, \ldots\}$, independent random variables with the common distribution function $F(x) := P\{X \leqq x\}$, $x \in \mathbf{R}$, where $F(\cdot)$ is assumed to be a *continuous* function on the whole real line $\mathbf{R}$ throughout this paper. Denoting by $F_n(x) := \#\{1 \leqq j \leqq n : X_j \leqq x\}/n$, $x \in \mathbf{R}$, the sample distribution function, Kolmogorov's well-known result from 1933 is that

$$P\left\{\sqrt{n} \sup_{x \in \mathbf{R}} |F_n(x) - F(x)| \leqq y\right\} \to K(y) := 1 + 2\sum_{k=1}^{\infty}(-1)^k e^{-2k^2y^2}$$

for every $y > 0$, where an unspecified convergence or asymptotic order relation is meant to hold as $n \to \infty$ everywhere in the paper. In 1949, Doob identified $K(\cdot)$ as the distribution function of the random variable $\sup_{0 \leqq s \leqq 1} |B(s)|$, where $\{B(s) : 0 \leqq s \leqq 1\}$ is a Brownian bridge, a sample-continuous Gaussian process with zero mean and covariance $E(B(s)B(t)) = \min(s,t) - st$, $0 \leqq s, t \leqq 1$. Thus, letting $\xrightarrow{\mathcal{D}}$ denote convergence in distribution, Kolmogorov's theorem may be written as

$$\sqrt{n} \sup_{x \in \mathbf{R}} |F_n(x) - F(x)| \xrightarrow{\mathcal{D}} \sup_{0 \leqq s \leqq 1} |B(s)| .$$

1991 *Mathematics Subject Classification.* Primary 62G05; Secondary 60F05.

Key words and phrases. Asymptotic confidence bands, distribution and survival functions.

[1] Read at the Rényi Memorial Session of the Hungarian Academy of Sciences, Budapest, November 29, 1996. Work partially supported by the NSF, USA, Grant DMS–96-25732.

Let $y_\alpha > 0$ be the unique number for which $K(y_\alpha) = 1 - \alpha$, where $\alpha \in (0,1)$ is a fixed number throughout. For testing simple goodness of fit or for estimating the unknown F, the resulting statement for the Kolmogorov confidence band is that $P\{F(x) \in [F_n(x) - y_\alpha n^{-1/2}, F_n(x) + y_\alpha n^{-1/2}], x \in \mathbf{R}\} \to 1 - \alpha$. The corresponding half-sided asymptotic confidence lines were derived by Smirnov in 1939.

Twenty years after the publication of Kolmogorov's theorem, dedicating his paper to Kolmogorov's fiftieth birthday, Rényi [12] proved that for each fixed $p \in (0,1)$ and all $y > 0$,

$$P\left\{\sqrt{\frac{np}{1-p}} \sup_{F(x) \leqq 1-p} \frac{F_n(x) - F(x)}{1 - F(x)} \leqq y\right\} \to 2\Phi(y) - 1, \tag{1}$$

where $\Phi(\cdot)$ stands for the standard normal distribution function, and

$$P\left\{\sqrt{\frac{np}{1-p}} \sup_{F(x) \leqq 1-p} \frac{|F_n(x) - F(x)|}{1 - F(x)} \leqq y\right\} \to L(y), \tag{2}$$

where

$$L(y) := \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-\frac{(2k+1)^2 \pi^2}{8y^2}}.$$

These are in fact the right-tail versions of Rényi's Theorems 5 and 6, motivated by the problem of estimation of the survival function $1 - F(x) = P\{X > x\}$, $x \in \mathbf{R}$. The (extended forms of the) corresponding, mathematically equivalent left-tail versions are in (5) and (6) below. Rényi's paper [12] has been eminently influential in the directions that the development of the Hungarian school of probability and mathematical statistics has taken.

Exposing the statistical idea behind (the left-tail versions), Rényi [12] writes:

> Kolmogorov's theorem considers the difference $|F_n(x) - F(x)|$ with the same weight, regardless to the value of $F(x)$; so e.g. the difference $|F_n(x) - F(x)| = 0.01$ has the same weight in a point x with $F(x) = 0.5$ (where this difference is 2% of the value of $F(x)$) as in a point x with $F(x) = 0.01$ (where this difference is 100% of the value of $F(x)$!). We can avoid this by considering the quotient $|F_n(x) - F(x)|/F(x)$ instead of $|F_n(x) - F(x)|$, that is to say, by considering the *relative error* of $F_n(x)$.

The resulting left-tail theorems then yield simple goodness-of-fit tests for the hypothesis that the underlying distribution function is indeed $F(\cdot)$:

> The character of these tests consists in that they give a band around $F(x)$ in which, if the hypothesis is true, the sample distribution function $F_n(x)$ has to lie with a certain probability and the width of this band in all points x is proportional to $F(x)$.

The results are also stated and discussed in Rényi's textbooks, see for example Chapter VIII, §10 in [14], his last book.

Alternative, asymptotically equivalent tests result from replacing the weight functions $F(\cdot)$ and $1-F(\cdot)$ by $F_n(\cdot)$ and $1-F_n(\cdot)$ in the denominators of the test statistics for the left-tail and right-tail tests, respectively. That these replacements can in fact be done was pointed out by M. Csörgő [2] along with several other variants. In his rejoinder to the developments inspired by his 1953 paper in the ensuing fifteen years, Rényi [13] himself has indicated an easier way to do this. Also, reducing Rényi's results to applications of Donsker's weak convergence theorem for the empirical process, M. Csörgő [3] obtained the two limiting distribution functions in (1) and (2) above as the distribution functions of the random variables $\sup_{0\leqq t\leqq 1} W(t)$ and $\sup_{0\leqq t\leqq 1} |W(t)|$, respectively, where $\{W(t): t\geqq 0\}$ is a standard Wiener process, a sample-continuous Gaussian process with zero mean and covariance $E(W(s)W(t)) = \min(s,t)$, $s,t\geqq 0$. (See also p. 165 in [6].) In what follows $\{W_*(t): t\geqq 0\}$ will denote another standard Wiener process, *independent* of $\{W(t): t\geqq 0\}$. A great amount of sophisticated work went into the determination of the exact distributions of Kolmogorov, Smirnov and Rényi-type statistics. A unified theory of this field has been given by Csáki [1], where the relevant references may also be found; references mentioned but not specifically given in the present paper are all included in Csáki's list of 109 items, or in [6].

The price Rényi's theorems pay for the consideration of relative errors is that they exclude a whole fixed proportion p of the sample from analysis even asymptotically, or $100p\,\%$, either the smallest or the largest observations. It is not possible to extend the supremum to the whole support of the distribution since by a classical theorem of Henry Daniels from 1945, for every sample size $n\in\mathbf{N}$,

$$P\left\{\sup_{F(x)>0}\frac{F_n(x)}{F(x)}>y\right\}=\frac{1}{y}\quad\text{for all } y\geqq 1.$$

(Note that Rényi [13] gives a half-page proof of this, based on his representation of order statistics, which representation is at the heart of his method in [12].) However, taking the limit in the corresponding exact formulae of his, Csáki [1] was able to show in his Theorems 2.8 and 2.9 that (1) and its left-tail version still remain true if p is changed to $p_n\in(0,1)$ such that $p_n\to 0$, provided $np_n\to\infty$.

As one of the easiest applications of the weighted approximations described in the next section, it was proved in [4] that if $\{p_n\}_{n=1}^{\infty}$ is a sequence such that $0<p_n\leqq p$ for some $p\in(0,1)$, for all $n\in\mathbf{N}$, and $np_n\to\infty$, then

$$\text{(3)}\qquad \sqrt{\frac{np_n}{1-p_n}}\sup_{F(x)\leqq 1-p_n}\frac{F_n(x)-F(x)}{1-F(x)}\xrightarrow{\mathcal{D}}\sup_{0\leqq t\leqq 1}W(t)\,,$$

$$\sqrt{\frac{np_n}{1-p_n}} \sup_{F(x)\leqq 1-p_n} \frac{|F_n(x)-F(x)|}{1-F(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} |W(t)|\,, \tag{4}$$

$$\sqrt{\frac{np_n}{1-p_n}} \sup_{p_n\leqq F(x)} \frac{F_n(x)-F(x)}{F(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} W_*(t)\,, \tag{5}$$

$$\sqrt{\frac{np_n}{1-p_n}} \sup_{p_n\leqq F(x)} \frac{|F_n(x)-F(x)|}{F(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} |W_*(t)|\,. \tag{6}$$

If $p_n \equiv p$, these give Rényi's theorems, the case $p_n \to 0$ in (3) and (5) are Csáki's results. To make the present paper self-contained, the proofs of (3) and (4) are given in the next section; the inclusion of these makes it easier to present the proofs of the main results. In addition to the individual four convergence statements in (3)–(6), Mason [11] has shown that these statements hold in fact *jointly* if $p_n \to 0$. The reason for the asymptotic independence of the left-tail and the right-tail statistics is the fact, as his proof reveals in an exact fashion, that the maximal deviations in (3) and (4) occur for $F(x)$ near $1-p_n$, while in (5) and (6) for $F(x)$ near p_n, and if $p_n \to 0$ the extreme order statistics determining these suprema, taken close to p_n and to $1-p_n$, become sufficiently remote to yield asymptotic independence. A version of the argument is in the proof of Theorem 1 below.

The "continuity" in the results in (3)–(6) is remarkable: the smaller p_n is, the larger is the stochastic order of the largest one-sided and two-sided deviations, but the distributional limits remain the same as long as $np_n \to \infty$. The condition that $np_n \to \infty$ is necessary for the latter: the second parts of Csáki's [1] Theorems 2.8 and 2.9 show for the cases (3) and (5) that while the "continuity" concerning stochastic order still holds even when $p_n \equiv v/n$ for any fixed $v>0$, as expected in view of Daniels' result above, the limiting distributions change drastically.

Simple goodness-of-fit tests may be built on the test statistics in (3)–(6) as before: in fact they all become consistent when $p_n \to 0$. However, if a null hypothesis does not specify F (and, for various well-argued reasons, *simple* goodness-of-fit tests do appear to have been abandoned in statistical practice in the last two decades or so), then the statistics in (3)–(6) are not determined. In particular, confidence bands for $1-F$ or F can be constructed only on intervals determined by the observations and not the unknown F. The question is whether the same results can be retained when the one-sided and two-sided maximal relative error of F_n is taken over the set $\{x: p_n \leqq F_n(x)\}$ rather than $\{x: p_n \leqq F(x)\}$ in (5) and (6), and when the maximal relative errors of $1-F_n$ are taken over the set $\{x: F_n(x) \leqq 1-p_n\}$ rather than $\{x: F(x) \leqq 1-p_n\}$ in (3) and (4). When $p_n \equiv p$, this is easy to do; in fact the very first step of Rényi's [12] original proof is to show that the two kinds of results are equivalent. However, it is far from obvious whether this is true for *all* sequences $\{p_n\}$ such that $p_n \to 0$ and $np_n \to \infty$. The aim of this paper is to pay tribute to Rényi's memory by showing that the answer is affirmative, and hence the construction of *extended* asymptotic Rényi confidence bands is in fact possible.

It is more suggestive from the statistical point of view when confidence bands are drawn on intervals determined directly by the order statistics $X_{1,n} \leqq \cdots \leqq X_{n,n}$ of the sample $X_1, \ldots, X_n$. This is why we state the main results in the form given below. Also, some recent results for confidence bands with censored data are formulated in a similar fashion in [7], so that comparisons will be the easiest this way. The theorem determines how far out Rényi confidence bands hold.

THEOREM 1. *Let $\{k_n\}_{n=1}^{\infty}$ be a sequence of integers such that $1 \leqq k_n \leqq np$, $n \geqq 1/p$, for some $p \in (0,1)$ and $k_n \to \infty$. Then the six convergence statements*

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \leqq X_{n-k_n,n}} \frac{F_n(x)-F(x)}{1-F(x)} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} W(t), \tag{7}$$

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \inf_{x \leqq X_{n-k_n,n}} \frac{F_n(x)-F(x)}{1-F(x)} \xrightarrow{\mathcal{D}} \inf_{0 \leqq t \leqq 1} W(t), \tag{8}$$

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \leqq X_{n-k_n,n}} \frac{F_n(x)-F(x)}{1-F_n(x)} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} W(t), \tag{9}$$

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \inf_{x \leqq X_{n-k_n,n}} \frac{F_n(x)-F(x)}{1-F_n(x)} \xrightarrow{\mathcal{D}} \inf_{0 \leqq t \leqq 1} W(t), \tag{10}$$

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \leqq X_{n-k_n,n}} \frac{|F_n(x)-F(x)|}{1-F(x)} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} |W(t)|, \tag{11}$$

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \leqq X_{n-k_n,n}} \frac{|F_n(x)-F(x)|}{1-F_n(x)} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} |W(t)| \tag{12}$$

take place jointly. Also, the six convergence statements

$$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \geqq X_{k_n,n}} \frac{F_n(x)-F(x)}{F(x)} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} W_*(t), \tag{13}$$

(14) $$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \inf_{x \geqq X_{k_n,n}} \frac{F_n(x)-F(x)}{F(x)} \xrightarrow{\mathcal{D}} \inf_{0\leqq t\leqq 1} W_*(t),$$

(15) $$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \geqq X_{k_n,n}} \frac{F_n(x)-F(x)}{F_n(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} W_*(t),$$

(16) $$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \inf_{x \geqq X_{k_n,n}} \frac{F_n(x)-F(x)}{F_n(x)} \xrightarrow{\mathcal{D}} \inf_{0\leqq t\leqq 1} W_*(t),$$

(17) $$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \geqq X_{k_n,n}} \frac{|F_n(x)-F(x)|}{F(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} |W_*(t)|,$$

(18) $$\sqrt{\frac{k_n}{1-\frac{k_n}{n}}} \sup_{x \geqq X_{k_n,n}} \frac{|F_n(x)-F(x)|}{F_n(x)} \xrightarrow{\mathcal{D}} \sup_{0\leqq t\leqq 1} |W_*(t)|$$

take place jointly. Furthermore, if $k_n/n \to 0$, then all twelve statements hold jointly.

Let $z_\alpha > 0$ be the unique value for which $L(z_\alpha) = 1-\alpha$. Linearly interpolating between neighbouring values of the table in [8] and rounding off to three decimals, we have $z_{0.01} = 2.806$, $z_{0.03} = 2.433$, $z_{0.05} = 2.241$, $z_{0.07} = 2.108$, $z_{0.1} = 1.960$, $z_{0.15} = 1.780$ and $z_{0.2} = 1.599$, for example. Let

$$c_{n,k_n}^-(\alpha) := 1 - z_\alpha \frac{\sqrt{1-\frac{k_n}{n}}}{\sqrt{k_n}} \quad \text{and} \quad c_{n,k_n}^+(\alpha) := 1 + z_\alpha \frac{\sqrt{1-\frac{k_n}{n}}}{\sqrt{k_n}}.$$

Then, for survival functions, (11) implies that

$$P\left\{\frac{1-F_n(x)}{c_{n,k_n}^+(\alpha)} \leqq 1-F(x) \leqq \frac{1-F_n(x)}{c_{n,k_n}^-(\alpha)}, \quad x \leqq X_{n-k_n,n}\right\} \to 1-\alpha$$

and (12) implies that

$$P\Big\{c_{n,k_n}^-(\alpha)[1-F_n(x)] \leqq 1-F(x) \leqq c_{n,k_n}^+(\alpha)[1-F_n(x)],\ x \leqq X_{n-k_n,n}\Big\} \to 1-\alpha.$$

Notice that the lower boundary of the first band lies everywhere above the lower boundary of the second band, while the upper boundary of the second band lies everywhere beneath the upper boundary of the first. This fact suggests to consider the inner envelope band $I_n^{[1-F]}(\cdot) := [\{1-F_n(\cdot)\}/$

$c^{+}_{n,k_n}(\alpha), c^{+}_{n,k_n}(\alpha)\{1-F_n(\cdot)\}]$ for the survival function $1-F(\cdot)$. (Ideas of this sort appeared first in [9] for estimation under censorship.) The same phenomenon occurs for the corresponding bands $[F_n(\cdot)/c^{+}_{n,k_n}(\alpha), F_n(\cdot)/c^{-}_{n,k_n}(\alpha)]$ and $[c^{-}_{n,k_n}(\alpha)F_n(\cdot), c^{+}_{n,k_n}(\alpha)F_n(\cdot)]$ for $F(\cdot)$, resulting from (17) and (18) and both having asymptotic coverage probability $1-\alpha$ on the half-line $[X_{k_n,n}, \infty)$. The suggested uniformly narrower inner envelope band for the distribution function $F(\cdot)$ is then

$$J_n^{[F]}(\cdot) := \left[\frac{F_n(\cdot)}{c^{+}_{n,k_n}(\alpha)}, c^{+}_{n,k_n}(\alpha)F_n(\cdot)\right].$$

That the idea works is the first statement of Theorem 2, in (19) and (20) below.

For each $x \leqq X_{n-k_n,n}$, the width of the band $I_n^{[1-F]}(x)$ is $d^{(\alpha)}_{n,k_n}\{1-F_n(x)\}$ and, for each $x \geqq X_{k_n,n}$, the width of the band $J_n^{[F]}(x)$ is $d^{(\alpha)}_{n,k_n}F_n(x)$, proportional to $1-F_n(x)$ and $F_n(x)$, respectively, where $d^{(\alpha)}_{n,k_n} := \{(c^{+}_{n,k_n}(\alpha))^2-1\}/c^{+}_{n,k_n}(\alpha)$. Of course, $I_n^{[1-F]}(\cdot)$ can also written as a band for $F(\cdot)$ rather than for $1-F(\cdot)$, namely as

$$I_n^{[F]}(\cdot) := \left[1-c^{+}_{n,k_n}(\alpha)\{1-F_n(\cdot)\}, 1-\frac{1-F_n(\cdot)}{c^{+}_{n,k_n}(\alpha)}\right].$$

Then $I_n^{[F]}(x)$ is expected to be good for large x's near $X_{n-k_n,n}$ while $J_n^{[F]}(x)$ for small x's near $X_{k_n,n}$. Indeed, for any $c_n > 1$, simple algebra shows that $1-c_n\{1-F_n(x)\} < F_n(x)/c_n$ if and only if $F_n(x) < c_n/(1+c_n)$, and $1-\{1-F_n(x)\}/c_n \leqq c_nF_n(x)$ if and only if $F_n(x) \geqq 1/(1+c_n)$. These facts explain the choice below of the lower and upper boundary curves of an all-purpose two-sided inner envelope band combined from the previous two combined bands.

To introduce these curves, using again the function $L(\cdot)$ in (2), let $z^*_\alpha > 0$ be the unique value for which $L(z^*_\alpha) = \sqrt{1-\alpha}$. In comparison with the values of z_α above, note that $z^*_{0.01} = 3.025$, $z^*_{0.03} = 2.671$, $z^*_{0.05} = 2.495$, $z^*_{0.07} = 2.369$, $z^*_{0.1} = 2.231$, $z^*_{0.15} = 2.064$ and $z^*_{0.2} = 1.937$. Then set

$$c^{*}_{n,k_n}(\alpha) := 1 + z^*_\alpha \frac{\sqrt{1-\frac{k_n}{n}}}{\sqrt{k_n}}$$

and let $F_n^{-1}(s) := \inf\{x \in \mathbf{R} : F_n(x) \geqq s\}$, $0 < s \leqq 1$, be the sample quantile function, so that $F_n^{-1}(s) = X_{j,n}$ if $\frac{j-1}{n} < s \leqq \frac{j}{n}$, $j = 1, \dots, n$, $F_n^{-1}(0) := X_{1,n}$.

Then the lower and upper boundaries will be

$$L_{n,k_n}^{(\alpha)}(x) := \begin{cases} \frac{F_n(x)}{c^*_{n,k_n}(\alpha)} & , \ X_{k_n,n} \leqq x < F_n^{-1}\Big(\frac{c^*_{n,k_n}(\alpha)}{1+c^*_{n,k_n}(\alpha)}\Big), \\ 1 - c^*_{n,k_n}(\alpha)[1-F_n(x)] & , \ F_n^{-1}\Big(\frac{c^*_{n,k_n}(\alpha)}{1+c^*_{n,k_n}(\alpha)}\Big) \leqq x \leqq X_{n-k_n,n}, \end{cases}$$

and

$$U_{n,k_n}^{(\alpha)}(x) := \begin{cases} c^*_{n,k_n}(\alpha) F_n(x) & , \ X_{k_n,n} \leqq x < F_n^{-1}\Big(\frac{1}{1+c^*_{n,k_n}(\alpha)}\Big), \\ 1 - \frac{1-F_n(x)}{c^*_{n,k_n}(\alpha)} & , \ F_n^{-1}\Big(\frac{1}{1+c^*_{n,k_n}(\alpha)}\Big) \leqq x \leqq X_{n-k_n,n}, \end{cases}$$

noting also that both $c^*_{n,k_n}(\alpha)/[1+c^*_{n,k_n}(\alpha)] \to 1/2$ and $1/[1+c^*_{n,k_n}(\alpha)] \to 1/2$ under the condition of Theorem 1 on $\{k_n\}$, for every choice of $\alpha \in (0,1)$.

THEOREM 2. *Let $\{k_n\}_{n=1}^{\infty}$ be a sequence of integers such that $1 \leqq k_n \leqq np$, $n \geqq 1/p$, for some $p \in (0,1)$ and $k_n \to \infty$. Then*

$$(19) \qquad P\left\{ \frac{1-F_n(x)}{c^+_{n,k_n}(\alpha)} \leqq 1 - F(x) \leqq c^+_{n,k_n}(\alpha)[1-F_n(x)], \ x \leqq X_{n-k_n,n} \right\} \to 1-\alpha$$

and

$$(20) \qquad P\left\{ \frac{F_n(x)}{c^+_{n,k_n}(\alpha)} \leqq F(x) \leqq c^+_{n,k_n}(\alpha) F_n(x), \ X_{k_n,n} \leqq x \right\} \to 1-\alpha .$$

Furthermore, if $k_n/n \to 0$, then

$$(21) \qquad P\left\{ L^{(\alpha)}_{n,k_n}(x) \leqq F(x) \leqq U^{(\alpha)}_{n,k_n}(x), \ X_{k_n,n} \leqq x \leqq X_{n-k_n,n} \right\} \to 1-\alpha .$$

Of course, the factor $\sqrt{(n-k_n)/n}$ can be replaced by 1 everywhere when $k_n/n \to 0$. However, we prefer to keep it because its presence unifies the results, narrows the bands somewhat and, as the proofs show, makes the asymptotic approximations more natural. In some statistical situations it may not be natural to discard the same number of lower and upper extremes. If $1 \leqq m_n < n - k_n < n$ for integers $m_n \to \infty$ and $k_n \to \infty$ such that both $m_n/n \to 0$ and $k_n/n \to 0$, then the twelve joint convergence relations in (7)–(18) remain valid if everywhere in (13)–(18) we replace k_n by m_n. Consequently, a suitably modified form of (21) for the combined band also remains true on the interval $[X_{m_n,n}, X_{n-k_n,n}]$, in which the "change points" of the two boundaries depend on both $c^*_{n,m_n}(\alpha)$ and $c^*_{n,k_n}(\alpha)$.

Mason's [11] asymptotic independence result mentioned above appears in fact in a more general context. He extended (3)–(6) by allowing weight functions more general than $1 - F(\cdot)$ and $F(\cdot)$. For example, when $p_n \to 0$ and $np_n \to \infty$ he shows that

$$\left(\frac{np_n}{1-p_n}\right)^{\theta-\frac{1}{2}} \sup_{F(x)\leq 1-p_n} \frac{F_n(x)-F(x)}{[1-F(x)]^\theta} \xrightarrow{\mathcal{D}} \sup_{0\leq t\leq 1} \frac{W(t)}{t^\theta},$$

$$\left(\frac{np_n}{1-p_n}\right)^{\theta-\frac{1}{2}} \sup_{F(x)\leq 1-p_n} \frac{|F_n(x)-F(x)|}{[1-F(x)]^\theta} \xrightarrow{\mathcal{D}} \sup_{0\leq t\leq 1} \frac{|W(t)|}{t^\theta},$$

$$\left(\frac{np_n}{1-p_n}\right)^{\theta-\frac{1}{2}} \sup_{p_n\leq F(x)} \frac{F_n(x)-F(x)}{[F(x)]^\theta} \xrightarrow{\mathcal{D}} \sup_{0\leq t\leq 1} \frac{W_*(t)}{t^\theta},$$

$$\left(\frac{np_n}{1-p_n}\right)^{\theta-\frac{1}{2}} \sup_{p_n\leq F(x)} \frac{|F_n(x)-F(x)|}{[F(x)]^\theta} \xrightarrow{\mathcal{D}} \sup_{0\leq t\leq 1} \frac{|W_*(t)|}{t^\theta}$$

hold jointly, for any constant $\theta \in (1/2, 1]$. (Subsequent results and references in this direction are in Sections 5.1 and 5.5 of [5].) Mason [11] also proves that if we assume only that $0 < p_n \leq p$, $n \in \mathbf{N}$, for some $p \in (0,1)$ and $np_n \to \infty$, then the first two joint statements here remain true provided the extra factor $[F(x)]^{1-\theta}$ is included in the denominators and the last two joint statements remain true provided the extra factor $[1-F(x)]^{1-\theta}$ is included in the denominators. Starting out from these generalizations instead of (3)–(6), every statement in Theorems 1 and 2 has a natural generalization which reduces to the present one when $\theta = 1$.

Typically, the bands in (19) and (20) will be uselessly wide on the left tail and on the right tail, respectively, for the usual nominal coverage probabilities such as $1-\alpha = 0.9$. This will happen even for x's for which $F_n(x) \approx 1/2$ if we want to go far out on the tail of interest, that is, if k_n is chosen small. Of course, this will be even more so for the middle portion of the bands in (21). Rényi bands are for tail estimation. For that purpose, the flexibility in the choice of k_n is a real advantage. It will be of interest to determine by an extensive simulation study what combinations of the sample size n and the choice of k_n make the actual and nominal coverage probabilities acceptably close, what is the direction of their deviations, and whether the weight functions $[1-F(\cdot)]^\theta$ and $[F(\cdot)]^\theta$ for $\theta \in (1/2, 1)$ are statistically useful in these questions. Since all the bands are distribution-free for any finite sample size n, as will be made clear in the next section, only one such study is needed. A student of mine will look into these problems*.

2. Proofs

For a sequence $\{\xi_n\}_{n=1}^\infty$ of random variables and a sequence $\{a_n\}_{n=1}^\infty$ of positive constants we write $\xi_n = O_P(a_n)$ if $\lim_{y\to\infty} \limsup_{n\to\infty} P\{|\xi_n| >$

* Z. Megyesi, Coverage probabilities of Rényi confidence bands, this volume, 317–332.

$ya_n\} = 0$, and write $\xi_n = o_P(a_n)$ if $\lim_{n\to\infty} P\{|\xi_n| > ya_n\} = 0$ for every $y > 0$, that is, if $\xi_n/a_n \xrightarrow{P} 0$. The proofs will use a specially constructed probability space $(\Omega, \mathcal{A}, P)$ that carries a sequence $U_1, U_2, \ldots$ of independent random variables uniformly distributed on $(0,1)$, with order statistics $U_{1,n} \leqq \cdots \leqq U_{n,n}$ pertaining to $U_1, \ldots, U_n$ for each $n \in \mathbf{N}$, and a sequence $B_1(\cdot), B_2(\cdot), \ldots$ of Brownian bridges such that if $G_n(s) := \#\{1 \leqq j \leqq n\colon U_j \leqq s\}/n$ and $U_n(s) := \inf\{t \in [0,1] : G_n(t) \geqq s\}$, $0 \leqq s \leqq 1$, denote the corresponding uniform sample distribution and quantile functions, so that $U_n(s) = U_{k,n}$ for $\frac{k-1}{n} < s \leqq \frac{k}{n}$, $k = 1, \ldots, n$, and $U_n(0) = U_{1,n}$, then for the corresponding empirical and quantile processes $\alpha_n(s) := \sqrt{n}\,[G_n(s) - s]$ and $\beta_n(s) := \sqrt{n}\,[s - U_n(s)]$, $0 \leqq s \leqq 1$, we have

$$\sup_{0<s<1} \frac{|\alpha_n(s) - B_n(s)|}{[s(1-s)]^{\frac{1}{2}-\gamma}} = \mathcal{O}_P\Big(\frac{1}{n^\gamma}\Big) \quad \text{and} \quad \sup_{\frac{\lambda}{n} \leqq s \leqq 1 - \frac{\lambda}{n}} \frac{|\beta_n(s) - B_n(s)|}{[s(1-s)]^{\frac{1}{2}-\delta}} = \mathcal{O}_P\Big(\frac{1}{n^\delta}\Big)$$

for all fixed $\gamma \in (0, 1/4)$, $\delta \in (0, 1/2)$ and $\lambda > 0$. This construction was accomplished in [4]. (That the first supremum can be extended to the whole $(0,1)$ was pointed out in [10].) Note right away that under the conditions of Theorem 1, from the second relation,

$$\frac{n}{\sqrt{k_n}}\Big\{U_{n-k_n,n} - \Big(1 - \frac{k_n}{n}\Big)\Big\} = -\sqrt{\frac{n}{k_n}}\, B_n\Big(1 - \frac{k_n}{n}\Big) + \mathcal{O}_P\Big(\frac{1}{k_n^\delta}\Big)$$

(22) and

$$\frac{n}{\sqrt{k_n}}\Big\{U_{k_n,n} - \frac{k_n}{n}\Big\} = -\sqrt{\frac{n}{k_n}}\, B_n\Big(\frac{k_n}{n}\Big) + \mathcal{O}_P\Big(\frac{1}{k_n^\delta}\Big),$$

where $\mathcal{O}_P(1/k_n^\delta) = o_P(1)$ as $k_n \to \infty$,

$$P\Big\{-\sqrt{\frac{n}{k_n}}\, B_n\Big(1 - \frac{k_n}{n}\Big) < -x\Big\} = P\Big\{-\sqrt{\frac{n}{k_n}} B_n\Big(\frac{k_n}{n}\Big) < -x\Big\} = \Phi\left(\frac{-x}{\sqrt{1 - \frac{k_n}{n}}}\right) \leqq \Phi(-x)$$

and

$$P\Big\{-\sqrt{\frac{n}{k_n}} B_n\Big(1 - \frac{k_n}{n}\Big) > x\Big\} = P\Big\{-\sqrt{\frac{n}{k_n}}\, B_n\Big(\frac{k_n}{n}\Big) > x\Big\} = 1 - \Phi\left(\frac{x}{\sqrt{1 - \frac{k_n}{n}}}\right) \leqq 1 - \Phi(x) = \Phi(-x)$$

for every $n \in \mathbf{N}$ and $x > 0$.

Since $\{F_n(x) : x \in \mathbf{R}\} \stackrel{\mathcal{D}}{=} \{G_n(F(x)) : x \in \mathbf{R}\}$, the distributional equality meaning the equality of all finite-dimensional distributions of the two processes, the continuity of $F(\cdot)$ implies that all the statistics in (3)–(6) are distribution free. Furthermore, introducing the quantile function $F^{-1}(s) := \inf\{x \in \mathbf{R} : F(x) \geqq s\}$, $0 < s \leqq 1$, $F^{-1}(0) := \lim_{s \downarrow 0} F^{-1}(s)$, we also see, for the left-hand side of (7), that

$$\sup_{x \leqq X_{n-k_n,n}} \frac{F_n(x) - F(x)}{1 - F(x)} \stackrel{\mathcal{D}}{=} \sup_{F(x) \leqq F(F^{-1}(U_{n-k_n,n}))} \frac{G_n(F(x)) - F(x)}{1 - F(x)} = \sup_{0 \leqq s \leqq U_{n-k_n,n}} \frac{G_n(s) - s}{1 - s}$$

for each meaningful n; in fact the equality in distribution holds jointly in n, but we do not need this in the paper. It is clear then, in exactly the same way, that all the statistics in (7)–(18) are distribution free. Thus we may and do assume in the proofs below that the underlying distribution is uniform on the interval $(0, 1)$. It is no loss of generality, either, to assume without further notice that we are on the special probability space described above and, in particular, our statistics are based on the Uniform$(0, 1)$ order statistics for which the approximations above hold.

PROOF OF (3)–(6). As was stated already, this proof is from [4], where the left-tail versions are detailed. Choose any $\gamma \in (0, 1/4)$. Then we have

$$\begin{aligned}
&\sup_{0 \leqq s \leqq 1-p_n} \left| \sqrt{\frac{np_n}{1-p_n}} \frac{G_n(s) - s}{1 - s} - \sqrt{\frac{p_n}{1-p_n}} \frac{B_n(s)}{1-s} \right| \\
&= \sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq 1-p_n} \left| \frac{\alpha_n(s)}{1-s} - \frac{B_n(s)}{1-s} \right| \\
&\leqq \sqrt{\frac{p_n}{1-p_n}} \sup_{p_n \leqq 1-s \leqq 1} \frac{1}{(1-s)^{\frac{1}{2}+\gamma}} \sup_{0 \leqq s \leqq 1} \frac{|\alpha_n(s) - B_n(s)|}{(1-s)^{\frac{1}{2}-\gamma}} \qquad (23) \\
&\leqq \frac{1}{\sqrt{1-p}} \frac{p_n^{\frac{1}{2}}}{p_n^{\frac{1}{2}+\gamma}} \mathcal{O}_P\Big(\frac{1}{n^\gamma}\Big) \\
&= \mathcal{O}_P\Big(\frac{1}{(np_n)^\gamma}\Big) = o_P(1)\,.
\end{aligned}$$

But since

$$\left\{ \sqrt{\frac{p_n}{1-p_n}} \frac{B_n(s)}{1-s} : 0 \leqq s \leqq 1 - p_n \right\} \stackrel{\mathcal{D}}{=} \left\{ W\Big(\frac{p_n}{1-p_n} \frac{s}{1-s}\Big) : 0 \leqq s \leqq 1 - p_n \right\}$$

for each $n \in \mathbf{N}$, we also have

$$\sup_{0 \leqq s \leqq 1-p_n} \sqrt{\frac{p_n}{1-p_n}} \frac{B_n(s)}{1-s} \stackrel{\mathcal{D}}{=} \sup_{0 \leqq s \leqq 1-p_n} W\Big(\frac{p_n}{1-p_n} \frac{s}{1-s}\Big) = \sup_{0 \leqq t \leqq 1} W(t)$$

and

$$\sup_{0\leq s\leq 1-p_n}\sqrt{\frac{p_n}{1-p_n}}\frac{|B_n(s)|}{1-s}\overset{\mathcal{D}}{=}\sup_{0\leq s\leq 1-p_n}\left|W\left(\frac{p_n}{1-p_n}\frac{s}{1-s}\right)\right|=\sup_{0\leq t\leq 1}|W(t)|\,.$$

These imply (3) and (4). The proofs of (5) and (6) are completely analogous and are formally given in Section 4.5 of [4]. □

The proof of (9), (12), (15) and (18) in Theorem 1 requires the following

LEMMA. *If the sequence* $\{k_n\}_{n=1}^{\infty}$ *satisfies the conditions of Theorem 1, then*

$$\sup_{0\leq s\leq U_{n-k_n,n}}\left|\frac{1-s}{1-G_n(s)}-1\right|\xrightarrow{P}0 \quad \textit{and} \quad \sup_{U_{k_n,n}\leq s\leq 1}\left|\frac{s}{G_n(s)}-1\right|\xrightarrow{P}0\,.$$

PROOF. Let $\varepsilon\in(0,1]$ be fixed. Then by the simple idea in §3 of Rényi [13],

$$\begin{aligned}
&P\left\{\sup_{0\leq s\leq U_{n-k_n,n}}\left|\frac{1-s}{1-G_n(s)}-1\right|>\varepsilon\right\}\\
&\quad=1-P\left\{-\varepsilon\leq\frac{1-s}{1-G_n(s)}-1\leq\varepsilon,\quad 0\leq s\leq U_{n-k_n,n}\right\}\\
&\quad=1-P\left\{-\frac{\varepsilon}{1+\varepsilon}\leq\frac{1-G_n(s)}{1-s}-1\leq\frac{\varepsilon}{1-\varepsilon},\quad 0\leq s\leq U_{n-k_n,n}\right\}\\
&\quad\leq 1-P\left\{-\frac{\varepsilon}{2}\leq\frac{1-G_n(s)}{1-s}-1\leq\frac{\varepsilon}{2},\quad 0\leq s\leq U_{n-k_n,n}\right\}\\
&\quad=P\left\{\sup_{0\leq s\leq U_{n-k_n,n}}\left|\frac{1-G_n(s)}{1-s}-1\right|>\frac{\varepsilon}{2}\right\}\\
&\quad\leq P\left\{\sup_{0\leq s\leq 1-q\frac{k_n}{n}}\left|\frac{1-G_n(s)}{1-s}-1\right|>\frac{\varepsilon}{2}\right\}+P\left\{U_{n-k_n,n}>1-q\frac{k_n}{n}\right\},
\end{aligned}$$

and similarly,

$$P\left\{\sup_{U_{k_n,n}\leq s\leq 1}\left|\frac{s}{G_n(s)}-1\right|>\varepsilon\right\}\leq P\left\{\sup_{q\frac{k_n}{n}\leq s\leq 1}\left|\frac{G_n(s)}{s}-1\right|>\frac{\varepsilon}{2}\right\}+P\left\{U_{k_n,n}<q\frac{k_n}{n}\right\}$$

for every $q\in(0,1)$. It can be seen in several different ways that the second terms in the upper bounds go to zero (in fact, for each $q\in(0,1)$, they are not greater than $e^{-C_q k_n}$, for some constant $C_q>$ such that $\lim_{q\downarrow 0}C_q=\infty$: cf. (4.2) and its proof in [7]). The equivalent statements that the first terms in the bound go to zero for all $q\in(0,1)$ are well known. This was first shown by Chang Li-Chien in 1955; perhaps the easiest direct proof is in [16], Theorem 0. □

PROOF OF THEOREM 1. Put $p_n := k_n/n$, $n \geqq 1/p$. Then $p_n \leqq p$ for the $p \in (0,1)$ in the condition and $np_n \to \infty$, while $\sqrt{k_n}/n = \sqrt{p_n/n} \to 0$. What we have to show to prove (7), (8) and (11) is

$$\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq U_{n-k_n,n}} \frac{[\![\alpha_n(s)]\!]_l}{1-s} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} [\![W(t)]\!]_l, \quad l = 1,2,3, \tag{24}$$

where $[\![z]\!]_1 := z$, $[\![z]\!]_2 := -z$ and $[\![z]\!]_3 := |z|$ for $z \in \mathbf{R}$.

Exactly as in (23), for any fixed $\gamma \in (0, 1/4)$,

$$\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq U_{n-k_n,n}} \left| \frac{\alpha_n(s)}{1-s} - \frac{B_n(s)}{1-s} \right| \leqq \frac{1}{\sqrt{1-p}} \frac{p_n^{\frac{1}{2}}}{[1-U_{n-k_n,n}]^{\frac{1}{2}+\gamma}} \mathcal{O}_P\Big(\frac{1}{n^\gamma}\Big).$$

Since for independent random variables $Y_1, Y_2, \ldots$, each having the exponential distribution with mean 1, we have $1/[1-U_{n-k_n,n}] \stackrel{\mathcal{D}}{=} [Y_1 + \cdots + Y_{n+1}]/[Y_{n-k_n+1} + \cdots + Y_n]$ for each n, we see that $p_n^{\frac{1}{2}}/[1-U_{n-k_n,n}]^{\frac{1}{2}+\gamma} = \mathcal{O}_P(n^\gamma/k_n^\gamma)$, and so the whole upper bound is $\mathcal{O}_P(1/k_n^\gamma) = o_P(1)$. Hence, to prove (24), it suffices to show that

$$\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq U_{n-k_n,n}} \frac{[\![B_n(s)]\!]_l}{1-s} \xrightarrow{\mathcal{D}} \sup_{0 \leqq t \leqq 1} [\![W(t)]\!]_l, \quad l = 1,2,3, \tag{25}$$

for the special construction.

For every $y \geqq 0$ and $x > 0$, and for any of $l = 1,2,3$, we have

$$P\left\{\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq 1-p_n - x\sqrt{\frac{p_n}{n}}} \frac{[\![B_n(s)]\!]_l}{1-s} \leqq y\right\} - P\left\{U_{n-k_n,n} < 1 - \frac{k_n}{n} - x\frac{\sqrt{k_n}}{n}\right\}$$
$$\leqq P\left\{\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq U_{n-k_n,n}} \frac{[\![B_n(s)]\!]_l}{1-s} \leqq y\right\}$$
$$\leqq P\left\{\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq 1-p_n + x\sqrt{\frac{p_n}{n}}} \frac{[\![B_n(s)]\!]_l}{1-s} \leqq y\right\} + P\left\{U_{n-k_n,n} > 1 - \frac{k_n}{n} + x\frac{\sqrt{k_n}}{n}\right\}$$

for all n large enough that make $x \leqq \sqrt{n/p_n}$ and $x\sqrt{n/p_n} \leqq 1 - p_n$. Since

$$\sqrt{\frac{p_n}{1-p_n}} \sup_{0 \leqq s \leqq 1-p_n \pm x\sqrt{p_n/n}} \frac{[\![B_n(s)]\!]_l}{1-s} \stackrel{\mathcal{D}}{=} \sup_{0 \leqq s \leqq 1-p_n \pm x\sqrt{p_n/n}} \left[\!\!\left[W\Big(\frac{p_n}{1-p_n}\frac{s}{1-s}\Big)\right]\!\!\right]_l$$
$$= \sup\left\{ [\![W(t)]\!]_l : 0 \leqq t \leqq \frac{1 \pm \frac{x}{\sqrt{n}}\sqrt{\frac{p_n}{1-p_n}}}{1 \mp \frac{x}{\sqrt{np_n}}} \right\}$$
$$\to \sup\{[\![W(t)]\!]_l : 0 \leqq t \leqq 1\},$$

where the last convergence is almost sure by the sample continuity of $W(\cdot)$, and since the distribution function of $\sup_{0\le t\le 1}[\![W(t)]\!]_l$ is continuous for every $l=1,2,3$, the right-tail part of (22) implies that

$$P\Big\{\sup_{0\leqq t\leqq 1}[\![W(t)]\!]_l\leqq y\Big\}-\Phi(-x)\leqq\liminf_{n\to\infty}P\left\{\sqrt{\frac{p_n}{1-p_n}}\sup_{0\leqq s\leqq U_{n-k_n,n}}\frac{[\![B_n(s)]\!]_l}{1-s}\leqq y\right\}$$

$$\leqq\limsup_{n\to\infty}P\left\{\sqrt{\frac{p_n}{1-p_n}}\sup_{0\leqq s\leqq U_{n-k_n,n}}\frac{[\![B_n(s)]\!]_l}{1-s}\leqq y\right\}$$

$$\leqq P\Big\{\sup_{0\leqq t\leqq 1}[\![W(t)]\!]_l\leqq y\Big\}+\Phi(-x)$$

for every $y\geqq 0$ and $x>0$, $l=1,2,3$. Now (25) follows upon letting $x\to\infty$.

Thus (7), (8) and (11) are now proved. The Uniform(0, 1) versions of (9), (10) and (12), respectively equivalent to (9), (10) and (12) themselves, follow from (7), (8) and (11) combined with the first statement of the Lemma. That the six statements hold jointly is clear from the structure of the proof.

The proof of the left-tail versions (13), (14) and (17) is completely analogous, or mathematically equivalent, to that of (7), (8) and (11), while (15), (16) and (18) follow again from (13), (14) and (17) and the second statement of the Lemma. It is again obvious, then, that (13)–(18) hold jointly.

To prove the last statement concerning asymptotic independence in the case when $k_n/n\to 0$, for $j=0$ and $j=1$, set

$$\xi_{n,k_n}^{(j)}(s):=\sqrt{\frac{k_n}{1-\frac{k_n}{n}}}\,\frac{G_n(s)-s}{1-G_n^{(j)}(s)}\quad\text{and}\quad \eta_{n,k_n}^{(j)}(s):=\sqrt{\frac{k_n}{1-\frac{k_n}{n}}}\,\frac{G_n(s)-s}{G_n^{(j)}(s)},$$

where $G_n^{(0)}(s)=s$, $0\leqq s\leqq 1$, and $G_n^{(1)}(s)=G_n(s)$, $0\leqq s\leqq 1$, for all $n\geqq 1/p$. The six convergence relations

$$V_{n,k_n}^{(3j+l)}:=\sup_{0\leqq s\leqq U_{n-k_n,n}}[\![\xi_{n,k_n}^{(j)}(s)]\!]_l\xrightarrow{\mathcal{D}}\sup_{0\leqq t\leqq 1}[\![W(t)]\!]_l:=V_l\,,\ j=0,1;\ l=1,2,3,$$

holding jointly, and the six convergence relations

$$W_{n,k_n}^{(3j+l)}:=\sup_{U_{k_n,n}\leqq s\leqq 1}[\![\eta_{n,k_n}^{(j)}(s)]\!]_l\xrightarrow{\mathcal{D}}\sup_{0\leqq t\leqq 1}[\![W_*(t)]\!]_l:=V_l^*\,,\ j=0,1;\ l=1,2,3,$$

also holding jointly, represent (the Uniform(0, 1) versions of) the six statements in (7)–(12) and the six statements in (13)–(18), respectively.

Now let $\{m_n\}_{n=1}^{\infty}$ be a sequence of integers such that $1\leqq k_n<m_n<n$ for each $n\geqq 3$ and $m_n/n\to 0$, but $m_n/k_n\to\infty$. Introduce

$$V_{n,k_n,m_n}^{(3j+l)}:=\sup_{U_{n-m_n,n}\leqq s\leqq U_{n-k_n,n}}[\![\xi_{n,k_n}^{(j)}(s)]\!]_l\,,\quad j=0,1;\ l=1,2,3.$$

and

$$W^{(3j+l)}_{n,k_n,m_n} := \sup_{U_{k_n,n}\leq s\leq U_{m_n,n}} [\![\eta^{(j)}_{n,k_n}(s)]\!]_l\,, \quad j=0,1;\ l=1,2,3.$$

Then, for all the six cases from $j=0,1$ and $l=1,2,3$,

$$\left|V^{(3j+l)}_{n,k_n,m_n} - V^{(3j+l)}_{n,k_n}\right| \leq \sqrt{\frac{k_n}{m_n}}\,\frac{\sqrt{1-\frac{m_n}{n}}}{\sqrt{1-\frac{k_n}{n}}}\,V^{(3j+3)}_{n,m_n} = o(1)\mathcal{O}_P(1) = o_P(1) \tag{26}$$

by (11) and (12), or, what is the same, by the cases $j=0$ and $j=1$ coupled with $l=3$ in the first group of convergence relations above, applied with $\{m_n\}$ replacing $\{k_n\}$. Similarly,

$$W^{(3j+l)}_{n,k_n,m_n} - W^{(3j+l)}_{n,k_n} \xrightarrow{P} 0\,, \quad j=0,1;\ l=1,2,3.$$

Therefore, since

$$\mathbf{V}_{n,k_n} := \left(V^{(1)}_{n,k_n},\dots,V^{(6)}_{n,k_n}\right) \xrightarrow{\mathcal{D}} (V_1,V_2,V_3,V_1,V_2,V_3) =: \mathbf{V} \quad \text{in } \mathbf{R}^6$$

and

$$\mathbf{W}_{n,k_n} := \left(W^{(1)}_{n,k_n},\dots,W^{(6)}_{n,k_n}\right) \xrightarrow{\mathcal{D}} (V_1^*,V_2^*,V_3^*,V_1^*,V_2^*,V_3^*) =: \mathbf{V}_* \quad \text{in } \mathbf{R}^6,$$

and what we have to show is that the convergence in these last two relations is in fact joint, where $\mathbf{V} \stackrel{\mathcal{D}}{=} \mathbf{V}_*$ and $\mathbf{V}$ and $\mathbf{V}_*$ are independent, we also have

$$\mathbf{V}_{n,k_n,m_n} := \left(V^{(1)}_{n,k_n,m_n},\dots,V^{(6)}_{n,k_n,m_n}\right) \xrightarrow{\mathcal{D}} \mathbf{V} \quad \text{in } \mathbf{R}^6$$

and

$$\mathbf{W}_{n,k_n,m_n} := \left(W^{(1)}_{n,k_n,m_n},\dots,W^{(6)}_{n,k_n,m_n}\right) \xrightarrow{\mathcal{D}} \mathbf{V}_* \quad \text{in } \mathbf{R}^6.$$

Note that the vector $\mathbf{V}_{n,k_n,m_n}$ is a function only of the upper extreme order statistics $U_{n-m_n,n},\dots,U_{n,n}$ while the vector $\mathbf{W}_{n,k_n,m_n}$ is a function only of the lower extreme order statistics $U_{1,n},\dots,U_{m_n,n}$. Thus, since $m_n\to\infty$ and $m_n/n\to 0$, Satz 4 of Rossberg [15] implies that the random vectors $\mathbf{V}_{n,k_n,m_n}$ and $\mathbf{W}_{n,k_n,m_n}$ are asymptotically independent. But since we have already established that

$$\left(\mathbf{V}_{n,k_n} - \mathbf{V}_{n,k_n,m_n},\, \mathbf{W}_{n,k_n} - \mathbf{W}_{n,k_n,m_n}\right) \xrightarrow{P} (0,\dots,0) \in \mathbf{R}^{12},$$

the random vectors $\mathbf{V}_{n,k_n}$ and $\mathbf{W}_{n,k_n}$ are also asymptotically independent, that is, $(\mathbf{V}_{n,k_n},\mathbf{W}_{n,k_n}) \xrightarrow{\mathcal{D}} (\mathbf{V},\mathbf{V}_*)$ in $\mathbf{R}^{12}$. □

PROOF OF THEOREM 2. Setting

$$\zeta_{n,k_n}(x) := \sqrt{\frac{k_n}{1-\frac{k_n}{n}}}\,\frac{F_n(x)-F(x)}{1-F(x)} \quad\text{and}\quad \rho_{n,k_n}(x) := \sqrt{\frac{k_n}{1-\frac{k_n}{n}}}\,\frac{F_n(x)-F(x)}{1-F_n(x)}\,,$$

the left-hand side of (19) is

$$P\{\{-\zeta_{n,k_n}(x) \leqq z_\alpha,\ x \leqq X_{n-k_n,n}\} \cap \{\rho_{n,k_n}(x) \leqq z_\alpha,\ x \leqq X_{n-k_n,n}\}\}$$

$$= P\left\{\left\{-z_\alpha \leqq \inf_{x \leqq X_{n-k_n,n}} \zeta_{n,k_n}(x)\right\} \cap \left\{\sup_{x \leqq X_{n-k_n,n}} \rho_{n,k_n}(x) \leqq z_\alpha\right\}\right\}$$

$$\to P\left\{\left\{-z_\alpha \leqq \inf_{0 \leqq t \leqq 1} W(t)\right\} \cap \left\{\sup_{0 \leqq t \leqq 1} W(t) \leqq z_\alpha\right\}\right\}$$

$$= P\left\{\sup_{0 \leqq t \leqq 1} |W(t)| \leqq z_\alpha\right\} = L(z_\alpha) = 1 - \alpha$$

by a joint application of (8) and (9). The left-tail statement (20) follows from the joint validity of (13) and (16) in the same way.

Finally, introducing the events

$$A_{n,k_n}(x) := \left\{\frac{1 - F_n(x)}{c^*_{n,k_n}(\alpha)} \leqq 1 - F(x) \leqq c^*_{n,k_n}(\alpha)[1 - F_n(x)]\right\}$$

and

$$B_{n,k_n}(x) := \left\{\frac{F_n(x)}{c^*_{n,k_n}(\alpha)} \leqq F(x) \leqq c^*_{n,k_n}(\alpha) F_n(x)\right\},$$

the left-hand side of (21) is

$$p_n := P\{\{A_{n,k_n}(x),\ X_{k_n,n} \leqq x \leqq X_{n-k_n,n}\} \cap \{B_{n,k_n}(x),\ X_{k_n,n} \leqq x \leqq X_{n-k_n,n}\}\}$$

by the argument motivating the introduction of the band in question. Since

$$P\{A_{n,k_n}(x),\ x < X_{k_n,n}\} \to 1 \quad \text{and} \quad P\{B_{n,k_n}(x),\ x > X_{n-k_n,n}\} \to 1$$

by (26) and its left-tail analogue, using the proof of (19) and (20) above we see that

$$p_n \to P\left\{\sup_{0 \leqq t \leqq 1} |W(t)| \leqq z^*_\alpha,\ \sup_{0 \leqq t \leqq 1} |W_*(t)| \leqq z^*_\alpha\right\} = L^2(z^*_\alpha) = 1 - \alpha$$

by a joint application of the four statements in (8), (9), (13) and (16). □

REFERENCES

[1] Csáki, E., Vizsgálatok az empirikus eloszlásfüggvényről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **23** (1977), 239–327. *MR* **57** #4415. [English translation: Investigations concerning the empirical distribution function, *Selected Translations in Mathematical Statistics and Probability* **15**, American Mathematical Society, Providence, RI, 1981, 229–317.]

[2] Csörgő, M., Some Rényi type limit theorems for empirical distribution functions, *Ann. Math. Statist.* **36** (1965), 322–326. *MR* **31** # 4130

[3] CSÖRGŐ, M., A new proof of some results of Rényi and the asymptotic distribution of the range of his Kolmogorov–Smirnov type random variables, *Canad. J. Math.* **19** (1967), 550–558. *MR* **35** #3712

[4] CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D. M., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[5] CSÖRGŐ, M. and HORVÁTH, L., *Weighted approximations in probability and statistics*, Wiley, Chichester, 1993. *MR* **94c**:60060

[6] CSÖRGŐ, M. and RÉVÉSZ, P., *Strong approximations in probability and statistics*, Akadémiai Kiadó, Budapest, and Academic Press, New York, 1981. *MR* **84d**:60050

[7] CSÖRGŐ, S., Universal Gaussian approximations under random censorship, *Ann. Statist.* **24** (1996), 2744–2778. *MR* **97k**:62075

[8] CSÖRGŐ, S. and HORVÁTH, L., The baboons come down from the trees quite normally, *Mathematical statistics and applications* (Bad Tatzmannsdorf, 1983). Proceedings of the Fourth Pannonian Symposium on Mathematical Statistics, Vol. B (W. Grossman, G. Ch. Pflug, I. Vincze and W. Wertz, eds.), D. Reidel, Dordrecht, 1985, 95–106. *MR* **87d**:62002

[9] CSÖRGŐ, S. and HORVÁTH, L., Confidence bands from censored samples, *Canad. J. Statist.* **14** (1986), 131–144. *MR* **87i**:62089

[10] CSÖRGŐ, S. and MASON, D. M., Bootstrapping empirical functions, *Ann. Statist.* **17** (1989), 1447–1471. *MR* **91a**:62044

[11] MASON, D. M., The asymptotic distribution of generalized Rényi statistics, *Acta Sci. Math. (Szeged)* **48** (1985), 315–323. *MR* **87e**:60042

[12] RÉNYI, A., On the theory of order statistics, *Acta Math. Acad. Sci. Hungar.* **4** (1953), 191–231. *MR* **15** #885e. [Hungarian version: A rendezett minták elméletéről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **3** (1953), 467–503. *MR* **15** #544d]

[13] RÉNYI, A., A rendezett minták elméletének egy problémaköréről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **18** (1968), 23–30. *MR* **39** #3630. [English version: On some problems in the theory of order statistics, *Bull. Inst. Internat. Statist.* **42** (1969), 165–176. *MR* **41** #2859]

[14] RÉNYI, A., *Probability theory*, Akadémiai Kiadó, Budapest and North-Holland, Amsterdam, 1970. *MR* **47** #4296

[15] ROSSBERG, H.-J., Über das asymptotische Verhalten der Rand- und Zentralglieder einer Variationsreihe. II, *Publ. Math. Debrecen* **14** (1967), 83–90. *MR* **37** #1047

[16] WELLNER, J. A., Limit theorems for the ratio of the empirical distribution function to the true distribution function, *Z. Wahrsch. Verw. Gebiete* **45** (1978), 73–88. *MR* **58** #31356

(Received January 10, 1997)

DEPARTMENT OF STATISTICS
THE UNIVERSITY OF MICHIGAN
1440 MASON HALL
ANN ARBOR, MI 48109–1027
U.S.A.

scsorgo@umich.edu

and

SZEGEDI EGYETEM
BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
H–6720 SZEGED
HUNGARY

csorgo@math.u-szeged.hu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 89–106

RANDOM FRACTAL FUNCTIONAL LAWS OF THE ITERATED LOGARITHM

P. DEHEUVELS[1] and D. M. MASON[2]

Dedicated to the memory of Alfréd Rényi on the occasion of the 75th anniversary of his birth

1. Introduction and statement of main results

We shall establish random fractal versions of Chung-type functional laws of the iterated logarithm [FLIL] for the local oscillations of the Wiener process. In the process we will disclose a general scheme for evaluating the Hausdorff dimension of a large variety of random fractals which arise from local [FLIL].

Let $\{W(t): t \geqq 0\}$ be a standard Wiener process, and $(C_0[0,1], \mathcal{U})$ denote the set $C_0[0,1]$ of all continuous functions f on $[0,1]$ with $f(0)=0$, endowed with the uniform topology $\mathcal{U}$ generated by the sup-norm $\|f\| := \sup_{0 \leqq s \leqq 1} |f(s)|$.

For any $f \in C_0[0,1]$, we set

$$|f|_H = \begin{cases} \left(\int_0^1 g^2(s)ds\right)^2, & \text{when } f \text{ is absolutely continuous on } [0,1] \\ & \text{with } g = \dfrac{df}{ds}, \\ \infty, & \text{otherwise.} \end{cases}$$

Further, introduce the following subset of $C_0[0,1]$, called the *Strassen set of functions*,

$$\mathcal{S} = \{f \in C_0[0,1] : |f|_H \leqq 1\}.$$

For use later on set $\log_2 u = \log_+ \log_+ u$ with $\log_+ u = \log(u \vee e)$.

The Strassen [21] functional law of the iterated logarithm [FLIL] for W, formulated in this notation, asserts that, with probability 1,

$$(1.1) \qquad \lim_{T\to\infty} \inf_{f \in \mathcal{S}} \|(2T \log_2 T)^{-1/2} W(T\cdot) - f\| = 0,$$

1991 *Mathematics Subject Classification*. Primary 60F06, 60F15; Secondary 60F17, 60G15.

Key words and phrases. Empirical processes, fractals, strong laws, functional laws of the iterated logarithm, Wiener process, Brownian motion.

[1]Partially supported by a US-France NSF/CNRS Grant.

[2]Partially supported by an NSF and US-France NSF/CNRS Grant.

and, for each $f \in S$,

$$\liminf_{T\to\infty} \|(2T\log_2 T)^{-1/2}W(T\cdot) - f\| = 0. \tag{1.2}$$

We shall next discuss analogous FLIL for the increments of the Wiener process. Towards this end, we introduce for each $t \geqq 0$ and $h \geqq 0$, the increment function of $s \in [0, 1]$

$$\xi(h, t; s) = W(t + hs) - W(t). \tag{1.3}$$

Recall the Lévy modulus of continuity theorem which says that with probability one

$$\lim_{h\downarrow 0} \sup_{0\leqq t\leqq 1-h} \sup_{0\leqq s\leqq 1} (2h|\log(h)|)^{-1/2}|\xi(h, t; s)| = 1. \tag{1.4}$$

De Acosta [2] (see also Révész [19], Mueller [17] and Deheuvels and Lifshits [9]) proved a functional version of the Lévy modulus of continuity theorem. Namely he showed that as $h \downarrow 0$ the class of functions

$$\{(2h|\log(h)|)^{-1/2}\xi(h, t; \cdot) : 0 \leqq t \leqq 1 - h\} \tag{1.5}$$

converges with probability 1 in the Hausdorff metric to $\mathcal{S}$. Moreover, de Acosta [2] established that for $f \in \mathcal{S}$ satisfying $|f|_H < 1$, with probability 1,

$$\begin{aligned}\lim_{h\downarrow 0} \inf_{0\leqq t\leqq 1} |\log(h)| \times \|(2h|\log(h)|)^{-1/2}\xi(h, t; \cdot) - f\| \\ = 2^{-1/2}\gamma_{\mathcal{U}}^{1/2}\left(1 - |f|_H^2\right)^{-1/2},\end{aligned} \tag{1.6}$$

where $\gamma_{\mathcal{U}} = \pi^2/8$. Setting $f = 0$ in (1.6) yields an earlier result of Csörgő and Révész [7].

Now, for each $f \in \mathcal{S}$, let $\mathcal{D}_{\mathcal{U}}(f)$ denote the set of all $t \in [0, 1]$ such that

$$\liminf_{h\downarrow 0} \|(2h|\log(h)|)^{-1/2}\xi(h, t; \cdot) - f\| = 0. \tag{1.7}$$

Deheuvels and Mason [10] showed that, with probability 1, the set $\mathcal{D}_{\mathcal{U}}(f)$ is a random fractal with Hausdorff dimension

$$\dim \mathcal{D}_{\mathcal{U}}(f) = 1 - |f|_H^2. \tag{1.8}$$

Recall (see e.g. Falconer [12]) that the Hausdorff dimension of a subset E of $[0, 1]$ is defined by

$$\dim E = \inf\{c > 0 : s^c - \text{mes}\, E = 0\}, \tag{1.9}$$

where $s^c-\text{mes}\, E$ denotes the s^c-measure of E equal to

$$s^c - \text{mes}\, E = \liminf_{\epsilon\downarrow 0}\Big\{\sum_{j\in J}|I_j|^c : E \subseteqq \bigcup_{j\in J} I_j,\ |I_j| \leqq \epsilon,\ j \in J\Big\}, \tag{1.10}$$

In words: the infimum in (1.10) is taken over all collections $\{I_j : j \in J\}$ of closed intervals with lengths $|I_j| \leqq \epsilon$ for all $j \in J$, and such that $E \subseteq \bigcup_{j\in J} I_j$.

For each $f \in \mathcal{S}$ and $c > 1$, let $\mathcal{S}_{\mathcal{U}}(f, c)$ denote the set of all $t \in [0, 1]$ such that

$$\liminf_{h\downarrow 0} |\log(h)| \times \|(2h|\log(h)|)^{-1/2}\xi(h, t; \cdot) - f\| \leqq c\,2^{-1/2}\gamma_{\mathcal{U}}^{1/2}(1 - |f|_H^2)^{-1/2}. \tag{1.11}$$

Orey and Taylor [18] stated in their remarkable paper that in the particular case when $f = 0$,

$$\dim \mathcal{S}_{\mathcal{U}}(0, c) = 1 - c^{-2} \text{ a.s..}$$

(See e.g. (6.11) in their paper.) Our first main result is the following theorem, which determines the Hausdorff dimension of $\mathcal{S}_{\mathcal{U}}(f, c)$ for any $c > 1$ and $f \in \mathcal{S}$ satisfying $|f|_H < 1$.

THEOREM 1.1. *Let $f \in \mathcal{S}$ satisfy $|f|_H < 1$. Then, for any $c > 0$ with probability* 1,

$$\dim \mathcal{S}_{\mathcal{U}}(f, c) = (1 - |f|_H^2)(1 - c^{-2}). \tag{1.12}$$

In the process of proving Theorem 1.1 in Section 2, we shall develop a general scheme for establishing results like (1.12). Related results that can be readily obtained using this technique are described in Section 3.

2. Proof of Theorem 1.1 and related results

2.1. Preliminary facts and notation

We keep the notation of Section 1. The following facts will turn out to be essential to our proofs.

FACT 1. *For any $C > 0$, we have almost surely*

$$\lim_{a\downarrow 0} \sup_{0\leqq t\leqq C} \sup_{0\leqq u\leqq a} (2a|\log(a)|)^{-1/2}|W(t + u) - W(t)| = 1.$$

This is the Lévy [16] modulus of continuity theorem for the Wiener process (see e.g. Taylor [22]).

FACT 2. *For any $f \in \mathcal{S}$ with $|f|_H < 1$ and for any $r > 0$, we have*

$$\lim_{\lambda\to\infty} \lambda^{-2} \log P(\|\lambda^{-1}W - f\| \leqq \lambda^{-2}r) = -\frac{\pi^2}{8r^2} - \frac{1}{2}|f|_H^2.$$

This is Theorem 3.3 of de Acosta [1].

2.2. A subsequence argument

We shall make use of the following discretization scheme. Let $\gamma > 0$ be an arbitrary constant whose value will be chosen later on. For each $n \geqq 1$, set $h_n = n^{-\gamma}$. Denote by $\lfloor u \rfloor \leqq u < \lfloor u \rfloor + 1$ the integer part of u. We set for each $n \geqq 1$,

$$M_n := \lfloor 1/(h_n(|\log(h_n)|)^{-K}) \rfloor,$$

and for each $i = 0, 1, \ldots, M_n$,

$$s_{i,n} = ih_n(|\log(h_n)|)^{-K},$$

where $K \geqq 3$ is arbitrary but fixed. We note for further use that, for any $t \in [0,1]$, there exists an $i \in \{1, \ldots, M_n\}$ such that $|t - s_{i,n}| \leqq h_n(|\log(h_n)|)^{-K}$.

LEMMA 2.1. *We have almost surely*

$$\lim_{n\to\infty} h_{n+1}^{-1/2}(|\log(h_{n+1})|)^{1/2} U_n = 0, \tag{2.1}$$

where

$$U_n = \max_{1 \leqq i \leqq M_n} \sup_{|t - s_{i,n}| \leqq h_n(|\log(h_n)|)^{-K}} \sup_{h_{n+1} < h \leqq h_n} \|\xi(h, t; \cdot) - \xi(h_n, s_{i,n}; \cdot)\|.$$

PROOF. Recalling the definition (1.4) of $\xi(h, t; \cdot)$, we first observe, via the triangle inequality, that

$$U_n \leqq 2 \sup_{0 \leqq t \leqq 2} \sup_{0 \leqq u \leqq a'_n} |W(t+u) - W(t)|,$$

where $a'_n := h_n(|\log(h_n)|)^{-K} + (h_n - h_{n+1})$. Now our definition of $h_n = n^{-\gamma}$ ensures that as $n \to \infty$

$$h_n - h_{n+1} = (1 + o(1))\gamma n^{-\gamma-1} = o\left(h_n(|\log(h_n)|)^{-K}\right).$$

Therefore, by setting $a''_n = 2h_n(\log(1/h_n))^{-K}$ we see that for all n sufficiently large

$$U_n \leqq 2 \sup_{0 \leqq t \leqq 2} \sup_{0 \leqq u \leqq a''_n} |W(t+u) - W(t)|.$$

Applying Fact 1, taken with $C = 2$ and $a = a''_n$, we infer from this inequality and $K \geqq 3$ that with probability 1,

$$U_n = O\left((a''_n|\log(a''_n)|)^{1/2}\right) = o\left(h_n^{1/2}(|\log(h_n)|)^{-1/2}\right),$$

which gives (2.1). □

2.3. Upper bounds for the dimension of the exceptional set

The following lemma yields the upper half of (1.12). (This is the easy part of proving results like (1.12).) Its method of proof based on computing moments is readily adapted to other situations. See for instance the proof of Theorem 1.1 (i.e. (1.8) above) in [10].

LEMMA 2.2. *Let $f \in \mathcal{S}$ satisfy $|f|_H < 1$. Then, with probability 1, for any $c > 1$,*

$$\dim \mathcal{S}_{\mathcal{U}}(f, c) \leqq (1 - |f|_H^2)(1 - c^{-2}). \tag{2.2}$$

PROOF. Fix an arbitrary $\epsilon > 0$, and choose any $\gamma > 1/\epsilon$. Keeping the notation of Subsection 2.2, we define $Y_{i,n}$, for $i = 1, \dots, M_n$, to be 1 or 0 according as the random variable

$$|\log(h_n)| \times \|(2h_n|\log(h_n)|)^{-1/2}\xi(h_n, s_{i,n}; \cdot) - f\| \leqq c(1+\epsilon)(\gamma_{\mathcal{U}}/2)^{1/2}(1 - |f|_H^2)^{-1/2}$$

or not. Making use of Fact 2, we have uniformly over $i = 1, \dots, M_n$, as $n \to \infty$,

$$\begin{aligned} P(Y_{i,n} = 1) &= P(Y_{1,n} = 1) = P\left(\|\frac{W}{\lambda_n} - f\| \leqq \frac{cr}{\lambda_n^2}\right) \\ &= \exp\left(\log(h_n)\left\{1 - \left\{1 - \frac{1}{c^2(1+\epsilon)^2}\right\}(1 - |f|_H^2) + o(1)\right\}\right). \end{aligned} \tag{2.3}$$

Consider now the (possibly empty and at most countable) collection $\{I_j : j \in J\}$ of closed intervals of the form $[s_{i,n} - h_n, s_{i,n} + h_n]$ for which $Y_{i,n} = 1$, where $n \geqq 1$ and $1 \leqq i \leqq M_n$. Set

$$E = \cup\{I_j : j \in J\},$$

and

$$\delta = 2\epsilon + \left\{1 - \frac{1}{c^2(1+\epsilon)^2}\right\}(1 - |f|_H^2).$$

Introduce the (possibly infinite valued) random variable

$$Z = \sum_{j \in J} |I_j|^\delta,$$

where $\sum_{j \in \emptyset} (\cdot)$ is defined to be 0 whenever $J = \emptyset$. Obviously, we have

$$EZ = \sum_{n \geqq 1} M_n (2h_n)^\delta P(Y_{1,n} = 1) =: \sum_{n \geqq 1} u_n. \tag{2.4}$$

Note that, as $n \to \infty$,

$$M_n = \lfloor 1/(h_n(|\log(h_n)|)^{-K}) \rfloor = \exp((1+o(1))|\log(h_n)|).$$

Thus we infer from (2.3) and (2.4) that for all large n

$$u_n = \exp\left(\{2\epsilon + o(1)\}\log h_n\right) \leqq h_n^{\epsilon} = n^{-\epsilon\gamma}.$$

Since our choice of $\gamma > 1/\epsilon$ entails that $\sum_{n=1}^{\infty} u_n < \infty$, we see that $EZ < \infty$, which, in turn, implies that $Z < \infty$ with probability 1. In view of (1.9) and (1.10), it follows that, with probability 1, the measure $s^{\delta} - \dim E < \infty$, and hence

$$\dim E \leqq \delta = 2\epsilon + \left(1 - c^{-2}(1+\epsilon)^{-2}\right)(1 - |f|_H^2). \tag{2.5}$$

We finish by comparing the sets E and $\mathcal{S}_{\mathcal{U}}(f,c)$. By Lemma 2.1, there exists almost surely an $n_0 < \infty$ such that for all $n \geqq n_0$

$$h_{n+1}^{-1/2}(|\log(h_{n+1})|)^{1/2} U_n \leqq c\frac{\epsilon}{3}(\gamma_{\mathcal{U}}/2)^{1/2}(1 - |f|_H^2)^{-1/2}.$$

Hence, whenever for some $n \geqq n_0, h_{n+1} < h \leqq h_n$ and $0 \leqq t \leqq 1$, we have

$$\begin{aligned} &|\log(h)| \times \|(2h|\log(h)|)^{-1/2}\xi(h,t;\cdot) - f\| \\ &\leqq c(1 + \frac{\epsilon}{3})(\gamma_{\mathcal{U}}/2)^{1/2}(1 - |f|_H^2)^{-1/2}, \end{aligned} \tag{2.6}$$

then there exists an $i \in \{1, \ldots, M_n\}$ such that both $Y_{i,n} = 1$ and $t \in [s_{i,n} - h_n, s_{i,n} + h_n]$.

Since $\mathcal{S}_{\mathcal{U}}(f,c)$ is a subset of the set of all points t such that (2.6) holds for some $h_{n+1} < h \leqq h_n$ for infinitely many indexes n, it follows that, with probability 1, $\mathcal{S}_{\mathcal{U}}(f,c) \subseteqq E$, which, via (2.5), implies

$$\dim \mathcal{S}_{\mathcal{U}}(f,c) \leqq 2\epsilon + \left\{1 - \frac{1}{c^2(1+\epsilon)^2}\right\}(1 - |f|_H^2).$$

We conclude (2.2) by observing that $\epsilon > 0$ may be chosen arbitrarily small.□

2.4. A binomial scheme for the computation of fractal dimension

The following argument will be instrumental in our proof of the lower half of (1.12) and should be of independent interest. In formulating the theorem in this section we were strongly motivated by the arguments in Orey and Taylor [18]. We begin by introducing some notation.

Let $\{H_n : n \geqq 1\}$ denote a sequence of constants satisfying the following conditions (H1) and (H2).

(H1) $$H_n \downarrow,\ 0 < H_n < 1 \quad \text{for all large} \quad n \geqq 1.$$

(H2) $$\sum_{n \geq 1} \exp\left(-H_n^{-\epsilon}\right) < \infty \quad \text{for each} \quad \epsilon > 0.$$

Assume that, for each $n \geqq 1, Z_{i,n},\ i = 1, \ldots, N_n := \lfloor 1/H_n \rfloor$, is a sequence of independent and identically distributed Bernoulli random variables. Set

$$p_n = P\left(Z_{1,n} = 1\right) = 1 - P\left(Z_{1,n} = 0\right).$$

Further assume that, for some $0 < \delta < 1$, as $n \to \infty$

(H3) $$p_n = H_n^{\delta + o(1)}.$$

For each $n \geqq 1$, set $s_n(i) = iH_n$, $i = 1, \ldots, N_n$, and introduce the disjoint closed intervals

$$I_{i,n} = \begin{cases} [s_n(i) - H_n/2, s_n(i)] & \text{when } Z_{i,n} = 1, \\ \emptyset & \text{when } Z_{i,n} = 0. \end{cases} \tag{2.7}$$

Our main result in Section 2 is the following theorem.

THEOREM 2.1. *Under* (H1), (H2) *and* (H3), *for any* $\epsilon > 0$, *there exist almost surely a sequence of integers* $1 \leqq q_1 < q_2 < \ldots$, *and sets* E_1, $E_2, \ldots$, *such that*

$$\dim E \geqq 1 - \delta - \epsilon, \tag{2.8}$$

where $E = \bigcap_{j=1}^{\infty} E_j$ *and for each* $j \geqq 1$, E_j *is a union of some intervals taken from the set* $\left\{I_{i,q_j} : 1 \leqq i \leqq N_{q_j}\right\}$.

The proof of the theorem is derived from the forthcoming sequence of lemmas. First, we require some more notation.

Throughout this subsection I will denote a closed interval contained in $[0, 1]$. For any such interval I, let

$$\mathcal{N}_n(I) = \#\left\{I_{i,n} \subseteqq I : I_{i,n} \neq \emptyset,\ 1 \leqq i \leqq N_n\right\},$$

and

$$\mathcal{N}_n := \mathcal{N}_n([0,1]) = \sum_{i=1}^{N_n} Z_{i,n}.$$

The following lemmas establish some useful properties of $\mathcal{N}_n(I)$ and $\mathcal{N}_n$. Introduce the function

$$h(u) = \begin{cases} u \log u - u + 1 & \text{for } u > 0, \\ 1 & \text{for } u = 0, \\ \infty & \text{for } u < 0. \end{cases}$$

We will need the following probability inequality.

FACT 4. *Let S_N be a binomial $B(N,p)$ random variable. Then for all $r \geqq 1$*

$$P(S_N \geqq Nrp) \leqq \exp(-Nph(r)), \tag{2.9}$$

and for all $r \leqq 1$

$$P(S_N \leqq Nrp) \leqq \exp(-Nph(r)). \tag{2.10}$$

For a proof of this fact see Lemma 3.8 of [11].

LEMMA 2.3. *Under* (H1), (H2) *and* (H3), *we have almost surely as $n \to \infty$*

$$\lim_{n\to\infty} \mathcal{N}_n / \left\{H_n^{-1} p_n\right\} = 1, \tag{2.11}$$

and

$$\mathcal{N}_n = H_n^{\delta-1+o(1)}. \tag{2.12}$$

PROOF. Since (2.12) is a direct consequence of (2.11) and (H3), we need only show that the latter holds. Choose any $\epsilon > 0$. Replacing N, p and r in (2.9) by $N_n = \lfloor 1/H_n \rfloor$, p_n and $1+\epsilon$, respectively, shows, via (H1), (H2) and (H3), that

$$P\left(\mathcal{N}_n \geqq (1+\epsilon)N_n p_n\right) \leqq \exp\left(-N_n p_n h(1+\epsilon)\right) = \exp\left(-H_n^{\delta-1+o(1)}\right),$$

which by (H2) is summable in n. Therefore, the Borel–Cantelli lemma implies that almost surely,

$$\limsup_{n\to\infty} \mathcal{N}_n/(H_n^{-1}p_n) \leqq 1+\epsilon.$$

A similar argument based on (2.10), which we omit, shows likewise that almost surely,

$$\liminf_{n\to\infty} \mathcal{N}_n/(H_n^{-1}p_n) \geqq 1-\epsilon.$$

Since $\epsilon > 0$ may be chosen arbitrarily small, we readily infer (2.11) from the above two inequalities. □

The next lemma gives a refinement of the upper bound half of Lemma 2.3.

LEMMA 2.4. *For each $0 < \epsilon \leqq 1-\delta$ there exists almost surely an $n_1 < \infty$, such that, for all $n \geqq n_1$ and for all closed intervals $I \subseteq [0,1]$, we have*

$$\mathcal{N}_n(I)/(\mathcal{N}_n|I|^{1-\delta-\epsilon}) \leqq 1+\epsilon, \tag{2.13}$$

where $0/0 := 0$.

PROOF. Choose any $0 < \epsilon \leqq 1 - \delta$. For any integer $1 \leqq k \leqq N_n - 1$ denote by $C_n(k)$ the class of all closed intervals of the form

$$I = [s_n(i), s_n(i+k)] \text{ with } 1 \leqq i \leqq N_n - k.$$

We will first derive upper bounds for $\mathcal{N}_n(I)$, when for appropriate choices of k_n, the intervals I are in $C_n(k)$. Let $\epsilon_1 > 0$ be an arbitrary constant such that $\epsilon_1/2 + \delta < 1$. Set $k_n = \lfloor H_n^{-\delta-\epsilon_1/2} \rfloor$. Note that the total number of intervals $I \in C_n(k_n)$ is bounded above by $N_n = \lfloor 1/H_n \rfloor \leqq H_n^{-1}$. Also observe that when $I \in C_n(k_n)$, the random variable $\mathcal{N}_n(I)$ follows a binomial $B(k_n, p_n)$ distribution. Thus we infer from (2.9) and (H3) that

$$\begin{aligned} P&\left(\sup_{I \in C_n(k_n)} \mathcal{N}_n(I)/|I| \geqq \sqrt{1+\epsilon_1/2}\, p_n/H_n \right) \\ &\leqq H_n^{-1} \exp\left(-k_n p_n h(\sqrt{1+\epsilon_1/2})\right) \\ &= \exp\left(-H_n^{-\epsilon_1/2+o(1)}\right), \end{aligned}$$

which by (H2) is summable in n. Therefore, the Borel–Cantelli lemma shows that for all large enough n and $I \in C_n(k_n)$,

$$\mathcal{N}_n(I) \leqq \sqrt{1+\epsilon_1/2}\, |I| p_n/H_n. \tag{2.14}$$

To treat arbitrary intervals $I \subseteqq [0,1]$ we need to examine three cases.

Case 1. Consider any closed interval $I \subseteqq [0,1]$ such that $|I| \geqq H_n^{1-\delta-\epsilon_1}$.

Letting $k_n = \lfloor H_n^{-\delta-\epsilon_1/2} \rfloor$ be as above, we see that for all large n, I contains less than

$$\sqrt{1+\frac{\epsilon_1}{2}}\ \frac{|I|}{k_n H_n}$$

intervals belonging to $C_n(k_n)$. Thus, by (2.14), we see that with probability 1, for all n sufficiently large, we have uniformly over all intervals $I \subseteqq [0,1]$ with $|I| \geqq H_n^{1-\delta-\epsilon_1}$,

$$\mathcal{N}_n(I) \leqq (1+\frac{\epsilon_1}{2})|I|\{H_n^{-1} p_n\} \leqq (1+\epsilon_1)|I|\mathcal{N}_n, \tag{2.15}$$

where we have used (2.11). It follows from (2.15) that with probability 1, for all n sufficiently large and uniformly over all closed intervals $I \subseteqq [0,1]$ with $|I| \geqq H_n^{1-\delta-\epsilon_1}$,

$$\mathcal{N}_n(I)/\{\mathcal{N}_n |I|^{1-\delta-\epsilon}\} \leqq (1+\epsilon_1)|I|^{\delta+\epsilon} \leqq 1+\epsilon_1. \tag{2.16}$$

For the purpose of treating Case 2, stated below, consider now closed intervals I such that $I \in C_n(k)$ for some $1 \leqq k \leqq K_n := \lfloor H_n^{-\delta-\epsilon_1} \rfloor$. Observe that the total number of these intervals is bounded above by

$$H_n^{-1-\delta-\epsilon_1} \geqq K_n N_n. \tag{2.17}$$

Moreover, for any $I \in C_n(k)$ with $1 \leqq k \leqq K_n$ the random variable $\mathcal{N}_n(I)$ follows a binomial $B(k, p_n)$ distribution. Therefore keeping in mind that $|I| = kH_n$, we infer from (2.9) that for any $d > 0$,

$$Q(I) \leqq \exp\left(-kp_n h\left((1-\epsilon)d(kH_n)^{-\delta-\epsilon}\right)\right), \tag{2.18}$$

where

$$Q(I) := P\left(\mathcal{N}_n(I) \geqq kp_n\{(1-\epsilon)d(kH_n)^{-\delta-\epsilon}\}\right).$$

Next, we observe that our assumptions imply that, uniformly over $1 \leqq k \leqq k_n$, we have

$$kH_n \leqq K_n H_n = H_n^{1-\delta-\epsilon_1},$$

which converges to 0 as $n \to \infty$. Making use of the inequality holding for any $d > 0$ and all large enough x,

$$h(x) = (1 + o(1))\, x \log x \geqq \frac{x}{d(1-\epsilon)},$$

we infer from (H3) and (2.18) that for all large n and uniformly over $1 \leqq k \leqq K_n$

$$Q(I) \leqq \exp\left(-k^{1-\delta-\epsilon} p_n H_n^{-\delta-\epsilon}\right) \leqq \exp\left(-H_n^{-\epsilon/2}\right).$$

This inequality, in turn, implies that, for all large n,

$$P\left(\max_{1 \leqq k \leqq K_n} \sup_{I \in C_n(k)} \frac{\mathcal{N}_n(I)}{\{(1-\epsilon)H_n^{-1} p_n\}|I|^{1-\delta-\epsilon}} \geqq d\right)$$

$$\leqq H_n^{-1-\delta-\epsilon_1} \exp\left(-H_n^{-\epsilon/2}\right) \leqq \exp\left(-H_n^{-\epsilon/4}\right),$$

which, by (H2) is summable in n. Thus the Borel–Cantelli lemma and (2.11) implies that with probability 1 for all n sufficiently large, and all

$$I \in \bigcup_{1 \leqq k \leqq K_n} C_n(k),$$

we have

$$\mathcal{N}_n(I)/\{\mathcal{N}_n |I|^{1-\delta-\epsilon}\} \leqq d. \tag{2.19}$$

We note for further use that $d > 0$ in (2.19) is an arbitrary positive constant.

Case 2. We now consider an arbitrary $I \subseteqq [0,1]$ with $H_n/2 \leqq |I| < H_n^{1-\delta-\epsilon_1}$.

It is readily checked in this case that there always exists an $I' \in C_n(k)$ for some $1 \leqq k \leqq K_n = \lfloor H_n^{-\delta-\epsilon_1} \rfloor$ such that both $I \subseteqq I'$ and $|I'| \leqq 3|I|$. This, in conjunction with (2.19), shows that for all large n

$$\mathcal{N}_n(I)/\left\{\mathcal{N}_n|I|^{1-\delta-\epsilon}\right\} \leqq \mathcal{N}_n(I')/(\mathcal{N}_n|I'|^{1-\delta-\epsilon}3^{-1+\delta+\epsilon}) \leqq 3^{1-\delta-\epsilon}d. \tag{2.20}$$

Case 3. In the only remaining case when $|I| < H_n/2$, we have $\mathcal{N}_n(I) = 0$.

Thus putting this last case together with (2.16) and (2.20), observing that $\epsilon_1 > 0$ and $d > 0$ may be chosen arbitrarily small, we conclude (2.13). □

The following lemma gives a lower bound for $\mathcal{N}_n(I)$ when I is restricted to lie within an appropriate class.

LEMMA 2.5. *For any $\epsilon > 0$, we have almost surely*

$$\lim_{n\to\infty} \inf\left\{\mathcal{N}_n(I)/\mathcal{N}_n|I| : I \subseteqq [0,1], |I| \geqq H_n^{1-\delta-\epsilon}\right\} = 1. \tag{2.21}$$

PROOF. The proof is very similar to the just-given proof of Lemma 2.4 in Case 1 (see the argument from (2.14) to (2.15)). Therefore, we omit details. □

Finally, note that combining Lemmas 2.4 and 2.5 with (2.15), we get the following lemma.

LEMMA 2.6. *For all $0 < \tau \leqq 1$, we have almost surely*

$$\lim_{n\to\infty} \sup\left\{|\frac{\mathcal{N}_n(I)}{\mathcal{N}_n|I|} - 1| : I \subseteqq [0,1], |I| \geqq \tau\right\} = 0. \tag{2.22}$$

The next fact is a version of Lemma 2.2 of Orey and Taylor [18] stated in a manner appropriate for our needs.

FACT 5. *Let $K \subseteqq [0,1]$ be such that $K = \cap_{m=1}^{\infty} E_m$, where $E_1 \supseteqq \cdots \supseteqq E_m \supseteqq \cdots$ for $m = 1, 2, \ldots$ and $E_m = \bigcup_{k=1}^{M_m} J_{k,m}$, with $\{J_{k,m} : 1 \leqq k \leqq M_m\}$ being, for each $m \geqq 1$, a collection of disjoint closed nonempty subintervals of $[0,1]$ such that $\max_{1\leqq k\leqq M_n} |J_{k,m}| \to 0$ and $M_m \to \infty$ as $M \to \infty$. If there exist two constants $\triangle > 0$ and $d > 0$ such that, for every interval $I \subseteqq [0,1]$ with $|I| \leqq \triangle$ there is a constant $m(I)$ such that for all $m \geqq m(I)$,*

$$M_m(I) := \#\{J_{k,m} \subseteqq I : 1 \leqq k \leqq M_m\} \leqq d|I|^c M_m, \tag{2.23}$$

then we have $s^c - \dim\ K > 0$.

PROOF OF THEOREM 2.1. We have now in hand all the ingredients for the proof of Theorem 2.1. We claim that, for any $\epsilon > 0$, there exists with probability 1 a sequence $1 \leqq q_1 < q_2 < \cdots$ of integers, together with a sequence $E_1 \supseteqq E_2 \supseteqq \cdots$ of sets fulfilling the assumptions of Fact 5 with $c = 1 - \delta - \epsilon$, where, for each $j \geqq 1$, E_j is a disjoint union of closed nonempty intervals taken among $\{I_{i,q_j} : 1 \leqq i \leqq N_{q_j}\}$, as are defined in (2.7). Once these sets are constructed, the theorem follows directly from (2.23) using this in combination with definition (1.10). To show this one follows the same arguments as in [18], pp. 182–184 (also see [11], pp. 375–386). For completeness we include the Orey and Taylor arguments with some modifications and clarifications here.

Choose any $0 < \epsilon < 1 - \delta$ and apply Lemma 2.3 and Lemma 2.4 to find with probability 1 an $n_1 < \infty$ such that for all $n \geqq n_1$ and for all closed intervals $I \subseteqq [0,1]$, we have

$$\mathcal{N}_n(I) \leqq (1+\epsilon)\mathcal{N}_n |I|^{1-\delta-\epsilon/4}. \tag{2.24}$$

Now choose a decreasing sequence of positive constants $\{\epsilon_k\}_{k \geq 1}$ such that $0 < \epsilon_k < 1$ and $\sum_{k=1}^{\infty} \epsilon_k < \infty$ (which implies $0 < \prod_{k=1}^{\infty}(1-\epsilon_k) \leqq \prod_{k=1}^{\infty}(1+\epsilon_k) < \infty$). Next we apply Lemma 2.3 to find with probability 1, $q_1 \geqq n_1$ such that for all $n \geqq q_1$,

$$(1-\epsilon_1)H_n^{-1} p_n \leqq \mathcal{N}_n \leqq (1+\epsilon_1) N_n p_n. \tag{2.25}$$

For each $n \geqq 1$, let $\mathcal{I}_n$ denote those intervals among $\{I_{i,n} : 1 \leqq i \leqq N_n\}$ which are nonempty. Note that $\#(\mathcal{I}_n) = \mathcal{N}_n$. Define E_1 to be the union of those intervals in $\mathcal{I}_{q_1}$.

We shall now define an increasing sequence of integers $\{q_k\}_{k \geq 1}$ inductively beginning with q_1. Each E_k, for $k \geqq 2$, will be the union of those intervals in $\mathcal{I}_{q_k}$ that are subsets of E_{k-1}. For any integer q_k set

$$p_{q_k} = H_{q_k}^{\delta(k)}. \tag{2.26}$$

Notice that by (H3)

$$\delta(k) \to \delta, \text{ as } k \to \infty. \tag{2.27}$$

We set $\gamma(k) = 1 - \delta(k)$. Without loss of generality we can assume that $\epsilon < \gamma(k) < 1$, for all k. Denote the length of the intervals forming E_k by

$$\eta_k = \frac{H_{q_k}}{2}.$$

By Lemmas 2.3 and 2.6 for each $0 < \tau < 1$ and $0 < \beta < 1$ with probability one there exists an integer $m(\tau, \beta) < \infty$ such that for all $n \geqq m(\tau, \beta)$ and closed intervals $I \subseteqq [0,1]$ satisfying $|I| \geqq \tau$

$$(1-\beta)|I| H_n^{-1} p_n \leqq \mathcal{N}_n(I) \leqq (1+\beta)|I| N_n p_n. \tag{2.28}$$

Suppose that for $k \geqq 2$, the integers $q_1, \ldots, q_{k-1}$ have been defined. Select q_k large enough so that simultaneously

$$q_k > m(\epsilon, \eta_{k-1}^{2\delta(k)/\epsilon}) \vee m(\epsilon_k, \eta_{k-1}) \vee q_{k-1}; \tag{2.29}$$

$$\eta_k^{\epsilon/2} \leqq \left(\frac{1}{2}\right)^{2(1-\delta)} \prod_{i=1}^{k-1} \eta_i^{\delta(i)} \left(\frac{1}{2}\right)^{\gamma(i)}. \tag{2.30}$$

For integers $k \geqq 1$ let

$$M_k(I) = \# \{J_{i,k} \subseteqq I : 1 \leqq i \leqq M_k\}, \tag{2.31}$$

where $\{J_{i,k} : 1 \leqq i \leqq M_k\}$ denote the intervals which form E_k.

Suppose that for some $k \geqq 1$ and $j \geqq 1$

$$\eta_{k+j} < |I| \leqq \eta_{k+j-1}. \tag{2.32}$$

We claim that there exists a constant C independent of $k \geqq 1$ and $j \geqq 1$ such that when (2.32) holds

$$M_{k+j}(I) \leqq C|I|^{1-\delta-\epsilon/4} (2\eta_{k+j})^{-\gamma(i+k)} \prod_{i=1}^{j-1} \eta_{i+k}^{\delta(i+k)} \left(\frac{1}{2}\right)^{\gamma(i+k)}. \tag{2.33}$$

We shall verify (2.33) by induction. For $j = 1$, since $q_{k+1} > q_1 \geqq n_1$, we have by (2.24) that

$$M_{k+1}(I) \leqq \mathcal{N}_{q_{k+1}}(I) \leqq (1+\epsilon)\mathcal{N}_{q_{k+1}} |I|^{1-\delta-\epsilon/4},$$

which in turn by (2.25) is

$$\leqq (1+\epsilon)(1+\epsilon_1)|I|^{1-\delta-\epsilon/4} N_{q_{k+1}} p_{q_{k+1}} \leqq C_1 |I|^{1-\delta-\epsilon/4} (2\eta_{k+1})^{-\gamma(k+1)}.$$

Now if (2.33) is valid for some $j \geqq 1$ and C_j, we can apply the stipulation that $q_{k+j+1} > m(\epsilon_k, \eta_{k+j})$ along with (2.28) to each of the intervals $J_{i,k+j}$ that make up E_{k+j}. Keeping this in mind, along with (2.32), we have

$$M_{k+j+1}(I) = \sum_{\{i: J_{i,k+j} \subseteqq I\}} \mathcal{N}_{q_{k+j+1}}(J_{i,k+j}), \tag{2.34}$$

which, in turn, is

$$M_{k+j+1}(I) \leqq (1+\epsilon_{k+j}) \sum_{\{i: J_{i,k+j} \subseteqq I\}} N_{q_{k+j+1}} p_{q_{k+j+1}} |J_{i,k+j}|$$

$$\leqq M_{k+j}(I)h_{k+j}\,(2\eta_{k+j+1})^{-\gamma(k+j+1)}\,(1+\epsilon_{k+j}).$$

This establishes (2.33) with C_j replaced by $C_{j+1}=C_1(1+\epsilon_{k+j})$. Thus by induction, (2.33) holds for all k,j if we set $C=\prod_{i=1}^{\infty}(1+\epsilon_i)$.

A similar induction argument using $q_{k+j}\geqq m(\epsilon,\eta_{k+j-1}^{2\delta(k)/\epsilon})$ shows that

$$M_{k+j}(I)\leqq C|I|\,(2\eta_{k+j})^{-\gamma(k+j)}\prod_{i=1}^{j-1}\eta_{i+k}^{\delta(i+k)}\left(\frac{1}{2}\right)^{\gamma(i+k)} \tag{2.35}$$

for all intervals I satisfying for some $k\geqq 1$ and $j\geqq 1$

$$\eta_{k+j-1}^{2\delta(k)/\epsilon}<|I|\leqq\eta_{k+j-1}. \tag{2.36}$$

Notice that from (2.34) we get that

$$M_m=M_m([0,1])=\sum_{\{i:J_{i,m-1}\subseteq I\}}\mathcal{N}_m(J_{i,m-1}),$$

which by (2.28) is

$$\geqq M_{m-1}([0,1])\eta_{m-1}(2\eta_m)^{-\gamma(m)}(1-\epsilon_m)$$

$$\geqq M_1([0,1])\prod_{i=1}^{m-1}\eta_i(2\eta_{i+1})^{-\gamma(i+1)}(1-\epsilon_{i+1})$$

$$=\mathcal{N}_{q_1}([0,1])\prod_{i=1}^{m-1}\eta_i(2\eta_{i+1})^{-\gamma(i+1)}(1-\epsilon_{i+1}),$$

which by (2.25) is

$$\geqq(1-\epsilon_1)H_{q_1}^{-1}p_{q_1}\prod_{i=1}^{m-1}\eta_i(2\eta_{i+1})^{-\gamma(i+1)}(1-\epsilon_{i+1})$$

$$\geqq\eta_1^{\gamma(1)}(2\eta_m)^{-\gamma(m)}\prod_{i=1}^{m-1}\eta_i^{\delta(i)}\left(\frac{1}{2}\right)^{\gamma(i)}(1-\epsilon_i).$$

Thus we see that for some constant $D>0$ uniformly in $m\geqq 1$,

$$M_m\geqq D\,(2\eta_m)^{-\gamma(m)}\prod_{i=1}^{m-1}\eta_i^{\delta(i)}\left(\frac{1}{2}\right)^{\gamma(i)}. \tag{2.37}$$

We are now ready to finish the proof of Theorem 2.1. By Fact 5 it suffices to show that for the set

$$E = \bigcap_{k=1}^{\infty} E_k,$$

with probability 1 there exist constants δ_0, c_0 and m_0 such that

$$M_m(I) \leqq c_0 |I|^{1-\delta-\epsilon} M_m, \tag{2.38}$$

for all closed intervals $I \subseteq [0,1]$ with $|I| \leqq \delta_0$ and $m \geqq m_0$.

In order to deduce (2.38) from (2.33), (2.34) and (2.37) it is enough to prove that when $\eta_m^{2\delta(m)/\epsilon} < |I| \leqq \eta_m$,

$$|I| \leqq C|I|^{1-\delta-\epsilon} \prod_{i=1}^{m-1} \eta_i^{\delta(i)} \left(\frac{1}{2}\right)^{\gamma(i)}, \tag{2.39}$$

and when $\eta_{m+1} < |I| \leqq \eta_m^{2\delta(m)/\epsilon}$,

$$|I|^{1-\delta-\epsilon/4} \leqq C|I|^{1-\delta-\epsilon} \prod_{i=1}^{m-1} \eta_i^{\delta(i)} \left(\frac{1}{2}\right)^{\gamma(i)}. \tag{2.40}$$

Both of these inequalities hold by (2.30). Using (2.33), (2.35), (2.37), (2.39) and (2.40) it is straightforward to show that (2.38) holds. This finishes the proof of Theorem 2.1. □

2.5. Lower bounds for the dimension of the exceptional set

Armed with the results of the preceding subsections, we will now complete the proof of Theorem 1.1 by showing that, whenever $|f|_H < 1$, we have almost surely for each $c > 1$ and $\epsilon > 0$, chosen so that the right-hand side of the inequality below is strictly positive

$$\dim \mathcal{S}_{\mathcal{U}}(f,c) \geqq (1 - |f|_H^2)(1 - \frac{1}{c^2(1-\epsilon)^2}) - \epsilon. \tag{2.41}$$

Since $\epsilon > 0$ may be selected as small as desired, the proof of (1.12) will follow readily from (2.41) and the upper bound result (2.2) in Subsection 2.3.

To establish (2.41), we apply Theorem 2.1, with the following special choices of $\{H_n : n \geqq 1\}$ and $\{Z_{1,n} : 1 \leqq i \leqq N_n\}$ fulfilling (H1), (H2) and (H3). Choose a constant $\gamma > 0$ and set $h_n = n^{-\gamma}$ and $H_n = h_n(|\log(h_n)|)^{-3}$ for $n \geqq 1$. Now let $Z_{i,n} = 1$ or 0 according as the random variable

$$|\log(h_n)| \times \|(2h_n|\log(h_n)|)^{-1/2}\xi(h_n, s_{i,n}; \cdot) - f\|$$

$$\leqq c(1-\epsilon)(\gamma_{\mathcal{U}}/2)^{1/2}(1-|f|_H^2)^{-1/2}$$

or not.

Applying Fact 2, we get, similarly as for (2.3), that

$$P\left(Z_{i,n}=1\right)=$$

$$\exp\left(\log(h_n)\left(1-\left\{1-\frac{1}{c^2(1-\epsilon)^2}\right\}(1-|f|_H^2)+o(1)\right)\right)$$

$$=\exp\left(\log(H_n)\left\{1-\left\{1-\frac{1}{c^2(1-\epsilon)^2}\right\}(1-|f|_H^2)+o(1)\right\}\right),$$

which shows that (H3) holds with $\delta=1-\left\{1-\frac{1}{c^2(1-\epsilon)^2}\right\}(1-|f|_H^2)$. Notice that, in addition, the assumptions (H1) and (H2) hold trivially. Hence we may apply Theorem 2.1 to establish the existence of a set E such that

$$\dim E\geqq 1-\delta-\epsilon=\left\{1-\frac{1}{c^2(1-\epsilon)^2}\right\}(1-|f|_H^2)-\epsilon.$$

To conclude, we observe from the definition of $Z_{i,n}$ and Lemma 2.1, that with probability 1, for all large n, whenever $Z_{i,n}=1$ we have

$$|\log(h_n)|\times\|(2h_n|\log(h_n)|)^{-1/2}\xi(h_n,t;\cdot)-f\|\leqq c(\gamma_{\mathcal{U}}/2)^{1/2}(1-|f|_H^2)^{-1/2}$$

for all $t\in I_{i,n}=[s_n(i)-H_n/2,s_n(i)]$. This readily implies that $E\subseteqq\mathcal{S}_{\mathcal{U}}(f,c)$, which yields (2.24). This completes the proof of Theorem 1.1. □

3. Other applications of the general scheme outlined in Section 2

Making use of the methodology of Deheuvels and Lifshits [8], [9] in combination with the methods of de Acosta [1], the arguments of this paper can be readily adapted to treat norms $\|\cdot\|_\tau$ for which there exist positive constants κ_τ and γ_τ such that

$$\lim_{\lambda\to\infty}\lambda^{-\kappa_\tau}\log P\left(\|W\|_\tau\leqq\lambda^{-1}r\right)=-\frac{\gamma_\tau}{r^{\kappa_\tau}}.$$

Aside from the uniform norm, there are very few examples where this condition is known to be hold. Refer, for example, to Theorem 4.4 of Baldi and Roynette [3] in the case of the Hölder norm. Moreover, in most cases, the only information available concerns the rate of convergence to infinity as $\lambda\to\infty$ of

$$-\log P\left(\|W\|_\tau\leqq\lambda^{-1}r\right).$$

See, in particular, Stolz [20], Kuelbs and Li [14], Kuelbs, Li and Talagrand [15], and the references therein. Furthermore, the versions of the Lévy modulus of continuity theorem (see e.g. Fact 2) which must hold for such general norms are not known in the literature outside of special cases. Therefore to extend our fractal Chung-type FLIL to a more general setting requires additional results, which are beyond the scope of this paper.

The basic ingredients that are needed to prove a result like (1.12) (respectively like (1.7)) are a functional small ball result (respectively a functional large deviation result) combined with a modulus of continuity theorem like (1.4). In fact, the Wiener process W that appears in the definition of the increment process (1.3) and the definitions of the random fractals (1.6) and (1.11) can be replaced by certain separable Banach space valued Wiener processes. Then all of the arguments that yielded (1.7) and (1.12) carry over nearly verbatim. For the appropriate functional small ball, functional large deviation and modulus of continuity results consult de Acosta [1], [2]. It is little more than a matter of bookkeeping to translate our results to the general setting given there.

We conclude with a remark about the case when $|f|_H = 1$. Notice that (1.7) shows that when this condition holds, we have almost surely $\dim \mathcal{D}_{\mathcal{U}}(f) = 0$. Hence, in this case, there is no hope to obtain an appropriate definition of $\mathcal{S}_{\mathcal{U}}(f,c)$, since $\dim A = 0$ for all subsets A of $\mathcal{D}_{\mathcal{U}}(f)$.

ACKNOWLEDGEMENT. The authors thank Alex de Acosta for kindly pointing out to them his paper de Acosta [2].

REFERENCES

[1] DE ACOSTA, A., Small deviations in the functional central limit theorem with applications to functional laws of the iterated logarithm, *Ann. Probab.* **11** (1983), 78–101. *MR* **84m**:60038

[2] DE ACOSTA, A., On the functional form of Lévy's modulus of continuity for Brownian motion, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **69** (1985), 567–579. *MR* **86i**:60088

[3] BALDI, P. and ROYNETTE, B., Some exact equivalents for the Brownian motion in Hölder norm, *Probab. Theory Related Fields* **93** (1992), 457–484. *MR* **94a**: 60117

[4] 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

[5] CSÁKI, E., A relation between Chung's and Strassen's law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **54** (1980), 287–301. *MR* **82d**:60052

[6] CSÁKI, E., A lim inf result in Strassen's law of the iterated logarithm, *Limit theorems in probability and statistics* (Pécs, 1989), Colloq. Math. Soc. János Bolyai, Vol. 57, North-Holland, Amsterdam, 1990, 83–93. *MR* **92j**:60096

[7] CSÖRGŐ, M. and RÉVÉSZ, P., How small are the increments of a Wiener process?, *Stochastic Process. Appl.* **8** (1978), 119–129. *MR* **80a**:60031

[8] DEHEUVELS, P. and LIFSHITS, M. A., Strassen-type functional laws for strong topologies, *Probab. Theory Related Fields* **97** (1993), 151–167. *MR* **94i**:60048

[9] DEHEUVELS, P. and LIFSHITS, M. A., On the Hausdorff dimension of the set generated by exceptional oscillations of a Wiener process, *Studia Sci. Math. Hungar.* **33** (1997), 75–110.

[10] 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, M. B. Marcus, ed.), Progr. Probab., 35, Birkhäuser, Boston, 1994, 73–89. *MR* **96b**:60191

[11] DEHEUVELS, P. and MASON, D. M., On the fractal nature of empirical increments, *Ann. Probab.* **23** (1995), 355–387. *MR* **96e**:60064

[12] FALCONER, K. J., *Fractal geometry. Mathematical foundations and applications*, Wiley, New York, 1990, 355–387. *MR* **92j**:28008

[13] GRILL, K., A lim inf result in Strassen's law of the iterated logarithm, *Probab. Theory Related Fields* **89** (1991), 149–157. *MR* **92k**:60063

[14] KUELBS, J. and LI, W. V., Small ball estimates for Brownian motion and the Brownian sheet, *J. Theoret. Probab.* **6** (1993), 547–577. *MR* **94h**:60121

[15] KUELBS, J., LI, W. V. and TALAGRAND, M., Lim inf results for Gaussian samples and Chung's functional LIL, *Ann. Probab.* **22** (1994), 1879–1903. *MR* **96h**:60047

[16] LÉVY, P., *Processus stochastiques et mouvement brownien*, Gauthier-Villars, Paris, 1948. *MR* **8**, 551a

[17] MUELLER, C., A unification of Strassen's law and Lévy's modulus of continuity, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete.* **56** (1981), 163–179. *MR* **82f**:60180

[18] OREY, S. and TAYLOR, S. J., How often on a Brownian path does the law of iterated logarithm fail?, *Proc. London Math. Soc.* (3) **28** (1974), 174–192. *MR* **50** #11486

[19] RÉVÉSZ, P., A generalization of Strassen's functional law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **50** (1979), 257–264. *MR* **82c**:60057

[20] STOLZ, W., Une méthode élémentaire pour l'évaluation de petites boules browniennes, *C. R. Acad. Sci. Paris Sér. I Math.* **316** (1993), 1217–1220. *MR* **94a**:60121

[21] STRASSEN, V., An invariance principle for the law of the iterated logarithm, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **3** (1974), 211–226. *MR* **30** #5379

[22] TAYLOR, S. J., Regularity of irregularities on a Brownian path, *Ann. Inst. Fourier (Grenoble)* **24** (1974), 195–203. *MR* **53** #14699

(Received May 5, 1997)

LABORATOIRE DE STATISTIQUE THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
F–75252 PARIS Cedex 05
FRANCE

pd@ccr.jussieu.fr

DEPARTMENT OF MATHEMATICAL SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DE 19716
U.S.A.

davidm@math.udel.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 107–126

REFINED GIBBS CONDITIONING PRINCIPLE FOR CERTAIN INFINITE DIMENSIONAL STATISTICS

A. DEMBO[1] and J. KUELBS[2]

To the memory of Professor Alfréd Rényi

Abstract

Let $X_1, X_2, X_3, \ldots$ be independent, identically distributed random observations taking values in a Polish space Σ, and θ a statistic on Σ with values in a separable Banach space E. We examine the limit law of $(X_1, \ldots, X_k)$ conditional on $n^{-1}\Sigma_{i=1}^n \theta(X_i)$ being in an open convex subset D of E. In this setting the conditional limit law is a k-fold product probability $(P^*)^k$, where P^* is determined by the Gibbs conditioning principle. Our results describe the allowed dependence of $k = k(n)$ on n in terms of explicit geometric conditions related to smoothness of ∂D at a dominating point.

1. Introduction

Let $X, X_1, X_2, \ldots$ be independent, identically distributed random observations with empirical measure $L_n = \frac{1}{n}\Sigma_{j=1}^n \delta_{X_j}$, and common law P_X. In statistical mechanics, and also in a number of other settings, it is of interest to determine the limiting distribution of the k-tuple $(X_1, \ldots, X_k)$ provided one conditions on some observation of the empirical measure, say $T(L_n)$ being in some set D. It is intuitively clear that in a number of situations this conditional limit law should be a k-fold product measure $(P^*)^k$, but what is P^*? Of course, when $\lim_n P(T(L_n) \in D) = 1$, it is trivial that $P^* = P_X$, and hence the situation of greatest interest is when the conditional constraint $\{T(L_n) \in D\}$ is a "rare event". There are a variety of such results in the literature, and we mention [1], [4], [5], [7], and [12], which also include further background and references. The paper [11] examines some analogous results for discrete parameter Markov processes, and [13] and [14] are also related.

In [5], these results are described in terms of the "Gibbs conditioning principle", which beyond confirmation of the previous intuition also prescribes P^* via a variational principle. More precisely, the Gibbs conditioning principle claims that the limit law of $(X_1, \ldots, X_k)$ conditioned on the event

1991 *Mathematics Subject Classification.* Primary 60B12, 60F10.

Key words and phrases. Large deviation probabilities, Gibbs conditioning principle, relative entropy, conditional limit theorems, infinite dimensional statistics.

[1]Supported in part by NSF Grant DMS-9403553 and a US-ISRAEL BSF Grant.

[2]Supported in part by NSF Grants DMS-9400024 and DMS-9703740.

$\{T(L_n) \in D\}$ is a k-fold product measure $(P^*)^k$, where P^* minimizes the relative entropy with respect to P_X over all laws Q satisfying the constraint $T(Q) \in D$. This is indeed the case in the situations examined in [5], where a preliminary study of the relation between properties of the set D and the growth of $k = k(n)$ with n is carried out. In particular, the explicit results of [5] are mostly in case D is a subset of $\mathbb{R}^d$ for some finite d. Here we look at the extension of these results to infinite dimensional (Banach space valued) statistics.

To precisely describe our results, and their relationship to [5], we now fix some notation. Throughout the i.i.d. observations $\{X_j\}$ have values in a measure space $(\Sigma, \mathcal{B}_\Sigma)$, where Σ is a Polish space, $\mathcal{B}_\Sigma$ denotes the Borel subsets of Σ, and P_X is their common law. Let $M_1(\Sigma)$ denote the probability measures on $(\Sigma, \mathcal{B}_\Sigma)$, with the topology of weak convergence, and for any measure Q let Q^k denote the k-fold product of Q. If $P(L_n \in \Pi) > 0$ for some subset Π of $M_1(\Sigma)$ and $n \geqq k$, then $P^n_{X^k|\Pi}$ denotes the law of $(X_1, \dots, X_k)$ conditioned on $\{L_n \in \Pi\}$. The relative entropy of μ with respect to ν is given by

$$H(\mu|\nu) = \begin{cases} \displaystyle\int \log\left(\frac{d\mu}{d\nu}\right) d\mu & \text{if } \mu \ll \nu \\ +\infty & \text{otherwise.} \end{cases}$$

Let E denote a real separable Banach space with dual E^*, norm $\|\cdot\|$, and assume $\theta : \Sigma \to E$ is Borel measurable. Let Q_X be the law P_X induces on E through θ. Since Q_X is a probability on $(E, \mathcal{B}_E)$ there exist increasing, compact, convex sets $K_m (m \geqq 1)$ such that $Q_X(\cup_{m=1}^\infty K_m) = 1$, so we may and shall assume throughout that $\theta(s) \in \cup_{m=1}^\infty K_m$ for all $s \in \Sigma$ (modifying θ on a set of measure zero if needed). The statistic of the empirical measure we condition on is

$$T(L_n) = \int \theta dL_n = n^{-1} \sum_{j=1}^n \theta(X_j), \tag{1.1}$$

and our constraint that $\{T(L_n) \in D\}$ for a convex open $D \subset E$ is equivalent to $\{L_n \in \Pi'\}$ for

$$\Pi' = \left\{ \nu \in M_1(\Sigma) : \nu \circ \theta^{-1}(K_m) = 1 \text{ for some } m \geqq 1, \int_\Sigma \theta d\nu \in D \right\}. \tag{1.2}$$

Corollary 2.7 of [5] examines the situation when $E = \mathbb{R}^d$, D is convex and Q_X is lattice or strongly non-lattice. Furthermore, the collection $(X_1, \dots, X_k)$ is such that $k = k(n)$ may go to infinity as $n \to \infty$. The basic ingredients in the proof involve refinements of large deviation probabilities and Csiszár's information theoretic identity for blocks of length $k = k(n)$.

The goal here is to establish analogous results when $T(L_n)$ is an average of infinite dimensional statistics. Our moment assumption on Q_X is that

$$\int_E e^{t\|x\|} dQ_X(x) < \infty \tag{1.3}$$

for all $t > 0$, which is usual for the study of large deviation probabilities in the infinite dimensional setting, but stronger than what one expects to assume in $\mathbb{R}^d$. However, we do not assume anything about Q_X beyond that, and hence the assumptions Q_X being lattice, or strongly non-lattice, need not arise in $\mathbb{R}^d$ when (1.3) holds.

The usual rate function for Q_X is

$$\lambda(x) = \sup_{h \in E^*} [h(x) - \log \hat{Q}_X(h)] \qquad (x \in E), \tag{1.4}$$

where

$$\hat{Q}_X(h) = \int_E e^{h(x)} dQ_X(x) \qquad (h \in E^*).$$

Throughout, we assume that

$$D \subset E \;\text{ open, convex }\; \inf_{x \in D} \lambda(x) < \infty \quad m = \int_E x dQ_X(x) \notin D. \tag{1.5}$$

Assuming (1.3) and (1.5), by [9, Theorem 1], there exists a unique point $a_0 \in \partial D$ such that

$$\lambda(a_0) = \inf_{z \in \overline{D}} \lambda(z) < \lambda(x) \quad \forall x \in D. \tag{1.6}$$

This point a_0 is the so-called dominating point of (D, Q_X).

By the Hahn–Banach theorem, in this case, there exist $f \in E^*$ such that

$$\sup_{\{z : \lambda(z) \leq \lambda(a_0)\}} f(z) \quad = f(a_0) < f(x) \quad \forall x \in D. \tag{1.7}$$

Suppose in addition to (1.3) and (1.5) that $f \in E^*$ satisfies (1.7). Then, when $m \notin \overline{D}$, by [9, Lemma 2.6] for $g = t_0 f$ with a unique $t_0 > 0$,

$$\lambda(a_0) = g(a_0) - \log \hat{Q}_X(g). \tag{1.8}$$

In case $m \in \partial D$ we have $\lambda(a_0) \leqq \lambda(m) = 0$, hence $a_0 = m$ and (1.8) holds, now with $t_0 = 0$ and g the zero linear functional.

Thus, associated with the dominating point are $P^* \ll P_X$ on $(\Sigma, \mathcal{B}_\Sigma)$ such that for $g = t_0 f$ satisfying (1.8)

$$\frac{dP^*}{dP_X} = \exp\{\langle g, \theta(\cdot)\rangle - \log \hat{Q}_X(g)\}. \tag{1.9}$$

If both $g_1 \in E^*$ and $g_2 \in E^*$ satisfy (1.8), then considering $h = (g_1 + g_2)/2$ in (1.4), it follows by Hölder's inequality that $g_1 - g_2$ is constant a.s. Q_X, hence P^* of (1.9) is unique. With Q^* denoting the law that P^* induces on E through θ we also have by [9, Lemma 2.6] that

$$a_0 = \int_E x dQ^*(x) \tag{1.10}$$

(and (1.10) clearly holds when $m \in \partial D$ for then $P^* = P_X$, $Q^* = Q_X$ and $a_0 = m$).

DEFINITION. Assume (1.3) and (1.5). Let a_0 be the unique dominating point of (D, Q_X). Then, D contains slices whose diameters near a_0 dominate the function $\tau(s)$ if for some $f \in E^*$ satisfying (1.7) there exist $x_0 \in E$, $\delta > 0$, and $\beta > 0$ such that $f(x_0) > 0$, and

$$\{y + sx_0 : f(y) = 0, \|y\| < \beta\tau(s), 0 < s < \delta\} \subseteqq (D - a_0). \tag{1.11}$$

With (1.7) and (1.11) invariant to scaling of f we may and shall assume hereafter that $f(x_0) = 1$.

REMARK. Note that (1.11) holds for $\tau(s) = s$, any $f \in E^*$ and any $x_0 \in (D - a_0)$. Indeed, $B_{x_0,\beta} = \{x : \|x - x_0\| < \beta\} \subset (D - a_0)$ for $\beta > 0$ small enough, and with $0 \in \partial(D - a_0)$, by the convexity of $\|\cdot\|$ and of $(D - a_0)$ it follows that $B_{sx_0,\beta s} = sB_{x_0,\beta} \subset (D - a_0)$ for all $s \in (0, 1]$.

Our version of the Gibbs' conditioning principle for E-valued statistics is the following theorem.

THEOREM 1. *Assume* (1.3) *and* (1.5). *Let* $a_0 \in \partial D$ *be the unique dominating point for* (D, Q_X) *and* $T_n = \sum_{i=1}^n (Z_i - a_0)$ *for* Z_i *i.i.d.* E*-valued random vectors of common law* Q^*. *Suppose* $\{T_n/\sqrt{n}\}$ *is bounded in probability. Then, for* P^* *of* (1.9) *and* Π' *of* (1.2),

$$\lim_{n\to\infty} H(P^n_{X^{k(n)}|\Pi'} | (P^*)^{k(n)}) = 0, \tag{1.12}$$

provided one of the following holds:

(i) $k(n) = o((n/\log n)^{1/2})$.

(ii) $k(n) = O(n^{1/2})$ *and* $\{T_n/\sqrt{n}\}$ *has the CLT property in* E.

(iii) $k(n) = o((n/\log n)^{(1+\alpha)/2})$ *and* D *contains slices near* a_0 *whose diameters dominate the function* $\tau_\alpha(s) = s^{1/(1+\alpha)}$ *for some* $\alpha \in (0, 1]$.

(iv) $k(n) = o(n)$, $m \in \partial D$ *and either* $\{T_n/\sqrt{n}\}$ *has the CLT property or* D *contains slices near* a_0 *whose diameters dominate some* $\tau(s)$ *such that* $s^{-1}\tau(s) \to \infty$ *for* $s \downarrow 0$.

REMARKS.

(I) Recall that $P^n_{X^{k(n)}|\Pi'}$ stands for the law of $(X_1, \ldots, X_{k(n)})$ conditional on $\{L_n \in \Pi'\}$, where by (1.3), (1.5) and the infinite dimensional version of Cramér's theorem, $P(L_n \in \Pi') = P(n^{-1}\sum_{j=1}^n \theta(X_j) \in D) > 0$ for n sufficiently large.

(II) By Pinsker's inequality ($\|\mu - \nu\|_{\text{var}} \leqq (2H(\mu|\nu))^{1/2}$, cf. [2, Theorem 4.1]), the convergence of relative entropies in (1.12) implies convergence to zero of the total variation norms $\|P^n_{X^{k(n)}|\Pi'} - (P^*)^{k(n)}\|_{\text{var}}$.

(III) If D contains slices near a_0 whose diameters dominate $\tau_\alpha(s)$ for some $\alpha \in [1, \infty)$, then the same applies for $\alpha = 1$.

(IV) As shown in (2.23), the measure P^* of (1.9) satisfies the Gibbs conditioning principle (that is $H(P^*|P_X) = \inf_{P \in \Pi'} H(P|P_X)$).

(V) For E a type 2 Banach space, the assumption in Theorem 1 that $\{T_n/\sqrt{n}\}$ is bounded in probability and even the assumption of it having the CLT property follow directly from the moment assumptions on Q_X. However, if E is not of type 2, this need not be the case. Boundedness in probability of $\{T_n/\sqrt{n}\}$, i.e. $\sup_n P(\|T_n\| > r\sqrt{n}) \to 0$ as $r \to \infty$, is important as it allows the application of the Fuk–Nagaev inequality of [8] in our proof. Of course, if E is uniformly 2-smooth (see below), then E is already type 2, and the assumptions simplify accordingly.

(VI) From the proof of Theorem 1 we have that (1.12) holds for $k(n) = o(n)$ and $P^* = P_X$ as soon as $P(L_n \in \Pi')$ is bounded away from zero. By the law of large numbers for E-valued empirical means this is the case when $\int \|x\| dQ_X(x) < \infty$, $m \in D$ and $D \subset E$ is open and convex.

(VII) Part (iv) of Theorem 1 holds for $m \in \partial D$, $D \subset E$ open and convex, even when assumptions (1.3) and (1.5) are relaxed to either $\int \|x\|^3 dQ_X(x) < \infty$ and $f \in E^*$ of (1.11) such that $f(x_0) = 1$, $\int f(x)^2 dQ_X(x+m) > 0$ when $\{T_n/\sqrt{n}\}$ is only assumed bounded in probability, or that D intersects the convex hull of Q_X when assuming that $n^{-1/2}T_n$ has the CLT property. In particular, since $Q_X = Q^*$ and $a_0 = m$ with $\lambda(m) = 0$, the third moment assumption allows the immediate application of the Berry–Esseen and Fuk–Nagaev inequality to (2.10), as in (2.16) and (2.17).

(VIII) If Q_X is concentrated at the single point b, then either $b = m \in D$ or else $\inf_{x \in D} \lambda(x) = \infty$. Thus, (1.5) never holds for Q_X concentrated at a single point. However, in this case $P(L_n \in \Pi')$ is either zero or one with the conditioning on $\{L_n \in \Pi'\}$ of no interest.

Condition (1.11) is a geometric smoothness property of ∂D at a_0. Any open convex set D contains slices whose diameters dominate $\tau_0(s) = s$ at all possible dominating points. In certain Banach spaces, for some $\alpha \in (0,1]$, any open ball D also contains slices whose diameters dominate $\tau_\alpha(s)$ at all possible dominating points. For this we recall the following definition.

A Banach space $(E, \|\cdot\|)$ is called uniformly $(1+\alpha)$-smooth for some $\alpha \in (0,1]$ if for some $C < \infty$ and all $t \in [0, \infty)$

$$\sup_{\|x\|=1, \|y\|=1} \{\|x+ty\| + \|x-ty\| - 2\} \leqq C|t|^{1+\alpha}. \tag{1.13}$$

For example, if E is the standard L^p space for some $1 \leqq p < \infty$, then it is known that E is uniformly $(1+\alpha)$-smooth with $\alpha = \min(p-1, 1)$. There are no Banach spaces except $E = \{0\}$ which are uniformly $(1+\alpha)$-smooth for $\alpha > 1$, but an open set at a *particular* boundary point may contain slices whose diameters dominate $\tau_\alpha(s)$ for any $\alpha > 0$. Here is our theorem when $(E, \|\cdot\|)$ is uniformly $(1+\alpha)$-smooth.

THEOREM 2. *Suppose D is any non-empty open ball of the uniformly $(1+\alpha)$-smooth Banach space $(E, \|\cdot\|)$ for some $\alpha \in (0,1]$. Then, D contains slices whose diameters dominate $\tau_\alpha(s)$ near every possible dominating point $a_0 \in \partial D$.*

The following corollary is therefore immediate from part (i) of Theorem 1 and Theorem 2.

COROLLARY 1. *Assume in addition to* (1.3) *that a non-empty open ball D satisfies* (1.5) *and $\{T_n/\sqrt{n}\}$ of Theorem 1 is bounded in probability. Then,* (1.12) *holds for P^* of* (1.9) *provided $k(n) = o((n/\log n)^{1/2})$. If in addition $(E, \|\cdot\|)$ is uniformly $(1+\alpha)$-smooth for some $\alpha \in (0,1]$ then* (1.12) *holds even for $k(n) = o((n/\log n)^{(1+\alpha)/2})$ when $m \notin \overline{D}$, or $k(n) = o(n)$ when $m \in \overline{D}$.*

We next provide a partial converse of Theorem 1.

PROPOSITION 1. *Assume in the setting of Theorem 1, parts* (ii), (iii) *and* (iv), *that the characteristic function of $f(Z - a_0)$ is in $L^p(\mathbb{R})$ for some $p \in [1, \infty)$ and $f \in E^*$ satisfying* (1.7). *Then, $k(n) = o(n)$ is necessary for* (1.12) *to hold for P^* of* (1.9) *and Π' of* (1.2).

Much of the proof of Theorem 1 is inspired by a CLT type intuition. It is therefore interesting to examine in more detail the special case of Q_X a Gaussian measure. As we next show, in this setting one can typically remove the $\log n$ terms out of $k(n)$, leading to a tight characterization of the maximal $k(n)$ when D is smooth enough.

PROPOSITION 2. *Suppose* (1.5) *holds for a Gaussian measure Q_X. Then,* (1.12) *holds for P^* of* (1.9) *and Π' of* (1.2) *iff $k(n) = o(n)$ when either D is a non-empty open ball in a separable Hilbert space $(E, \|\cdot\|)$ or $(D - a_0)$ equals the left-side of* (1.11) *for $\tau(s) = s^{1/(1+\alpha)}$, $\alpha = 1$ and some x_0, $\beta > 0$, $\delta > 0$. In contrast,* (1.12) *fails for some $k(n) = o(n)$ in the latter case, whenever*

$m \notin \partial D$, $\alpha < 1$, and the support of Q_X is an infinite dimensional linear subspace of E.

Based on the above intuition, it seems that with more effort the $\log n$ terms might also be removed in the general (non Gaussian) case (see (5.13) for details). However, typically, at most $k(n) \ll o(n)$ in (1.12) when $m \notin \overline{D}$ and the diameters of slices of D near a_0 dominate only $\tau_\alpha(s)$ for some $\alpha < 1$.

2. Proof of Theorem 1

Let a_0 be the dominating point of (D, Q_X). The following lower bounds on $P(L_n \in \Pi')$ which are of some independent interest play a key role in the proof.

PROPOSITION 3. *Assume* (1.3), (1.5) *and that* $\{T_n/\sqrt{n}\}$ *is bounded in probability.*

(i) *For* $\alpha = 0$, *some* C_1 *finite and all* n *large enough,*

$$\log P(L_n \in \Pi') \geqq -n\lambda(a_0) - C_1(\log n)^{(1+\alpha)/2} n^{(1-\alpha)/2}. \tag{2.1}$$

(ii) *If* $\{T_n/\sqrt{n}\}$ *also has the CLT property in* E, *then one can remove the* $(\log n)^{1/2}$ *term in* (2.1) *and have* $C_1 > 0$ *arbitrarily small.*

(iii) *The bound* (2.1) *holds for* $\alpha \in (0, 1]$ *when* D *contains slices near* a_0 *whose diameters dominate the function* $\tau_\alpha(s)$.

(iv) *For* $m \in \partial D$, *if* $\{T_n/\sqrt{n}\}$ *has the CLT property or* D *contains slices near* a_0 *whose diameters dominate some* $\tau(s)$ *such that* $s^{-1}\tau(s) \to \infty$ *for* $s \downarrow 0$, *then*

$$\rho = \liminf_{n\to\infty} P(L_n \in \Pi') > 0. \tag{2.2}$$

PROOF. (i) There exists $x \in D$ and $\beta > 0$ such that $B_{x,3\beta} \subset D$ and (1.5) holds for D replaced by the open ball D_0 of radius β, centered at x. Let $a_0 + x_0$ denote the dominating point of (D_0, Q_X) as in (1.6) with P_0^* the measure associated with it via (1.9) and Q_0^* the measure P_0^* induces on E through θ. In particular, $B_{x_0,2\beta} \subset (D - a_0)$ and by the remark preceding Theorem 1,

$$\{z : \|z - a_0 - sx_0\| < 2\beta s,\ 0 < s < 1\} \subset D. \tag{2.3}$$

Let $V_r = \sum_{i=1}^r (Y_i - a_0 - x_0)$ for Y_i i.i.d. E-valued of common law Q_0^*. Then, $EY = \int x dQ_0^* = a_0 + x_0$ (compare with (1.10)). Moreover, $dQ_0^*/dQ_X \leqq ce^{c\|x\|}$ for some $c < \infty$ (see (1.9)). Hence, $E(\exp(t\|Y - a_0 - x_0\|)) < \infty$ for all t by (1.3), implying that for some $\eta > 0$ and all r large enough

$$P(\|V_r\| \geqq \beta r) \leqq e^{-\eta r} \tag{2.4}$$

(cf. [6, Exercise 6.2.21]). The duality identity of [6, (6.2.14)] for $h(x) = -nK1_{L_n \notin \Pi'} \in B(\Sigma^n)$ and $(P_X)^n \in M_1(\Sigma^n)$ results with

$$\begin{aligned}&\log[P(L_n \in \Pi') + e^{-nK} P(L_n \notin \Pi')] \\ = &\sup_{R \in M_1(\Sigma^n)} \{-KnR(L_n \notin \Pi') - H(R|(P_X)^n)\}.\end{aligned}$$

For $R = (P^*)^{n-r} \otimes (P_0^*)^r$, with $r < \delta n$ integer, and $N = H(P_0^*|P_X) - H(P^*|P_X) < \infty$ (see (2.20)), we have

$$\log[P(L_n \in \Pi') + e^{-nK}] \geqq -KnR(T(L_n) \notin D) - Nr - n\lambda(a_0). \tag{2.5}$$

By (2.3) for $s = r/n < \delta$,

$$\begin{aligned} R(T(L_n) \notin D) &= P\Big(n^{-1}(T_{n-r} + V_r) + a_0 + \frac{r}{n}x_0 \notin D\Big) \\ &\leqq P(\|T_{n-r} + V_r\| \geqq 2\beta r) \\ &\leqq P(\|T_{n-r}\| \geqq \beta r) + P(\|V_r\| \geqq \beta r). \end{aligned} \tag{2.6}$$

Note that $T_n = \sum_{i=1}^n (Z_i - a_0)$, with $Z_i - a_0$ i.i.d. of zero mean and exponential moments, in view of (1.10) and (1.3), respectively. The assumed boundedness in probability of $\{T_n/\sqrt{n}\}$ implies that $\sup_n E\|T_n/\sqrt{n}\| < \infty$ (see for example, [10, Proposition 2.3]). Hence, setting $r = r_n = [(An \log n)^{1/2}]$, by the Fuk–Nagaev inequality as given in [8, p. 338], we have for $\Lambda = E\|Z - a_0\|^2$ and all n large enough

$$P(\|T_{n-r_n}\| \geqq \beta r_n) \leqq r_n^{-3} O(n) + \exp\{-\beta^2 r_n^2/(96 n\Lambda)\}. \tag{2.7}$$

Taking $A > 96\Lambda/\beta^2$, by (2.4), (2.6) and (2.7) we see that $nR(T(L_n) \notin D) = o(r_n)$. Therefore, considering $K = \lambda(a_0) + 1$ in (2.5) we have $\log P(L_n \in \Pi') \geqq -n\lambda(a_0) - C_1 r_n$ for some $C_1 < \infty$ and all n large enough.

(ii) Subject to (1.3) and (1.5) holding, [9, Theorem 1] provides the representation

$$J_n = P(L_n \in \Pi') e^{n\lambda(a_0)} = E[e^{-t_0 f(T_n)} 1_{T_n \in n(D - a_0)}]. \tag{2.8}$$

In case $m \notin \overline{D}$, $t_0 > 0$ is specified by [9, Lemma 2.6] so that $g = t_0 f$ satisfies (1.8), while $t_0 = 0$ otherwise. Since $(D - a_0)$ is convex and $0 \in \partial(D - a_0)$, the open sets $\Gamma_b = b(D - a_0) \cap \{z : f(z) < b\}$ increase monotonically. In particular, (2.8) implies for all $n \geqq b^2$,

$$J_n \geqq e^{-t_0 n^{1/2} b} P(n^{-1/2} T_n \in \Gamma_b). \tag{2.9}$$

Recall that $n^{-1/2} T_n \to G$ weakly in $(E, \|\cdot\|)$ for $G = G(Z - a_0)$ an E-valued Gaussian variable with the same covariance structure as $(Z - a_0)$. The condition $\inf_{x \in D} \lambda(x) < \infty$ of (1.5) implies that D intersects the convex hull of the support of Q_X (see for example [6, (6.1.4) and Exercise 6.2.21]). Therefore, $(D - a_0)$ intersects the closed convex hull of the support of $Q^*(\cdot + a_0)$ and hence also the E-closure of the Hilbert space associated with G, denoted $\overline{H}_{\mathcal{L}(G)}$. Fix $z \in (D - a_0) \cap \overline{H}_{\mathcal{L}(G)}$. Since $f(z) > 0$, it follows that $tz \in \Gamma_b$ for any positive $t < b/(f(z) \vee 1)$. In particular, each of the open sets Γ_b also intersects $\overline{H}_{\mathcal{L}(G)}$. By the assumed CLT, for any $b > 0$,

$$\liminf_{n\to\infty} P(n^{-1/2}T_n \in \Gamma_b) \geqq P(G(Z - a_0) \in \Gamma_b) > 0,$$

and by (2.9) we get (2.1) for $\alpha = 0$ without the $(\log n)^{1/2}$ term and with $C_1 = (t_0 + 1)b$ arbitrarily small.

(iii) Suppose now that (1.11) holds for $\tau_\alpha(s)$, $\alpha \in (0, 1]$, and some $f \in E^*$ satisfying (1.7), $x_0 \in E$, $\delta > 0$, $\beta > 0$ (with $f(x_0) = 1$). Since $f(y) = 0$ for $y = (x - f(x)x_0)/t$ and every $t > 0$, $x \in E$, taking $s = f(x)/t$ it follows from (1.11) that

$$\{x : 0 < f(x) < t\delta, \|x - f(x)x_0\| < \beta t \tau_\alpha(f(x)/t)\} \subseteq t(D - a_0). \tag{2.10}$$

In particular, for $t = n$ and any $0 \leqq A_n \leqq B_n \lesssim n\delta$, by (2.8) and (2.10)

$$\begin{aligned} J_n &\geqq e^{-t_0 B_n} P(A_n < f(T_n) < B_n, \|T_n - f(T_n)x_0\| < \rho_n) \\ &\geqq e^{-t_0 B_n}[P(A_n < f(T_n) < B_n) - P(\|T_n - f(T_n)x_0\| \geqq \rho_n)], \end{aligned} \tag{2.11}$$

where $\rho_n = \beta n \tau_\alpha(A_n/n)$. Set $B_n = 2A_n$ and $A_n = (A \log n)^{(1+\alpha)/2} n^{(1-\alpha)/2}$ so that

$$\rho_n = \beta(An \log n)^{1/2}. \tag{2.12}$$

Note that $\int f(x)dQ^*(x+a_0) = 0$ (see (1.10)) and $\sigma_f^2 = \int f(x)^2 dQ^*(x+a_0) > 0$. Indeed, $\sigma_f^2 = 0$ implies $f(x) = f(a_0)$ a.s. Q_X, hence $\lambda(x) = \infty$ whenever $f(x) > f(a_0)$, in contradiction to (1.5) and (1.7).

Since $\sigma_f^2 > 0$, by the Berry–Esseen inequality, for some $C > 0$ and all n large enough

$$P(A_n < f(T_n) < B_n) \geqq C(B_n - A_n)/n^{1/2} \geqq 2n^{-\alpha/2}. \tag{2.13}$$

Applying the Fuk–Nagaev inequality (see (2.7)), we have for all n large enough

$$P(\|T_n - f(T_n)x_0\| \geqq \rho_n) \leqq \rho_n^{-3}O(n) + \exp\{-\rho_n^2/(96n\Lambda)\}, \tag{2.14}$$

where $\Lambda = E\|(Z - a_0) - f(Z - a_0)x_0\|^2$. Taking $A > 48\Lambda/\beta^2$ we have by (2.12) and (2.14) that

$$P(\|T_n - f(T_n)x_0\| \geqq \rho_n) = o(n^{-1/2}). \tag{2.15}$$

and hence as $n \to \infty$ combining (2.11), (2.13) and (2.15) we have

$$J_n \geqq \exp\{-2t_0(A \log n)^{(1+\alpha)/2} n^{(1-\alpha)/2}\} n^{-\alpha/2}.$$

Considering $\log J_n$, we establish (2.1) for $\alpha \in (0, 1]$, $C_1 > 2t_0 A^{(1+\alpha)/2} + 0.5$ and n large enough.

(iv) For $m \in \overline{D}$ we have $t_0 = 0$, $a_0 = m$, $\lambda(a_0) = 0$ and $Q^* = Q_X$. In this case we get (2.2) out of (2.9) when $\{T_n/\sqrt{n}\}$ has the CLT property. Otherwise, suppose D contains slices near a_0 whose diameters dominate some $\tau(s)$ such that $s^{-1}\tau(s) \to \infty$ for $s \downarrow 0$. Since $\sigma_f^2 > 0$, for $A_n = An^{1/2}$, $A > 0$ small, $B_n = Bn^{1/2}$, B large and all n large enough, the Berry–Esseen inequality implies

$$P(A_n < f(T_n) < B_n) \geqq 1/3. \tag{2.16}$$

Replacing $\tau_\alpha(\cdot)$ by $\tau(\cdot)$ in (2.10), we now have $\rho_n = \beta A_n \gamma_n$ in (2.11) for $\gamma_n = \inf\limits_{s < B_n/n} s^{-1}\tau(s) \uparrow \infty$. Thus, applying the Fuk–Nagaev inequality as in (2.14), for some $C = C(A) > 0$,

$$P(\|T_n - f(T_n)x_0\| \geqq \rho_n) \leqq Cn^{-1/2} + \exp(-C\gamma_n^2). \tag{2.17}$$

Since here $P(L_n \in \Pi') = J_n$ we establish (2.2) by combining (2.11), (2.16) and (2.17). □

The proof of (1.12) starts with the inequality

$$H(P^n_{X^{k(n)}|\Pi'} | Q^{k(n)}) \leqq [n/k(n)]^{-1} H(P^n_{X^n|\Pi'} | Q^n) \tag{2.18}$$

which holds for any probability measure Q and $k(n) \leqq n$ (cf. [5, (2.5)] for $n/k(n)$ integer). Considering in particular $Q = P^*$, the identity

$$H(P^n_{X^n|\Pi'} | (P_X)^n) = -\log P(L_n \in \Pi'), \tag{2.19}$$

yields the conclusions of Theorem 1 for $m \in \partial D$ and $P^* = P_X$, when combined with the lower bounds of Proposition 3 on $P(L_n \in \Pi')$.

Assuming hereafter that $m \notin \overline{D}$, by (1.8) – (1.10)

$$H(P^* | P_X) = H(Q^* | Q_X) = \lambda(a_0). \tag{2.20}$$

Hence, we have the conclusions of Theorem 1 by combining

$$H(P^n_{X^n|\Pi'} | (P^*)^n) \leqq -\log P(L_n \in \Pi') - nH(P^* | P_X), \tag{2.21}$$

with Proposition 3, (2.18) for $Q = P^*$, and (2.20).

To prove (2.21) we turn to Theorem 1 in [4]. Let K_m be the increasing, compact convex sets such that $\theta(s) \in \cup_m K_m$ for every $s \in \Sigma$. We first verify that all the conditions of [4, Theorem 1] apply for the convex set

$$\Pi = \left\{ \nu \in M_1(\Sigma) : \nu \circ \theta^{-1}(K_m) = 1 \text{ for some } m \geq 1, \int_\Sigma \theta d\nu \in \overline{D} \right\} \tag{2.22}$$

which contains Π' of (1.2). Indeed, by [4, Definition 2.3], Π being the union of the increasing completely convex sets $\Pi_m = \{\nu \in M_1(\Sigma) : \nu \circ \theta^{-1}(K_m) = 1, \int \theta d\nu \in \overline{D}\}$ is an almost completely convex set. By Remark (I), $P(L_n \in \Pi') > 0$ for n sufficiently large, whereas $\{(s_1, \ldots, s_n) : n^{-1} \sum_{i=1}^n \delta_{s_i} \in \Pi'\} = \{(s_1, \ldots, s_n) : n^{-1} \sum_{i=1}^n \theta(s_i) \in D\} \in (\mathcal{B}_\Sigma)^n$ for all n.

Recall that for Π a convex subset of $M_1(\Sigma)$, the generalized I-projection of P_X on Π is the unique element Q of $M_1(\Sigma)$ with $\lim_m \|P_m - Q\|_{\mathrm{var}} = 0$ for every $P_m \in \Pi$ such that

$$\lim_m H(P_m | P_X) = \inf_{P \in \Pi} H(P|P_X).$$

The existence and uniqueness of the generalized I-projection is given by [3, Theorem 2.1], which also shows that

$$H(Q|P_X) \leq \inf_{P \in \Pi} H(P|P_X).$$

In view of [4, Theorem 1], we now are able to conclude the proof by verifying that

LEMMA 2.1. *For $m \notin \overline{D}$ the probability measure P^* of (1.9) is the generalized I-projection of P_X on Π of (2.22). Moreover, for $\tilde{\Pi} = \{\nu \in M_1(\Sigma) : \int_\Sigma \theta d\nu \in \overline{D}\}$,*

$$\begin{aligned} \inf_{P \in \tilde{\Pi}} H(P|P_X) &= \inf_{P \in \Pi} H(P|P_X) = \inf_{P \in \Pi'} H(P|P_X) \\ &= H(P^*|P_X) = \lambda(a_0) < \infty. \end{aligned} \tag{2.23}$$

In the process of proving Lemma 2.1 we use the following simple relation.

LEMMA 2.2. *Let Q be a probability on $(E, \mathcal{B}_E)$ such that $\int_E h(x) dQ(x) = h(m_Q)$ for some $m_Q \in E$ and all $h \in E^*$. Then,*

$$H(Q|Q_X) \geq \lambda(m_Q). \tag{2.24}$$

PROOF. If $H(Q|Q_X) = \infty$ there is nothing to prove, so assume

$H(Q|Q_X) < \infty$ with $q = \dfrac{dQ}{dQ_X}$ and fix $h \in E^*$. Then,

$$
\begin{aligned}
H(Q|Q_X) &= \int_E q \log(q) dQ_X \\
&= h(m_Q) - \int_E q(x) I_{(0,\infty)}(q(x)) \log(q^{-1}(x) e^{h(x)}) dQ_X(x) \\
&\geqq h(m_Q) - \log \int_E e^{h(x)} dQ_X(x).
\end{aligned} \tag{2.25}
$$

The inequality (2.25) follows from Jensen's inequality since $\int_E qI_{(0,\infty)}(q)dQ_X = 1$. With $h \in E^*$ arbitrary, (1.4) implies (2.24). □

Proof of Lemma 2.1. Let $m_Q = \int_E x dQ(x)$ if $\int h(x)dQ(x) = h(m_Q)$ for every $h \in E^*$ (and otherwise $\int_E x dQ(x)$ undefined). In particular, $m_Q = x$ for $Q = \delta_x$ and any $x \in E$. Let $\Pi(C) = \{Q \in M_1(E) : \int_E x dQ(x) \in C\}$ and $\Pi_0(C) = \Pi(C) \cap \{Q \in M_1(E) : Q(K_m) = 1 \text{ for some } m \geqq 1\}$. The condition (1.5) implies in particular that D intersects the convex hull of the support of Q_X. Therefore, applying [4, Theorem 3] with $C = \overline{D}$ and noting that $\Pi_0(D) \subseteqq \Pi_0(\overline{D}) \subseteqq \Pi(\overline{D})$ we have that

$$\inf_{Q \in \Pi(\overline{D})} H(Q|Q_X) = \inf_{Q \in \Pi_0(\overline{D})} H(Q|Q_X) = \inf_{Q \in \Pi_0(D)} H(Q|Q_X) < \infty. \tag{2.26}$$

Then, by (2.26) and [4, Lemma 3.3]

$$\inf_{Q \in \Pi(\overline{D})} H(Q|Q_X) = \inf_{P \in \tilde{\Pi}} H(P|P_X) = \inf_{P \in \Pi} H(P|P_X) = \inf_{P \in \Pi'} H(P|P_X). \tag{2.27}$$

By (1.6) and Lemma 2.2 we have

$$\inf_{Q \in \Pi(\overline{D})} H(Q|Q_X) \geqq \inf_{x \in \overline{D}} \lambda(x) = \lambda(a_0). \tag{2.28}$$

Furthermore, $P^* \in \tilde{\Pi}$ since by (1.10)

$$\int_\Sigma \theta dP^* = \int_E x dQ^*(x) = a_0 \in \overline{D}.$$

Thus, we get (2.23) by combining (2.20), (2.27) and (2.28). In particular, $H(P^*|P_X) = \inf_{P \in \tilde{\Pi}} H(P|P_X) < \infty$ making $P^* \in \tilde{\Pi}$ the I-projection of P_X on $\tilde{\Pi}$.

Suppose $P_m \in \Pi \subseteqq \tilde{\Pi}$ is such that

$$\lim_m H(P_m|P_X) = \inf_{P \in \Pi} H(P|P_X).$$

Since P^* is the (generalized) I-projection of P_X on $\bar{\Pi}$, (2.27) implies that $\lim_m \|P_m - P^*\|_{\text{var}} = 0$. Consequently, by definition, P^* is also the generalized I-projection of P_X on Π. □

REMARK. The proof of Lemma 2.1 implies that P^* of (1.9) is also the generalized I-projection of P_X on Π' of (1.2).

3. Proof of Theorem 2

Since E is uniformly $(1+\alpha)$-smooth with respect to $\|\cdot\|$, by scaling (1.13), for all $x, y \in E$

$$\|x\|^\alpha[\|x+y\| + \|x-y\|] \leqq 2\|x\|^{1+\alpha} + C\|y\|^{1+\alpha}. \tag{3.1}$$

Let $a_0 \in \partial D$ be a possible dominating point for $D = \{x : \|x-a\| < R\}$ and $R > 0$. Set $x_0 = a - a_0$, with $D - a_0$ the open ball of radius $R = \|x_0\|$ centered at x_0. Then, $f \in E^*$ satisfying (1.7) is such that $(D - a_0) \subseteqq \{y : f(y) > 0\}$ and in particular $f(x_0) > 0$. Scale f so that $f(x_0) = 1$. Set $\delta = 1/2$ and $\beta = 0.5RC^{-1/(1+\alpha)}$ for C of (3.1).

Let y be such that $f(y) = 0$. Then, $f(-y) = 0$ so $-y \notin (D - a_0)$ implying $\|y + x_0\| \geqq R$. For $0 < s < \delta$ and $x = sx_0 - x_0$ we have

$$\|x-y\| = \|y-x\| \geqq \|y+x_0\| - \|sx_0\| \geqq (1-s)R = \|x\|. \tag{3.2}$$

In particular, for our choice of s, x and y, by (3.1) and (3.2)

$$\|x+y\| \leqq (1-s)R + 2C\|y\|^{1+\alpha}R^{-\alpha}.$$

By our choice of β, for $\|y\| < \beta\tau_\alpha(s)$ the above implies

$$\|y + sx_0 - x_0\| \leqq (1 - s + 2^{-\alpha}s)\|x_0\| < \|x_0\|,$$

and consequently,

$$\{y + sx_0 : f(y) = 0, \|y\| < \beta\tau_\alpha(s)\} \subseteqq (D - a_0). \tag{3.3}$$

Since each possible dominating point $a_0 \in \partial D$ satisfies (3.3) for $f \in E^*$ satisfying (1.7) and the above choices of x_0, β, δ, it follows that D contains slices whose diameters dominate $\tau_\alpha(s)$ near every possible dominating point $a_0 \in \partial D$. □

4. Proof of Proposition 1

For $m \notin \overline{D}$ this is a special case of [5, Proposition 2.12]. Indeed, by Lemma 2.1, P^* of (1.9) is then the generalized I-projection of P_X on the almost completely convex Π of (2.22) with $\log(dP^*/dP_X) - H(P^*|P_X) = g(\theta(\cdot) - a_0)$ by (1.8), (1.9) and (2.20). In particular, the latter random variable is in $L^2(P^*)$ by (1.3), and its characteristic function assumed in $L^p(\mathbb{R})$ for some $p \in [1, \infty)$. In the setting of parts (ii), (iii) and (iv) of Theorem 1, $n^{-1/2}|\log P(L_n \in \Pi') + n\lambda(a_0)| \to 0$ (see Proposition 3), which implies the lower bound of [5, (2.13)] and the conclusion follows from [5, (2.14)].

We next extend [5, (2.14)] to the case of $m \in \partial D$, that is, fixing $l = \epsilon^{-1}$ a positive integer and $n = lk$ we prove that $\|P^n_{X^k|\Pi'} - (P_X)^k\|_{\mathrm{var}}$ is bounded away from zero. To this end, the assumption that for $f \in E^*$ satisfying (1.7) the characteristic function of $f(Z - a_0)$ is in $L^p(\mathbb{R})$, some $p \in [1, \infty)$ allows for scaling such that $\int f(z - a_0)^2 dQ_X(z) = 1$. Moreover, with ϕ denoting the standard Normal density, as in [5, p.10] the conditional probability densities $p_n(y|v)$ of $Y_n = k^{-1/2} f(T_k)$ given $V_n = n^{-1/2} f(T_n)$ converge to $\psi_\epsilon(y, v) = \phi((y - \sqrt{\epsilon} v)/\sqrt{1-\epsilon})/\sqrt{1-\epsilon}$ uniformly in y and uniformly on compacts in v. For $h \colon \mathbb{R} \to [0, 1]$ monotone increasing, so is $v \mapsto \int h(y)\psi_\epsilon(y, v)dy$. Fix $h = 1_{[-1,\infty)}$ and $\eta > 0$ for which $(1 - \eta)\int h(y)\psi_\epsilon(y, 0)dy > \int h(y)\phi(y)dy$. By our assumptions, (2.2) holds, so we set $K < \infty$ such that $\int_K^\infty \phi(y)dy \leqq \rho\eta$ for $\rho > 0$ of (2.2). Then, by (2.2) and the CLT for V_n

$$\liminf_{n\to\infty} P(V_n \leqq K | L_n \in \Pi') \geqq 1 - \limsup_{n\to\infty} \frac{P(V_n \geqq K)}{P(L_n \in \Pi')} \geqq 1 - \eta.$$

Since $L_n \in \Pi'$ implies that $V_n > 0$, it follows that

$$\liminf_{n\to\infty} E(h(Y_n)|L_n \in \Pi') \geqq \liminf_{n\to\infty} P(V_n \leqq K | L_n \in \Pi') \inf_{0 \leqq v \leqq K} \int h(y) p_n(y|v) dy$$

$$\geqq (1 - \eta) \inf_{v \geqq 0} \int h(y)\psi_\epsilon(y, v)dy > \int h(y)\phi(y)dy. \tag{4.1}$$

The CLT for Y_n implies that $E(h(Y_n)) \to \int h(y)\phi(y)dy$. In view of (4.1), $\|\mathcal{L}(Y_n|L_n \in \Pi') - \mathcal{L}(Y_n)\|_{\mathrm{var}}$ is bounded away from zero, hence so is $\|P^n_{X^k|\Pi'} - (P_X)^k\|_{\mathrm{var}}$. □

5. Proof of Proposition 2

From Proposition 1 we know that $k(n) = o(n)$ is necessary for (1.12), where Q_X Gaussian implies the same for Q^* (see (1.9)) and in particular $f(Z - a_0)$ being Normal$(0, \sigma_f^2)$, some $\sigma_f > 0$, has a characteristic function in $L^p(\mathbb{R})$, all p.

In case $m \in \overline{D}$, the sufficiency of $k(n) = o(n)$ for (1.12) has already been shown. Assuming hereafter that $m \notin \overline{D}$ we rely on the following representation.

LEMMA 5.1. *For* Π' *of* (1.2), P^* *of* (1.9) *and* $1 \leqq k \leqq n$,

$$\frac{dP_{X^k|\Pi'}}{d(P^*)^k}(X_1,\dots,X_k) = \frac{E[h_n(T_n)|T_k]}{E[h_n(T_n)]}, \tag{5.1}$$

where $T_k = \sum_{i=1}^k (\theta(X_i) - a_0)$ *for* X_i *i.i.d. of common law* P^* *and* $h_n(x) = 1_{n(D-a_0)}(x)e^{-g(x)}$.

PROOF. By (1.9),

$$\frac{d(P_X)^n}{d(P^*)^n}(X_1,\dots,X_n) = e^{-n\lambda(a_0)}e^{-g(T_n)},$$

whereas by (2.8),

$$E[h_n(T_n)] = P(L_n \in \Pi')e^{n\lambda(a_0)},$$

so that for every $A \subseteqq \Sigma^k$ measurable,

$$\begin{aligned} E[h_n(T_n)]P_{X^k|\Pi'}(A) &= \int\limits_{A\times\Sigma^{n-k}} h_n(T_n)d(P^*)^n(X_1,\dots,X_n) \\ &= \int\limits_A E[h_n(T_n)|T_k]d(P^*)^k(X_1,\dots,X_k), \end{aligned}$$

out of which (5.1) follows. □

Since $Q^*(\cdot + a_0)$ is a centered Gaussian measure, we observe that $T_k = rtW'$, $T_n = r(tW' + \sqrt{1-t^2}W)$ for W, W' i.i.d. of law $Q^*(\cdot + a_0)$ where $r = \sqrt{n}$ and $t = \sqrt{k}/\sqrt{n}$. In particular, with $g \in E^*$, the law of $g(T_n)$ given T_k is Normal$(rtg(W'), r^2(1-t^2)\sigma_g^2)$ where $\sigma_g^2 = \int g(w)^2 dQ^*(w + a_0) > 0$.

For $v = \sigma_g^{-2} \int wg(w)dQ^*(w + a_0)$, clearly $W_o = W - g(W)v$ is independent of $g(W)$ with $g(v) = 1$ and $g(W_o) = 0$. For $u \geqq 0$ and $z \in E$ such that $g(z) = 0$ let

$$p_r(t,z,u) = P(r^{-1}uv + tz + \sqrt{1-t^2}W_o \in r(D - a_0)). \tag{5.2}$$

Then, with $W'_o = W' - g(W')v$ we have

$$E[h_n(T_n)|T_k] = \frac{N_r(t,W')}{\sigma_g r} \tag{5.3}$$

for

$$N_r(t,W') = (1-t^2)^{-1/2} \int_0^\infty e^{-u}\phi\left(\frac{r^{-1}u - tg(W')}{\sigma_g\sqrt{1-t^2}}\right) p_r(t,W'_o,u)du \tag{5.4}$$

(recall that $n(D-a_0) \subseteqq \{x : g(x) > 0\}$ and ϕ denotes the standard Normal density). Moreover, $E[\bar{h}_n(T_n)]$ is obtained by setting $t=0$ in (5.3) and (5.4) (in which case both $N_r(0,W')$ and $p_r(0,W'_o,u)$ are non-random).

Suppose next that for any $z \in E$ such that $g(z)=0$ the limit

$$h(u) = \lim_{t\to 0}\lim_{r\to\infty} p_r(t,z,u), \tag{5.5}$$

exists, does not depend on z and is such that $\int_0^\infty e^{-u}h(u)du = \eta > 0$. In particular, considering $t=0$ we have $N_r(0,W') = N_r(0) \to \eta$ (by bounded convergence). Moreover, $(2\pi(1-t^2))^{1/2}N_r(t,W') \le \sup_{u\geqq 0} p_r(t,W'_o,u) \leqq 1$, hence (5.5) implies that $N_r(t,W')/N_r(0) \to 1$ in L^2 as $r\to\infty$ followed by $t\to 0$. In particular, by Lemma 5.1,

$$H(P_{X^k|W'}|(P^*)^k) = E\left[\frac{N_r(t,W')}{N_r(0)}\log\left(\frac{N_r(t,W')}{N_r(0)}\right)\right] \to 0 \tag{5.6}$$

as $r=n^{1/2}\to\infty$ followed by $t=(k/n)^{1/2}\to 0$. Since $k\mapsto H(P_{X^k|W'}|(P^*)^k)$ is monotone non-decreasing (for n fixed), the convergence to zero in (5.6) implies that (1.12) holds for all $k=o(n)$.

We turn now to verify (5.5), first in case $(E,\|\cdot\|)$ is a Hilbert space and $(D-a_0) = \{x : \|x-x_0\| < \|x_0\|\}$. Since $g(x)>0$ for all $x\in(D-a_0)$, in particular, $\|y+x_0\| \geqq \|x_0\|$ when $g(y)=0$. Hence, x_0 is orthogonal to the closed, linear subspace $\{y : g(y)=0\}$ and for $c=g(x_0)^{-1}>0$, by the Pythagorean theorem,

$$\begin{aligned} p_r(t,z,u) = P(&\|r^{-1}u(v-cx_0)+tz+\sqrt{1-t^2}W_o\|^2 \\ &< [2cu-(cu/r)^2]\|x_0\|^2). \end{aligned} \tag{5.7}$$

The representation (5.7) implies that (5.5) holds for

$$h(u) = P(\|W_o\| < (2cu)^{1/2}\|x_0\|)$$

with $\eta = E(\exp(-\|W_o\|^2/(2c\|x_0\|^2))) > 0$.

Next suppose

$$(D-a_0) = \{y+sx_0 : f(y)=0, \|y\| < \beta s^{1/(1+\alpha)}, 0<s<\delta\} \tag{5.8}$$

for $\alpha=1$, some $\beta,\delta>0$ and f satisfying (1.7) such that $f(x_0)>0$. With no loss of generality take $f(\cdot)=g(\cdot)$ and assume $g(x_0)=1$. It is not hard to check that then, for $v'=v-x_0$ and z such that $g(z)=0$,

$$p_r(t,z,u) = P(\|r^{-1}uv'+tz+\sqrt{1-t^2}W_o\| < \beta u^{1/2})1_{(0,\delta r^2)}(u). \tag{5.9}$$

In case $W_o=0$ a.s. we get (5.5) for $h(u)=1_{u>0}$ with $\eta=1$. Otherwise, the continuity of $\rho\mapsto P(\|W_o\|\leqq\rho)$ yields (5.5) for $h(u)=P(\|W_o\|\leqq\beta u^{1/2})$ such that $\eta = E(\exp(-\|W_o\|^2/\beta^2)) > 0$.

Finally, we turn to show that (1.12) fails for some $k(n) = o(n)$ when $\alpha < 1$ in (5.8) and Q_X (hence also Q^*) is supported on an infinite dimensional subspace of E. First, let $Q_{\infty,t}$, $t \in [0,1)$ denote the mutually singular centered Gaussian laws that $Q^*(\cdot + a_0)$ induce on $\sqrt{1-t^2}(W - g(W)v)$. For the rest of the proof we forgo the change of measure, instead carrying out all computations for X_i i.i.d. P_X. In particular, for $U_n = g(S_n)$ and $S_n = n^{-1/2}\sum_{i=1}^n(\theta(X_i) - a_0)$ (with X_i i.i.d. P_X), note that $Q_{\infty,0}$ is also the law of $W_o = S_n - U_n v$ (independent of n) and that $U_n \sim N(-\mu_1 r, \mu_2)$ for some $\mu_1 > 0$, $\mu_2 > 0$, is independent of W_o. Moreover, with W_o' an i.i.d. copy of W_o, both independent of U_n, the law $Q_{r,t}$ that $P_{X^k|\Pi'}$ induces on $(S_k - U_k v)$ is the same as the law of $tW_o + \sqrt{1-t^2}W_o'$ conditional upon $\mathcal{A}_r = \{W_o + U_n v \in r(D - a_0)\}$.

We next show that the conditional law of W_o given $\mathcal{A}_r$ concentrates at 0 for $r \to \infty$. To this end, by the independence of U_n and W_o,

$$\begin{aligned} q_r(w) &= P(\mathcal{A}_r \mid W_o = w) = P(w + U_n v \in r(D - a_0)) \\ &= P(\|w + U_n v'\| < \beta(r^\alpha U_n)^{1/(1+\alpha)}, U_n < \delta r) \end{aligned}$$

(compare to (5.9)). If $\|w\| \geqq 4b$ then

$$q_r(w) \leqq P(U_n \geqq r^{-\alpha}(3b/\beta)^{1+\alpha}) + P(U_n \geqq b/\|v'\|),$$

whereas if $\|w\| \leqq b$ then

$$q_r(w) \geqq P(U_n \geqq r^{-\alpha}(2b/\beta)^{1+\alpha}) - P(U_n \geqq b/\|v'\|) - P(U_n \geqq \delta r).$$

Since $P(U_n \geqq a_r)/P(U_n \geqq c_r) \to 0$ for every $a_r, c_r \geqq 0$ such that $r(a_r - c_r) \to \infty$, it follows that for any $b > 0$,

$$\lim_{r\to\infty} \frac{\sup_{\|w\|\geqq 4b} q_r(w)}{\inf_{\|w\|\leqq b} q_r(w)} = 0. \tag{5.10}$$

For every $b > 0$, both $P(\|W_o\| \leqq b) > 0$ and

$$P(\|W_o\| \geqq 4b|\mathcal{A}_r) \leqq \frac{P(\mathcal{A}_r \mid \|W_o\| \geqq 4b)}{P(\mathcal{A}_r \mid \|W_o\| \leqq b)} \frac{}{P(\|W_o\| \leqq b)}.$$

Hence, by (5.10) also

$$P(\|W_o\| > b|\mathcal{A}_r) = P(\|W_o\| > b|W_o + U_n v \in r(D - a_0)) \to 0 \quad \forall b > 0. \tag{5.11}$$

This in turn implies that $Q_{r,t} \to Q_{\infty,t}$ with respect to the $C_b(E)$-topology on $M_1(E)$ for any fixed $t \in (0,1)$. In particular, by the lower semi-continuity of $H(\cdot|Q_{\infty,0})$ and mutual singularity of $\{Q_{\infty,t}\}$ we have

$$\liminf_{r\to\infty} H(Q_{r,t}|Q_{\infty,0}) \geqq H(Q_{\infty,t}|Q_{\infty,0}) = \infty. \tag{5.12}$$

Since $H(Q_{r,t}|Q_{\infty,0})$ is the relative entropy between the measures induced on $(S_k - U_k v)$ by $P_{X^k|\Pi'}$ and by $(P^*)^k$, it follows from (5.12) that

$$\lim_n H(P_{X^k|\Pi'}|(P^*)^k) = \infty$$

for $k = nt^2$, and arbitrary fixed $t > 0$. Of course, (1.12) then fails for some $k(n) = o(n)$. □

REMARKS.

(I) By the above proof, (1.12) fails for some $k(n) = o(n)$ in any set D for which (5.11) holds.

(II) Suppose Q_X is non-Gaussian, the characteristic function of $g(Z - a_0)$ is in $L^p(\mathbb{R})$ for some $p < \infty$ and (5.5) holds for

$$(5.13) \qquad p_r(t,z,u) = P(r^{-1}T_{r^2} \in r(D - a_0)|r^{-1}T_{t^2r^2} = tz, g(T_{r^2}) = u)$$

with $\int_0^\infty e^{-u}h(u)du > 0$. Assume moreover that $p_r(0,u) = P(r^{-1}T_{r^2} \in r(D - a_0)|g(T_{r^2}) = u) \to h(u)$ when $r \to \infty$. Then, (5.3) and (5.4) hold, now with $W_o' = W'$ and ϕ denoting the density of $(n-k)^{-1/2}g(T_n - T_k)/\sigma_g$. Since these densities converge uniformly to the standard Normal density (see [5, p. 10]), similarly to the above proof we again have (1.12) holding for $k(n) = o(n)$.

(III) Suppose $Q^*(\cdot + a_0)$ is the standard Gaussian measure on $E = \mathbb{R}^d$, $d \geq 3$, equipped with the Euclidean norm. Consider the open cone $(D - a_0) = \{y + sx_0 : g(y) = 0, \|y\| < s\}$ corresponding to $\alpha = 0$ in (1.11), with $g(\cdot)$ the 1-st coordinate projection and x_0 the associated (1-st coordinate) unit vector. Then, $\sigma_g = 1$, $v = x_0$ with W_o a standard Gaussian variable on the $(d-1)$-dimensional linear subspace $\{y : g(y) = 0\}$. Here we get $p_r(t,z,u) = P(\|tz + \sqrt{1-t^2}W_o\| < r^{-1}u)$ (compare with (5.9)), and (5.5) is no longer useful. Nevertheless, for some $C_{d-1} > 0$

$$\begin{aligned}\lim_{t\to 0}\lim_{r\to\infty} r^{d-1}p_r(t,z,u) &= \lim_{t\to 0} C_{d-1}(u/\sqrt{1-t^2})^{d-1}\exp(-0.5\|tz\|^2/(1-t^2)) \\ &= C_{d-1}u^{d-1}.\end{aligned}$$

Hence, we obtain (5.6) and thus (1.12) holds for all $k = o(n)$. Note that here

$$n^{d/2}P(L_n \in \Pi')e^{n\lambda(a_0)} = \frac{r^{d-1}N_r(0)}{\sqrt{2\pi}} \to \frac{C_{d-1}}{\sqrt{2\pi}}\int_0^\infty e^{-u}u^{d-1}du.$$

So, in this example the conclusion of [5, Proposition 2.15] holds although condition (2.16) of [5] fails.

(IV) Suppose Q_X is Gaussian and D satisfying (1.5) contains slices whose diameters near a_0 dominate $\tau_\alpha(s)$ for some $\alpha \in (0,1)$. Then, for $\delta > 0$ small enough and all $u \in (0, \delta r^2)$

$$\begin{aligned} p_r(0,z,u) &\geqq P(\|r^{-1}uv' + W_o\| < \beta(r^{\alpha-1}u)^{1/(1+\alpha)}) \\ &\geqq P(\|W_o\| < 0.5\beta(r^{\alpha-1}u)^{1/(1+\alpha)}) \end{aligned}$$

(compare to (5.9)). In particular (see (5.4)), considering $br^{1-\alpha} \leqq u \leqq 2br^{1-\alpha}$, for any $b \in (0,1)$ and r large enough

$$\begin{aligned} P(L_n \in \Pi')e^{n\lambda(a_0)} &= \frac{1}{\sigma_g r}\int_0^\infty e^{-u}\phi(\frac{u}{\sigma_g r})p_r(0,z,u)du \\ &\geqq b\exp(-3r^{1-\alpha}b)P(\|W_o\| < 0.5\beta b), \end{aligned}$$

implying that (2.1) holds without the $\log n$ terms and with $C_1 > 0$ arbitrarily small. As in the proof of Theorem 1, it follows that $k(n) = O(n^{(1+\alpha)/2})$ suffices for (1.12) to hold. In particular, this is the case for non-empty open balls in a uniformly $(1+\alpha)$-smooth $(E, \|\cdot\|)$.

REFERENCES

[1] Bolthausen, E., Maximum entropy principles for Markov processes, *Stochastic processes and their applications in mathematics and physics* (Bielefeld, 1985), Math. Appl., 61, Kluwer Acad. Publ., Dordrecht, 1990. *MR* **91m**:60049

[2] Csiszár, I., Information-type measure of difference of probability distributions and indirect observations, *Studia Sci. Math. Hungar.* **2** (1967), 299–318. *MR* **36** #2428

[3] Csiszár, I., *I*-divergence geometry of probability distributions and minimization problems, *Ann. Probab.* **3** (1975), 146–158. *MR* **51** #2050

[4] Csiszár, I., Sanov property, generalized *I*-projection and a conditional limit theorem, *Ann. Probab.* **12** (1984), 768–793. *MR* **86h**:60065

[5] Dembo, A. and Zeitouni, O., Refinements of the Gibbs conditioning principle, *Probab. Theory Related Fields* **104** (1996), 1–14. *MR* **97k**:60078

[6] Dembo, A. and Zeitouni, O., *Large deviations techniques and applications*, Jones and Bartlett Publishers, Boston, MA, 1993. *MR* **95a**:60034

[7] Diaconis, P. and Freedman, D., Conditional limit theorems for exponential families and finite versions of de Finetti's theorem, *J. Theoret. Probab.* **1** (1988), 381–410. *MR* **89g**:60127

[8] Einmahl, U., Stability results and strong invariance principles for partial sums of Banach space valued random variables, *Ann. Probab.* **17** (1989), 333–352. *MR* **90e**:60008

[9] Einmahl, U. and Kuelbs, J., Dominating points and large deviations for random vectors, *Probab. Theory Related Fields* **105** (1996), 529–543. *MR* **97k**:60008

[10] Ledoux, M. and Talagrand, M., Characterization of the law of the iterated logarithm in Banach spaces, *Ann. Probab.* **16** (1988), 1242–1264. *MR* **89i**:60016

[11] Schroeder, C., *I*-projection and conditional limit theorems for discrete parameter Markov processes, *Ann. Probab.* **21** (1993), 721–758. *MR* **94g**:60054

[12] Campenhout, J. M. van and Cover, T. M., Maximum entropy and conditional probability, *IEEE Trans. Inform. Theory* **27** (1981), 483–489. *MR* **83h**:94008

[13] Zabell, S. L., Rates of convergence for conditional expectations, *Ann. Probab.* **8** (1980), 928–941. *MR* **82d**:60047

[14] Zabell, S. L., A limit theorem for expectations conditional on a sum, *J. Theoret. Probab.* **6** (1993), 267–283. *MR* **94b**:60043

(Received June 10, 1997)

DEPARTMENT OF MATHEMATICS AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CA 94305
U.S.A.

amir@math.stanford.edu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
480 LINCOLN DRIVE
MADISON, WI 53706
U.S.A.

kuelbs@math.wisc.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 127–140

FUNCTIONAL DEPENDENCIES IN RANDOM DATABASES

J. DEMETROVICS, G. O. H. KATONA, D. MIKLÓS
O. SELEZNJEV and B. THALHEIM

To the memory of Alfréd Rényi

1. Introduction

A database can be considered as a matrix, where the rows contain the data of one individual (object, etc.) and the columns contain the data of the same type: last name, first name, date of birth, etc. The types of data are called *attributes*. These data are sometimes logically dependent. Consider the following example, where the attributes are the last name (denoted by a), the first name (b), the year of the birth (c), the month of the birth (d), the day of the birth (e), the age in years (f), the age in months (g) and the age in days (h). It is obvious that c determines f. On the other hand, the pair $\{c,d\}$ determines both f and g, finally the set $\{c,d,e\}$ determines all of f, g and h.

This is formalized in the following way. Let R be an $m\times n$ matrix with different rows and Ω denote the set of its columns, that is, $|\Omega|=n$. Suppose that $A\subseteqq\Omega, b\in\Omega$. We say that b *functionally depends* on A and write $A\to b$ if R contains no two rows containing equal entries in the columns belonging to A and different entries in b.

In most of the database theory it is supposed that the *functional dependencies* $A\to b$ are a priori known by the logic of the data, as in the above example. Our way of looking at the situation is different. We suppose that we have to find the functional dependencies in a large database (both m and n are large). If nothing is known about R, it is natural to assume that the entries are independently chosen. The question is: what the typical size of the minimal sets A such that $A\to b$ is.

Thus the first mathematical question is the following. Choose the entries of the matrix R totally independently, following the probability distribution $(q_1,\dots,q_d)$. What is the minimum size l of A such that $A\to b$ holds with

---

1991 *Mathematics Subject Classification.* Primary 68P15, 68R05, 15A52.
Key words and phrases. Functional dependency, random database, random matrix, sieve method.

Supported by the Hungarian Foundation for Scientific Research (OTKA) Grant T 016524, T 016389 and the German Natural Science Research Council Contract BB-II-B1-3141-211(94).

high probability for any set $A \subset \Omega$, $|A| \geqq l$ and any column $b \in \Omega$? The answer is

$$\frac{2m}{-\log_2(q_1^2 + \cdots + q_d^2)},$$

as it is given precisely in Corollary 1. Theorem 2 generalizes this result for the case when the entries have different distributions in the different columns.

Section 2 develops a sieve method for estimating the probability of the event that all the outcomes of a many times repeated experiment are different. This result is applied for the rows of a random matrix in Section 3: Theorem 1 determines the asymptotic probability of the event that the rows of the random matrix are different. This theorem is of crucial importance in proving Theorem 2.

If A is larger than the above critical size then $A \to b$ holds with high probability for any given b. However, it will not be true for each element b of a large set Ω. Theorem 3 determines the asymptotic size of the A's satisfying $A \to \Omega$.

The method of the present paper is combinatorial. Paper [2] of the same authors contains similar (but not identical) results. The method of that paper is probabilistic, and uses the so-called Poisson approximation technique (Stein–Chen method, see [1]).

2. A sequence of experiments with different outcomes

We may obtain a counterexample for $A \to b$ if the entries of two rows in the submatrix determined by A are equal. So the critical situation is when all these rows are different. This is why this section is devoted to the probability of the event that all the outcomes of a repeated experiment are different.

Let $E_1, \ldots, E_s$ be mutually exclusive events with respective probabilities $p_1, \ldots, p_s$, where $\sum_{i=1}^{s} p_i = 1$. The distribution is denoted by μ. Choose independently, m times, from these events with this distribution. That is, $\mathbf{P}(\xi_i = E_j) = p_j$ is supposed for all $1 \leqq i \leqq m$ and $1 \leqq j \leqq s$. Moreover, the ξ's are totally independent. Let $\mathbf{P}(\mu, m)$ be the probability of the event that $\xi_1, \ldots, \xi_m$ are all different.

Lemma 3 is the main result of the section giving good estimates on $\mathbf{P}(\mu, m)$.

For an arbitrary sequence of outcomes a trivial graph can be defined. The outcomes are the vertices and two vertices are adjacent if they have the same value. This is why we consider the following graphs. Our goal is actually to estimate the probability that this graph is empty.

The vertex-disjoint union of complete graphs with $m_1, \ldots, m_r$, resp., vertices is denoted by $G(m_1, \ldots, m_r)$. A graph consisting of vertex-disjoint edges is a *matching*. The vertex-disjoint union of a matching and a path consisting of two edges is called a *V-matching*. Finally, the vertex-disjoint union of a matching and a path consisting of three edges is an *N-matching*.

LEMMA 1. *Let* $m_1, \ldots, m_r$ $(0 \leqq r)$ *be non-negative integers. Then*

$$\sum_{\text{matching of } j \text{ edges}} (-1)^j + \sum_{\text{V-matching}} 1 + \sum_{\text{N-matching}} 1 \geqq 0, \tag{1}$$

where the matchings, V-matchings and N-matchings are arbitrary subgraphs of $G(m_1, m_2, \ldots, m_r)$.

PROOF. $2 \leqq m_i$ $(1 \leqq i \leqq r)$ can be supposed. Two cases will be distinguished.

(i) $m_1 = m_2 = \cdots = m_r = 2$. The number of matchings of j edges in $G(2, \ldots, 2)$ is $\binom{r}{j}$ therefore the left-hand side of (1) is

$$\sum_{j=0}^{r} \binom{r}{j} (-1)^j,$$

which is 0 if $0 < r$ and 1 if $r = 0$.

(ii) One of the m's > 2. An injection will be given from the set of all negative terms into a set of some positive terms in (1). Actually the injection will be defined on sets of subgraphs of $G(m_1, m_2, \ldots, m_r)$. A negative term is generated by a matching M of j edges, where j is odd. Suppose that there are at least two edges of M in one of the components of $G(m_1, m_2, \ldots, m_r)$. Join any two endpoints of these two edges by a new edge. The injection assigns this N-matching to M.

Suppose that no component of $G(m_1, m_2, \ldots, m_r)$ contains at least two edges of M but there is a component with at least 3 vertices and containing exactly one edge of M. Then this edge will be replaced by a pair of adjacent edges in the same component. As the number of such pairs is $\geqq$ the number of edges in a complete graph on $\geqq 3$ vertices, this can be defined as a part of an injection. (Actually the assignment can be made in such a way that the pair contains the edge, however, this fact is not needed and its proof is somewhat more difficult.)

The only remaining case is when all components with at least three vertices are disjoint to M. Then add an edge of this component to M. This matching contains an even number of edges therefore it generates 1 in (1).

It is easy to see that the function defined above is an injection and it assigns positive terms to negative terms, proving (1). □

LEMMA 2. *Let* $m_1, \ldots, m_r$ $(0 < r)$ *be non-negative integers, at least one of them is* $\geqq 2$. *Then*

$$\sum_{\text{matching of } j \text{ edges}} (-1)^j + \sum_{\text{V-matching of}} (-1) + \sum_{\text{N-matching}} (-1) \leqq 0, \tag{2}$$

where the matchings, V-matchings and N-matchings are subgraphs of $G(m_1, m_2, \ldots, m_r)$.

PROOF. The proof is analogous to the previous one. The only difference is that here the injection assigns a negative term to a positive term generated by a matching of even number of edges. □

LEMMA 3.

$$1+\sum_{j=1}^{\lfloor \frac{m}{2}\rfloor}\frac{(-1)^j}{j!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2j+2}{2}\Big(\sum_{i=1}^{s}p_i^2\Big)^j-$$

$$-\sum_{j=0}^{\lfloor \frac{m-3}{2}\rfloor}\frac{1}{j!}\binom{m}{3}\binom{m-3}{2}\binom{m-5}{2}\cdots\binom{m-2j-1}{2}\Big(\sum_{i=1}^{s}p_i^3\Big)\Big(\sum_{i=1}^{s}p_i^2\Big)^j-$$

$$-\sum_{j=0}^{\lfloor \frac{m-4}{2}\rfloor}\frac{1}{j!}\binom{m}{4}\binom{m-4}{2}\binom{m-6}{2}\cdots\binom{m-2j-2}{2}\Big(\sum_{i=1}^{s}p_i^4\Big)\Big(\sum_{i=1}^{s}p_i^2\Big)^j\leqq$$

$$\leqq \mathbf{P}(\mu, m)\leqq \tag{3}$$

$$\leqq 1+\sum_{j=1}^{\lfloor \frac{m}{2}\rfloor}\frac{(-1)^j}{j!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2j+2}{2}\Big(\sum_{i=1}^{s}p_i^2\Big)^j+$$

$$+\sum_{j=0}^{\lfloor \frac{m-3}{2}\rfloor}\frac{1}{j!}\binom{m}{3}\binom{m-3}{2}\binom{m-5}{2}\cdots\binom{m-2j-1}{2}\Big(\sum_{i=1}^{s}p_i^3\Big)\Big(\sum_{i=1}^{s}p_i^2\Big)^j+$$

$$+\sum_{j=0}^{\lfloor \frac{m-4}{2}\rfloor}\frac{1}{j!}\binom{m}{4}\binom{m-4}{2}\binom{m-6}{2}\cdots\binom{m-2j-2}{2}\Big(\sum_{i=1}^{s}p_i^4\Big)\Big(\sum_{i=1}^{s}p_i^2\Big)^j.$$

PROOF. $\mathbf{P}(\mu, m)$ is the probability of the event that $\xi_1, \xi_2, \ldots, \xi_m$ are all different, that is, one minus the sum of the probabilities

$$\mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k), \tag{4}$$

where $C_1, C_2, \ldots, C_t$ is a partition of $\{1, 2, \ldots, m\}$ with at least one C having more than one element, and $v_1, v_2, \ldots, v_t$ are different elements of $\{1, 2, \ldots, s\}$. Such partitions will be called *non-elementary*.

$\mathbf{P}(\mu, m)$ contains the probabilities in (4) with zero weight, therefore if they are counted with the weight given in (1) then it leads to an upper estimate. Consider the sum

$$\sum_{\text{partition}}\Bigg(\sum_{\text{matching of } j \text{ edges}}(-1)^j+\sum_{\text{V-matching}}1+\sum_{\text{N-matching}}1\Bigg)\times$$

$$\times \mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k). \tag{5}$$

If the partition is the elementary one, then the inner sums are empty with one exception, the empty matching. This leads to the sum of the probabilities where all ξ's are different. Therefore (5) is the sum in which the probabilities of the events, where all ξ's are different stand with weight 1, while the other probabilities stand with a non-negative weight. Consequently, (5) is an upper estimate on $\mathbf{P}(\mu, m)$.

Change the order of sums in (5).

$$\begin{aligned}
&\sum_{\text{matching of } j \text{ edges}} (-1)^j \sum_{\text{partition}} \mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k) + \\
&+ \sum_{\text{V-matching}} \sum_{\text{partition}} \mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k) + \\
&+ \sum_{\text{N-matching}} \sum_{\text{partition}} \mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k),
\end{aligned} \tag{6}$$

where those partitions are taken for which the given matching is a subgraph of the graph generated by the partition. Consider

$$\sum_{\text{partition}} \mathbf{P}(\xi_u = E_{v_k} \text{ if } u \in C_k)$$

for a given matching of j edges. This is nothing else but the probability of the event that the ξ's adjacent in the matching are equal:

$$\left(\sum_{i=1}^{s} p_i^2\right)^j.$$

The number of matchings with j edges is

$$\frac{1}{j!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2j+2}{2}.$$

This gives the fifth row of (3). The second and third rows of (6) lead, in a similar manner, to the sixth and seventh rows of (3), resp.

The lower estimate is proved in the same way. □

3. Random matrix with different rows

The Lemma 3 will be used for random matrices. Let R be a random matrix with m rows and z columns, where the entries of the jth column can have d_j different values with probabilities $q_{j1}, \dots, q_{jd_j}$, respectively. All the entries are chosen totally independently. Then the probability of the occurrence of a certain row in R is $q_{1i_1}q_{2i_2}\cdots q_{zi_z}$, where i_j is arbitrary between 1 and d_j. The probability distribution of these sequences will be denoted by π_z. The following trivial observation will be used later.

LEMMA 4. *If* $m \leqq m'$ *then* $\mathbf{P}(\pi_z, m) \geqq \mathbf{P}(\pi_z, m')$.

We want to study the probability of the event that the rows of the above matrix are different. Therefore the probabilities $q_{1i_1} q_{2i_2} \cdots q_{zi_z}$ will be taken as p's in Lemma 3. Consider $\sum_{i=1}^{s} p_i^k$ for these probabilities:

$$\sum_{1 \leqq i_1 \leqq d_1, \ldots, 1 \leqq i_z \leqq d_z} (q_{1i_1} q_{2i_2} \cdots q_{zi_z})^k = \sum_{1 \leqq i_1 \leqq d_1, \ldots, 1 \leqq i_z \leqq d_z} q_{1i_1}^k q_{2i_2}^k \cdots q_{zi_z}^k = \prod_{i=1}^{z} (q_{i1}^k + \cdots + q_{id_i}^k). \tag{7}$$

Our investigations will be of asymptotic nature. From now on it is supposed that m tends to the infinity and the other parameters depend on m: $z = z(m), d_i = d_i(m), q_{ij} = q_{ij}(m)$. Our asymptotic assumption on them will be such that the first non-trivial term in the Lemma 3, that is,

$$m^2 \sum_{i=1}^{s} p_i^2 = m^2 \prod_{i=1}^{z} (q_{i1}^2 + \cdots + q_{id_i}^2) \tag{8}$$

tends to a non-zero constant. It will be done in a logarithmic way, therefore the quantities $\log(q_{i1}^2 + \cdots + q_{id_i}^2)$ will play an important role.(log will always mean log of base 2.) Denote the distribution $(q_{i1}, \ldots, q_{id_i})$ by κ_i. Rényi [3] introduced the so-called entropy of order α. For $\alpha = 2$ it is $H_2(\kappa) = -\log(q_1^2 + \cdots + q_d^2)$ if $\kappa = (q_1, \ldots, q_d)$.

LEMMA 5. *If*

$$2 \log m - \sum_{i=1}^{z} H_2(\kappa_i) \to a \tag{9}$$

when $m \to \infty$ *then*

$$1 + \sum_{j=1}^{\lfloor \frac{m}{2} \rfloor} \frac{(-1)^j}{j!} \binom{m}{2} \binom{m-2}{2} \cdots \binom{m-2j+2}{2} \Big(\sum_{i=1}^{s} p_i^2 \Big)^j \tag{10}$$

tends to

$$e^{-2^{a-1}}$$

for the distribution π_z.

PROOF. Consider the limit of one term for a fixed j.

$$\binom{m}{2} \binom{m-2}{2} \cdots \binom{m-2j+2}{2}$$

can be replaced by

$$\frac{m^{2j}}{2^j}.$$

On the other hand

$$\left(\sum_{i=1}^{s} p_i^2\right)^j = 2^{-j\sum_{i=1}^{z} H_2(\kappa_i)}$$

follows by the definition of the entropy of order 2 and (7). Therefore the limit of the jth term in (10) is the same as the limit of

$$\frac{(-1)^j}{j!} 2^{j(2\log m - \sum_{i=1}^{z} H_2(\kappa_i) - 1)}, \tag{11}$$

that is,

$$\frac{(-1)^j}{j!} 2^{j(a-1)}.$$

(9) implies that the sum of (11) and therefore (10) are uniformly convergent, hence the limit of (10) is equal to the infinite sum of the limits of its terms, that is,

$$\sum_{j=0}^{\infty} \frac{(-1)^j}{j!} 2^{j(a-1)} = e^{-2^{a-1}}.$$ □

We want to show that the other terms in the lower and upper estimates of (3) tend to zero under condition (9). Before proving that some other lemmas are needed.

LEMMA 6. *If $\kappa = (q_1, \dots, q_d)$ is a probability distribution, where $\varepsilon \leqq q_1, q_2$ $(0 < \varepsilon \leqq \frac{1}{2})$ then*

$$\frac{\left(\sum_{i=1}^{d} q_i^3\right)^2}{\left(\sum_{i=1}^{d} q_i^2\right)^3} \leqq 1 - 4\varepsilon^6. \tag{12}$$

PROOF. Consider the difference of the denominator and the numerator:

$$\sum_{i=1}^{d} q_i^6 + 3\sum_{i<j} q_i^4 q_j^2 + 3\sum_{i<j} q_i^2 q_j^4 + 6\sum_{i<j<k} q_i^2 q_j^2 q_k^2 - \left(\sum_{i=1}^{d} q_i^6 + 2\sum_{i<j} q_i^3 q_j^3\right) \geqq$$

$$\geqq \left(\sum_{i<j} q_i^4 q_j^2 + \sum_{i<j} q_i^2 q_j^4 - 2\sum_{i<j} q_i^3 q_j^3\right) + 2\sum_{i<j}(q_i^4 q_j^2 + q_i^2 q_j^4) =$$

$$= \sum_{i<j} q_i^2 q_j^2 (q_i^2 + q_j^2 - 2q_i q_j) + 2 \sum_{i<j} (q_i^4 q_j^2 + q_i^2 q_j^4) \geqq$$

$$\geqq 2q_1^4 q_2^2 + 2q_1^2 q_2^4) \geqq 4\varepsilon^6.$$

Using the fact that the denominator is at most 1, (12) easily follows. □

LEMMA 7. *If* $\kappa = (q_1, \ldots, q_d)$ *is a probability distribution, where* $\varepsilon \leqq q_1, q_2$ $(0 < \varepsilon \leqq \frac{1}{2})$ *then*

$$\frac{\sum_{i=1}^{d} q_i^4}{\left(\sum_{i=1}^{d} q_i^2\right)^2} \leqq 1 - 2\varepsilon^4. \tag{13}$$

PROOF. The proof is similar but easier than the previous one:

$$\sum_{i=1}^{d} q_i^4 + 2 \sum_{i<j} q_i^2 q_j^2 - \sum_{i=1}^{d} q_i^4 =$$

$$= 2 \sum_{i<j} q_i^2 q_j^2 \geqq 2q_1^2 q_2^2 \geqq 2\varepsilon^4.$$

□

LEMMA 8. *If* (9) *and*

(14) $\varepsilon \leqq q_{i1}, q_{i2}$ *hold for all* i *with a fixed* ε $\left(0 < \varepsilon \leqq \frac{1}{2}\right)$

then the second and third rows of (3) *tend to zero.*

PROOF. The jth term of the second row of (8) can be upperbounded by

$$\left(m^3 \prod_{i=1}^{z} (q_{i1}^3 + \cdots + q_{id_i}^3)\right) \left(m^2 \prod_{i=1}^{z} (q_{i1}^2 + \cdots + q_{id_i}^2)\right)^j. \tag{15}$$

The second factor tends to

$$\frac{2^{j(a-1)}}{j!}$$

as we have seen in the proof of Lemma 5. (9) implies

$$m^2 \prod_{i=1}^{z} (q_{i1}^2 + \cdots + q_{id_i}^2) \to 2^a \tag{16}$$

therefore the first factor of (15) can be expressed as

$$m^3 \prod_{i=1}^{z} (q_{i1}^3 + \cdots + q_{id_i}^3) = \left(m^2 \prod_{i=1}^{z} \Big(\sum_{j=1}^{d_i} q_{ij}^2 \Big)^3 \right)^{\frac{3}{2}} \prod_{i=1}^{z} \frac{\sum_{j=1}^{d_i} q_{ij}^3}{\Big(\sum_{j=1}^{d_i} q_{ij}^2 \Big)^{\frac{3}{2}}}. \tag{17}$$

The first factor of (17) tends to $2^{\frac{3}{2}a}$ while Lemma 6 gives the upper bound $(1-4\varepsilon^6)^{\frac{z}{2}}$ for the second factor.

The conditions of the lemma imply $-\log(2\varepsilon^2) \leqq H(\kappa_i)$ thus (9) results in $z \to \infty$ when $m \to \infty$. $(1-4\varepsilon^6)^{\frac{z}{2}}$ and consequently (17) tend to zero. By the uniform convergence, the infinite sum of (15) and the second row of (3) also tend to zero.

The convergence of the third row can be proved in the same way, using Lemma 7. □

THEOREM 1. *Let R be a random matrix with m rows and z columns, where the entries of the jth column can have d_j different values with probabilities $q_{j1}, \ldots, q_{jd_j}$, respectively. All the entries are chosen totally independently. Suppose that* (14) *holds. Then the probability of the event that the rows of R are all different satisfies*

$$\mathbf{P}(\pi_z, m) \to \begin{cases} 0, & \text{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to +\infty, \\ e^{-2^{a-1}}, & \text{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to a, \\ 1, & \text{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to -\infty. \end{cases}$$

PROOF. The middle row of the statement follows by Lemmas 3 and 8. The first and third rows are consequences of Lemma 4. □

In [4] Rényi proved a theorem on random matrices in connection with search theory (see also [5] and [6]). It is basically equivalent to the special case of the above theorem when κ_i's are the same. His method was different.

REMARK. The condition that each distribution contains two "large" probabilities ($\varepsilon \leqq q_{i1}, q_{i2}$) was important in the proof. This is shown by the following example. Let $\kappa_i = (\frac{1}{2}, \frac{1}{2m}, \ldots, \frac{1}{2m})$. Then the left-hand side of (12) is

$$\left(1 - \frac{2m}{(m+1)^2}\right)\left(1 - \frac{m-1}{m^2+m}\right),$$

which is not bounded from 1. Take $z = \log m$. As $H_2(\kappa_i) \to 2$, (9) holds with zero. However, the second factor of (17) does not tend to zero. (3) cannot be used.

Another example is $\kappa_i = (\frac{1}{m}, \frac{1}{m^2}, \dots, \frac{1}{m^2})$. We do not know, however, if the statement of the theorem holds for these and similar distributions. □

4. Typical sizes of functional dependencies and minimal keys

Let $\mathbf{P}(\mu, m, k)$ denote the probability of the event that exactly k pairs of $\xi_1, \dots, \xi_m$ are equal to each other and all other pairs are different. (More precisely: there are $2k$ distinct indices $i_1, \dots, i_k$ and $j_1, \dots, j_k$ such that $\xi_{i_l} = \xi_{j_l}$ for all $1 \leq l \leq k$, but $\xi_{i_l} \neq \xi_{j_m}$ for all $l \neq m$, $\xi_i \neq \xi_{i_l}$ and $\xi_i \neq \xi_{j_l}$ if $i \notin \{i_1, \dots, i_k, j_1, \dots, j_k\}$.)

LEMMA 9. *Suppose that k is fixed, m tends to infinity and* (14) *holds. Then*

$$\mathbf{P}(\pi_z, m, k) \to \frac{1}{k!} 2^{k(a-1)} e^{-2^{a-1}}, \quad \text{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to a.$$

PROOF. There are

$$\frac{1}{k!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2k+2}{2}$$

ways to choose the set $\{i_1, \dots, i_k, j_1, \dots, j_k\}$. Suppose that $i_1 = 1, j_1 = 2, \dots, i_k = 2k-1, j_k = 2k$ and determine the probability of the event that $\xi_1 = \xi_2, \dots, \xi_{2k-1} = \xi_{2k}$. The probabilities for the other choices of pairs will be the same. It is easy to see that

$$\mathbf{P}(\xi_1 = \xi_2) = \prod_{i=1}^{z} \sum_{j=1}^{d_i} q_{ij}^2.$$

We need the kth power of this expression. Finally, $\xi_1, \xi_3, \dots, \xi_{2k-1}, \xi_{2k+1}, \xi_{2k+2}, \dots, \xi_m$ must be all different. The probability of this event is $\mathbf{P}(\pi_z, m-k)$.

$$\mathbf{P}(\pi_z, m, k)$$

$$= \frac{1}{k!}\binom{m}{2}\binom{m-2}{2}\cdots\binom{m-2k+2}{2}\left(\prod_{i=1}^{z}\Big(\sum_{j=1}^{d_i} q_{ij}^2\Big)\right)^k \mathbf{P}(\pi_z, m-k). \tag{18}$$

The last factor is asymptotically equal to $\mathbf{P}(\pi_z, m)$ since $\log m - \log(m-k) \to 0$. Therefore Theorem 1 gives its limit. The limit of the product of the other factors of (18) was determined in the proof of Lemma 5:

$$\frac{1}{k!} 2^{k(a-1)}.$$ □

LEMMA 10. *If (9) and (14) hold then the probability of the event that there are three equal ξ's tends to zero.*

PROOF. The sum of the probabilities in Lemma 9 tends to 1. □

Let Ω denote the set of columns of the matrix R. Suppose $A \subset \Omega, b \in \Omega - A$. We say that b *functionally depends* on A if R contains no two rows equal in the columns belonging to A and different in b. In notation: $A \to b$. For sake of simplicity b is supposed to be the bth column.

THEOREM 2. *Let R be a random matrix with m rows and $n = n(m)$ columns with the distribution described above ((14) holds, again). Suppose that A_z is a set of $z = z(m)$ columns of R and b is a column not in A_z.*

$$\mathbf{P}(A_z \to b, m) \to \begin{cases} 0, & \textit{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to +\infty, \\ e^{2^{a-1}(2^{-H_2(\kappa_b)}-1)}, & \textit{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to a, \\ 1, & \textit{if } 2\log m - \sum_{i=1}^{z} H_2(\kappa_i) \to -\infty. \end{cases}$$

PROOF. Consider the restrictions of the rows of R within A_z. These rows of length z define a random partition $\gamma = (m_1, \dots, m_r)$ of m, where one class consists of the equal rows. Suppose $(m_1 \geqq \dots \geqq m_r)$. Start with the well known equation

(19)
$$\mathbf{P}(A_z \to b, m) = \sum_{m_1,\dots,m_r} \mathbf{P}(A_z \to b | \gamma = (m_1, \dots, m_r))\mathbf{P}(\gamma = (m_1, \dots, m_r)).$$

The right-hand side of (19) will be divided into two parts: (i) $m_1 \leqq 2$, (ii) $m_1 \geqq 3$. For case (ii) the following trivial inequality is needed:

$$\begin{aligned} &\sum_{m_1 \geqq 3, m_2,\dots,m_r} \mathbf{P}(A_z \to b, m | \gamma = (m_1, \dots, m_r))\mathbf{P}(\gamma = (m_1, \dots, m_r)) \\ (20) \quad &\leqq \sum_{m_1 \geqq 3, m_2,\dots,m_r} \mathbf{P}(\gamma = (m_1, \dots, m_r)) = \mathbf{P}(\text{there are 3 equal } \xi\text{'s}). \end{aligned}$$

The last quantity tends to zero under condition (9) therefore case (i) should only be considered. More precisely, if (9) holds then the limit of $\mathbf{P}(A_z \to b, m)$ is equal to the limit of

$$\sum_{m_1 \leqq 2, m_2,\dots,m_r} \mathbf{P}(A_z \to b, m | \gamma = (m_1, \dots, m_r))\mathbf{P}(\gamma = (m_1, \dots, m_r)).$$

This expression can be rewritten in the form

$$\sum_k \mathbf{P}(A_z \to b, m|\gamma = (2,\dots,1) \text{ (the number of 2's is } k))$$

(21) $$\times \mathbf{P}(\gamma = (2,\dots,1) \text{ (the number of 2's is } k))$$

$$= \sum_k \mathbf{P}(A_z \to b, m|\gamma = (2,\dots,1) \text{ (the number of 2's is } k))\mathbf{P}(\pi_z, m, k).$$

Here

$$\mathbf{P}(A_z \to b, m|\gamma = (2,\dots,1) \text{ (the number of 2's is } k)) = \Big(\sum_{j=1}^{d_b} q_{bj}^2\Big)^k = 2^{-kH_2(\kappa_b)}.$$

On the other hand, the limit of $\mathbf{P}(\pi_z, m, k)$ is given by Lemma 9. Therefore the limit of (21) is

$$e^{-2^{a-1}} \sum_{k=0}^{\infty} \frac{1}{k!} \left(2^{(a-1)-H_2(\kappa_b)}\right)^k = e^{2^{a-1}(2^{-H_2(\kappa_b)}-1)}.$$

The middle row of the statement is proved. The first and third rows are consequences of the inequality $\mathbf{P}(A_z \to b, m) \geqq \mathbf{P}(A_z \to b, m')$ for $m \leqq m'$. □

COROLLARY 1. *Let R be a random matrix with m rows and $n = n(m)$ columns, where the entries are chosen totally independently with probabilities $q_1, \dots, q_d$. Suppose that A_z is a set of $z = z(m)$ columns of R and b is a column not in A_z. Use the notation $H_2 = -\log \sum_{i=1}^{d} q_i^2$. Then*

$$\mathbf{P}(A_z \to b, m) \to \begin{cases} 0, & \text{if } \dfrac{2\log m}{H_2} - z \to +\infty, \\ e^{2^{aH_2-1}(2^{-H_2}-1)}, & \text{if } \dfrac{2\log m}{H_2} - z \to a, \\ 1, & \text{if } \dfrac{2 log m}{H_2} - z \to -\infty. \end{cases}$$

The main content of the latter statement is that if A is a set of columns of size definitely larger than $\frac{2\log m}{H_2}$, then $A \to b$ holds with high probability for any b.

We say, in general, that B *functionally depends* on A and write $A \to B (A. B \subseteq \Omega)$ if $A \to b$ holds for each element b of B. Theorem 2 can be easily generalized for this case. We only have to imagine the set of columns in B as one column. It is worth supposing that $A \cap B = \emptyset$. Then $H_2(\kappa_b)$ can be replaced by $H_2(\kappa_B) = \sum_{b \in B} H_2(\kappa_b)$.

Let us turn back to the case when κ_i does not depend on i. If the size of B is finite, say u, then the Consequence can be generalized for $A \to B$, only $-H_2$ should be multiplied by u. However, if $|B|$ tends to infinity, then the middle probability becomes simply $e^{-2^{aH_2-1}}$.

We say that $A \subseteq \Omega$ is a *key* if $A \to \Omega$ (or equivalently $A \to \Omega - A$) holds. A is a *minimal key* if it is a key and no proper subset is a key. The above reasoning proves the following statement.

THEOREM 3. *Let R be a random matrix with m rows and $n = n(m)$ columns, where the entries are chosen totally independently following the distribution κ. Suppose that $n - \frac{2\log m}{H(\kappa)}$ tends to infinity and A_z is a set of columns of R. Then*

$$\mathbf{P}(A_z \textit{ is a key}) \to \begin{cases} 0, & \textit{if } \dfrac{2\log m}{H_2} - z \to +\infty, \\ e^{-2^{aH_2-1}}, & \textit{if } \dfrac{2\log m}{H_2} - z \to a, \\ 1, & \textit{if } \dfrac{2\log m}{H_2} - z \to -\infty. \end{cases}$$

It can be briefly said that the sets A of size somewhat larger than $\frac{2\log m}{H_2}$ are keys with high probability.

REFERENCES

[1] BARBOUR, A. D., HOLST L. and JANSON, S., *Poisson approximation*, Oxford Studies in Probability, **2**, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1992. *MR* **93g**:60043

[2] DEMETROVICS, J., KATONA, G. O. H., MIKLÓS, D., SELEZNJEV, O. and THALHEIM, B., Asymptotic properties of keys and functional dependencies in random databases, *Theoret. Comput. Sci.* **190** (1998), 151–166.

[3] RÉNYI, A., A few fundamental problems of information theory, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **10** (1960), 251–282 (in Hungarian). *MR* **26** #1218

[4] RÉNYI, A., A general method for proving theorems in probability theory and some applications, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **11** (1961), 79–105 (in Hungarian).

[5] RÉNYI, A., Statistical laws of accumulation of information, *Bull. Inst. Internat. Statist.* **39** (1962), livraison 2, 311–316. *MR* **28** #5845

[6] RÉNYI, A., On the statistical regularities of information accumulation, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **12** (1962), 15–33 (in Hungarian). *MR* **28** #1023

(Received June 29, 1998)

J. Demetrovics

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI
KUTATÓ INTÉZETE
KENDE U. 13–17
H–1111 BUDAPEST
HUNGARY

dj@ilab.sztaki.hu

G. O. H. Katona and D. Miklós

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

ohkatona@math-inst.hu
dezso@math-inst.hu

O. Seleznjev

FACULTY OF MATHEMATICS AND MECHANICS
MOSCOW STATE UNIVERSITY
RU–119 899 MOSCOW
RUSSIA

seleznev@top.rector.msu.su

B. Thalheim

FACHBEREICH INFORMATIK
TECHNISCHE UNIVERSITÄT COTTBUS
POSTFACH 101344
D–03013 COTTBUS
GERMANY

thalheim@informatik.tu-cottbus.de

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 141–149

BIASED POSITIONAL GAMES ON HYPERGRAPHS

D. DUFFUS, T. ŁUCZAK and V. RÖDL

Dedicated to the memory of Alfréd Rényi

Abstract

Let $\mathbf{G}(p, q, H)$ be the game played on a hypergraph H by two players, who alternately choose p and q vertices, respectively. The object of the first player is to claim all vertices of a hyperedge of H, while the second player tries to prevent him from doing so. We give a sufficient condition for the first player to win $\mathbf{G}(p, q, H)$ played on an r-uniform hypergraph H and argue that this condition is close to optimal. Furthermore, we answer a question of Galvin by proving that the first player has a winning strategy in $\mathbf{G}(1, q, H)$ for each 3-uniform hypergraph H with chromatic number large enough.

1. Introduction

Erdős and Selfridge [4] introduced the following unbiased game played on set systems: two players alternately pick elements of the sets, the first who chooses all elements of some set in the system is the winner. Csirmaz [3] introduced this biased version. Let p and q be positive integers and let H be a finite hypergraph. The game $\mathbf{G}(p, q. H)$ is played by two players – the first, suggestively called Maker, chooses at most p as yet unchosen vertices of H. the second, Breaker, chooses at most q unchosen vertices. Maker's objective is to choose all vertices of an edge of H, while Breaker wants to prevent this. Players alternate choices until all vertices of some edge are chosen by Maker, a win for Maker, or until all vertices are chosen and there is no Maker edge, a win for Breaker.

Erdős and Selfridge [4] provided a sharp sufficient condition for Breaker to win $\mathbf{G}(1, 1, H)$. Beck accomplished the same for $\mathbf{G}(p, q, H)$; earlier, Csirmaz [3] obtained a weaker result. Beck proved the following [1, Theorem 1]. Given a hypergraph H, let $E(H)$ denote its set of edges. If

$$\sum_{A \in E(H)} (1+q)^{-|A|/p} < \frac{1}{1+q}$$

1991 *Mathematics Subject Classification.* Primary 90D46, 05C65.
Key words and phrases. Biased positional games, hypergraphs.

The research of the second author was supported in part by NSF International Programs Grant INT-9406971.
The research of the third author was supported in part by NSF Grant DMS-9401559.

then Breaker has a winning strategy for $\mathbf{G}(p,q,H)$. Moreover, the result is sharp in the strong sense that for all p and q there are infinitely many H for which equality holds above and for which Maker wins $\mathbf{G}(p,q,H)$ (see [1]).

In addition, Beck obtained a sufficient condition on H for Maker to win $\mathbf{G}(p,q,H)$. Given a hypergraph H on $v(H)$ vertices with $d(H)$ the maximum number of edges containing a pair of vertices, he proved that if

$$\sum_{A\in E(H)}\left(1+\frac{q}{p}\right)^{-|A|}>p^2q^2(p+q)^{-3}d(H)v(H) \tag{1}$$

then Maker has a winning strategy for $\mathbf{G}(p,q,H)$ [1, Theorem 2]. In the case of r-uniform, simple hypergraphs H with $e(H)=|E(H)|$, this result amounts to following: if

$$e(H)>\frac{q^2(p+q)^{r-3}}{p^{r-2}}v(H) \tag{2}$$

then Maker has a winning strategy for $\mathbf{G}(p,q,H)$.

In this paper, we improve the second result of Beck in the case of r-uniform simple hypergraphs for q large in comparison to p and r (see Theorem 1, §2). As well, we show that this result is quite sharp by proving there are Breaker-win hypergraphs with only marginally fewer edges (see Theorem 2, §3). In the last section, we show how Beck's results can be used to answer a question of Galvin [5] (see Theorem 3, §4) and we pose related problems.

2. A sufficient condition for a Maker win

In this section we obtain an improvement of [1, Theorem 2] in the special case of uniform, simple hypergraphs. To that end, we employ some terminology for analysis of the game $\mathbf{G}(p,q,H)$. Let us say that, during play, an edge is *surviving* if none of its vertices has yet been chosen by Breaker; the *size* of a surviving edge is the number of its vertices not yet chosen by Maker. A *round* of the game consists of Maker's selection of p vertices and Breaker's selection of q vertices.

THEOREM 1. *Let $r\geqq 1$, $r\geqq p$ and $q\geqq 2$ be integers and let H be an r-uniform simple hypergraph. If*

$$e(H)\geqq\frac{1}{p+1}(2q+1)^{r-p}v(H)$$

then Maker has a winning strategy for $\mathbf{G}(p,q,H)$.

Compare Theorem 1 to Beck's result in this special case, (2), to see the improvement for q large in comparison to p and r. Indeed, for q large, the right-hand side of (2) is of the form $c_{p,r}v(H)q^{r-1}$, while Theorem 1 ensures a Maker win with $c'_{p,r}v(H)q^{r-p}$.

PROOF. During the play of $\mathbf{G}(p,q,H)$, after a round has been completed, three cases may occur.

Case 1. There is a surviving edge of size at most p.

Case 2. There is a vertex of H which belongs to more than q surviving edges, each of size at most $p+1$.

Case 3. All surviving edges have size greater than p and there is no vertex as in Case 2.

It is obvious that Maker has a winning strategy in both Cases 1 and 2. We argue by showing that it is not possible for Case 3 to hold after every round of $\mathbf{G}(p,q,H)$.

Let H_i be the hypergraph with vertex set $V(H_i)=V(H)$ and edge set consisting of all surviving edges of H after round i. Given an edge $e\in E(H_i)$, we let $|e|$ denote the size of this surviving edge. Set $\rho=\frac{1}{2q+1}$ and define

$$f(H_i)=\sum_{e\in E(H_i)}\rho^{|e|}.$$

Observe that with $H_0=H$,

$$f(H_0)\geqq\frac{1}{p+1}(2q+1)^{r-p}\rho^r v(H).$$

We complete the proof by showing that after round i, if Case 3 holds then there is a choice of p vertices so that no matter which q vertices Breaker chooses to complete round $i+1$, $f(H_{i+1})\geqq f(H_i)$. We proceed by induction on i and use the lower bound on $f(H_0)$ above.

For each $v\in V(H_i)$, let $d(v)=\sum(\rho^{|e|}:v\in e\in E(H_i))$. Since Case 3 holds and $f(H_i)\geqq f(H_0)$,

$$\sum_{v\in V(H_i)}d(v)\geqq(p+1)f(H_i)\geqq(2q+1)^{r-p}\rho^r v(H).$$

So, there is a vertex v with

$$d(v)\geqq(2q+1)^{r-p}\rho^r. \tag{3}$$

Maker chooses p vertices, including a vertex v_0 such that $d(v_0)$ maximizes $d(v)$. Breaker completes round $i+1$ by selecting $w_1,w_2,\dots,w_q$. Let $d'(v_0)$ be the analogous value to $d(v_0)$ for H_{i+1}. Maker's choice of v_0 replaces each summand $\rho^{|e|}$ in $d(v_0)$ by $\rho^{|e|-1}$ in $d'(v_0)$. Since H_i is a simple hypergraph, there are at most q edges in H_i containing v_0 and one or more of the w_i's. As Case 3 holds, each edge of H_i has at least $p+1$ unchosen vertices before round $i+1$. Thus,

$$d'(v_0)-d(v_0)\geqq 2q(d(v_0)-q\rho^{p+1}).$$

Hence,

$$f(H_{i+1}) \geqq f(H_i) + 2q(d(v_0) - q\rho^{p+1}) - \sum_{i=1}^{q} d(w_i).$$

To prove that $f(H_{i+1}) \geqq f(H_i)$, it suffices to show that

$$2qd(v_0) - \sum_{i=1}^{q} d(w_i) \geqq 2q^2\rho^{p+1}. \tag{4}$$

But since $d(w_i) \leqq d(v_0)$ for each i, (4) follows from

$$(2q - q)d(v_0) > 2q^2\rho^{p+1}$$

which is a consequence of (3). □

3. Examples of Breaker-win hypergraphs

To construct Breaker-win hypergraphs with comparatively many edges, it is convenient to use a "continuous" variant of the Box Game analyzed by Chvátal and Erdős [2]. In this game $B(N;p,q)$, there are two players, again Maker and Breaker. The game begins with N empty bins of unlimited capacity. A move by Maker consists of pouring at most p units of a substance into any number of existing bins. Maker's objective is to maximize the amount of the substance in any single bin. Breaker's move is to destroy up to q bins. Breaker's objective is to minimize the maximum amount in any single bin. Let $b(N;p,q)$ denote the maximum amount of the substance that ever appears in a single bin, with measurement taken only after Breaker's turns. Maker's best strategy is, basically, to distribute the substance uniformly among all bins not yet destroyed by Breaker; Breaker's optimal strategy is to destroy the q bins with the greatest content. It follows from [2, §2] that

$$b(N;p,q) \leqq \frac{p}{q}\log N. \tag{5}$$

LEMMA 1. *Let H be a simple r-uniform hypergraph on n vertices, let p and q be positive integers, with $q \geqq 2$ and $r > p$. Suppose that $q \geqq c_0 n^{\frac{1}{r-p}}(\log n)^{\frac{r-p-1}{r-p}}$, where $c_0 = c_0(p,r)$. Then Breaker has a winning strategy in $\mathbf{G}(p,q,H)$.*

PROOF. Let $H_0 = H$ and let H_i be the subhypergraph of H with $V(H_i) = V(H)$ and $E(H_i)$ all edges of H surviving after round i. In order to describe Breaker's play in $\mathbf{G}(p,q,H)$, we define $r-j-1$ games $B_j(n;p_j,q_0)$ $(j = 1,\ldots,r-p-1)$, each an instance of the Box Game variant described above. Play in each round of each of these games is tied to play in $\mathbf{G}(p,q.H)$; we show how Breaker's strategies in $B_j(n; p_j, q_0)$ $(j = 1, \ldots, r-p-1)$ yield a Breaker win in $\mathbf{G}(p,q,H)$.

To define each $B_j(n; p_j, q_0)$, for each $v \in V(H)$, let $C_j(v)$ denote the bin corresponding to v. Edges of H will be placed in these bins as play progresses; we declare that $C_j(v) = \emptyset$ once v is chosen by Breaker and let $|C_j(v)|$ denote the number of edges in bin $C_j(v)$ at a specified point in play. Also, set $q_0 = \left\lfloor \frac{q - \binom{p}{2}}{r-p} \right\rfloor$.

Proceed by induction on the rounds of $\mathbf{G}(p,q,H)$ to define play in it and in each $B_j(n; p_j, q_0)$. Here is ith round play.

After Maker's move in round i of $\mathbf{G}(p,q,H)$, Breaker's play in the ith round of $\mathbf{G}(p,q,H)$ has three components.

(a_i) Breaker chooses a vertex from each edge that contains at least two vertices chosen by Maker in round i. As H_{i-1} is simple, this requires at most $\binom{p}{2}$ vertices.

Maker's play in round i of $B_j(n; p_j, q_0)$ consists of placing in $C_j(v)$ all $e \in H_{i-1}$ such that $v \in e$, and such that e contains exactly j vertices chosen by Maker, a vertex chosen by Maker in round i of $\mathbf{G}(p,q,H)$, and none chosen by Breaker in earlier rounds or as described in (a_i).

(b_i) Breaker continues play in the ith round of $\mathbf{G}(p,q,H)$: for each $j = 1, \ldots, r-p-1$, Breaker chooses those q_0 vertices v which maximize $|C_j(v)|$ after Maker's play in round i of $B_j(n; p_j, q_0)$.

(c_i) Breaker completes play in the ith round of $\mathbf{G}(p,q,H)$ by choosing a vertex from each edge of H_{i-1} which contains $r-p$ Maker vertices after Maker's move in round i of $\mathbf{G}(p,q,H)$.

Complete the ith round of $B_j(n; p_j, q_0)$ by having Breaker set the bins $C_j(v) = \emptyset$ for each of the vertices v it has chosen in round i of $\mathbf{G}(p,q,H)$.

We determine the values of p_j and prove that Breaker's choice of q vertices is sufficient to allow play as described in (a_i), (b_i), and (c_i).

How many edges are put into the bins of $B_1(n; p_1, q_0)$ as the result of Maker's choice of p vertices $x_1, \ldots, x_p$? Each x_k can belong to at most $\frac{n-1}{r-1}$ edges e, and each such e is placed in r bins $C_1(v)$, so we may set $p_1 = 2pn$. From (5), it follows that

$$b_1(n; p_1, q_0) \leqq \frac{2pn}{q_0} \log n.$$

Let $d_1 = \frac{2pn}{q_0} \log n$.

To determine p_2, suppose Maker chooses vertices $x_1, \ldots, x_p$ during round i of $B_2(n; p_2, q_0)$. For an edge e to be added to some bin $C_2(v)$ by Maker during round i, e must belong to some $C_1(x_k)$ after round $i-1$. Otherwise, two vertices of e would have been chosen by Maker in round i of $\mathbf{G}(p,q,H)$, so Breaker would have countered by spoiling e as described in (a_i). Since $|C_1(x_k)| \leqq d_1$, and each edge e can be added to at most r bins $C_2(v)$, we may take $p_2 = pd_1r$. Thus (5) yields

$$b_2(n; p_2, q_0) \leqq \frac{pd_1r}{q_0} \log n. \tag{7}$$

Let $d_2 = \frac{pd_1r}{q_0} \log n$.

The same reasoning about $B_j(n; p_j, q_0)$ shows that

$$b_j(n; p_j, q_0) \leqq \frac{pd_{j-1}r}{q_0} \log n, \tag{8}$$

for $j \leqq r-p-1$. Let $d_j = \frac{pd_{j-1}r}{q_0} \log n$.

We claim that Breaker has a winning strategy in $\mathbf{G}(p,q,H)$ provided that

$$d_{r-p-1} \leqq \frac{q_0}{p}. \tag{9}$$

The components of Breaker's moves in (a_i) and (b_i) require at most

$$\binom{p}{2} + (r-p-1)q_0$$

vertices in each round of $\mathbf{G}(p,q,H)$, so we have at least q_0 left to complete (c_i). No matter how Maker chooses its p vertices $x_1, \dots, x_p$ in its part of round i, each x_k can belong to at most q_0/p edges which, prior to round i, contained $r-p-1$ vertices selected by Maker, by (9). So there can be at most q_0 edges with $r-p$ Maker vertices after Maker's move in round i. Breaker selects a vertex from each edge, as described in (c_i), to prevent Maker's win in the next round.

To obtain a lower bound on q that will ensure that (9) holds, note that (8) gives

$$d_{r-p-1} < 2\left(\frac{p}{q_0} \log n\right)^{r-p-1} r^{r-p-2} n.$$

Thus, the following bound on q_0 suffices:

$$2pr^{\frac{r-p-2}{r-p}} n^{\frac{1}{r-p}} (\log n)^{\frac{r-p-1}{r-p}} < q_0 .$$

□

Recall that Theorem 1 guarantees Maker a winning strategy for $\mathbf{G}(p,q,H)$ for any r-uniform hypergraph H which satisfies $e(H) \geqq c_{p,r} q^{r-p} v(H)$. We shall use Lemma 1 to show that this bound is not far from the best possible, that is, we construct Breaker-win hypergraphs with many edges – not too many fewer than the lower bound in Theorem 1. We would also like to "de-couple" n and q, which are tied together in the lemma.

THEOREM 2. *Let p and r be positive integers with $r > p$. For $q \geqq q_0(p,r)$ and $n \geqq n_0(r)$, there is a constant $c_1 = c_1(p,r)$ and there is a simple r-uniform hypergraph H on n vertices and $e(H)$ edges such that*

$$e(H) \geqq c_1 \frac{q^{r-p}}{(\log q)^{r-p-1}} n \tag{10}$$

and Breaker has a winning strategy for $\mathbf{G}(p,q,H)$.

PROOF. Let p and r be given. It follows from a well known result of Wilson [6] that there is an integer $l_0(r)$ such that for all $l \geqq l_0(r)$ there exists a simple r-uniform hypergraph F_l on l vertices and with at least l^2/r^2 edges. Set $n_0 = 2l_0$ and

$$q_0 = \left\lceil c_0 l_0^{\frac{1}{r-p}} (\log l_0)^{\frac{r-p-1}{r-p}} \right\rceil,$$

where $c_0 = c_0(p,r)$ is the constant in Lemma 1.

Suppose that $n \geqq n_0$ and $q \geqq q_0$ are given. Let $l \geqq l_0$ be the largest integer such that

$$q \geqq \left\lceil c_0 l^{\frac{1}{r-p}} (\log l)^{\frac{r-p-1}{r-p}} \right\rceil \geqq q_0. \tag{11}$$

Hence,

$$l \geqq \left\lceil c' \frac{q^{r-p}}{(\log q)^{r-p-1}} \right\rceil, \tag{12}$$

where $c' = c'(p,r)$.

Observe that if $n < l$ then (11) and Lemma 1 show that Breaker has a winning strategy in $\mathbf{G}(p,q,H)$ for any simple r-uniform hypergraph H on n vertices. If $l < n < 2l$ we can define H to be the n-vertex hypergraph obtained from F_l by adding $2l - n$ isolated vertices. Then Lemma 1 shows that Breaker has a winning strategy for $\mathbf{G}(p,q,H)$. So, assume that $n \geqq 2l$ and let H be the simple r-uniform hypergraph on n vertices comprised of $\lfloor n/l \rfloor$ disjoint copies of F_l and the required number of isolated vertices. Then

$$e(H) \geqq \left\lfloor \frac{n}{l} \right\rfloor \frac{l^2}{r^2} \geqq \left(\frac{n}{l} - 1 \right) \frac{l^2}{r^2} \geqq \frac{nl}{2r^2}. \tag{13}$$

Apply (12) to (13) and conclude that

$$e(H) \geqq c_1 \frac{q^{r-p}}{(\log q)^{r-p-1}} n.$$

Breaker's winning strategy for $\mathbf{G}(p,q,H)$ is this: Breaker pursues its strategy for $\mathbf{G}(p,q,F)$ in whichever component of H in which Maker has just moved. □

4. A problem of Galvin and related questions

We became interested in the games $\mathbf{G}(p,q,H)$ because of this problem due to Galvin [5].

PROBLEM 1. Do there exist 3-uniform hypergraphs H of arbitrarily high chromatic number for which Breaker can win $\mathbf{G}(1,2,H)$.

The answer is no; this can be deduced from [1, Theorem 2] with a little work.

THEOREM 3. *Let H be a 3-uniform hypergraph with chromatic number $\chi(H) \geqq (6q^3+1)(q+1)$. For all $q \geqq 2$, Maker has winning strategy for* $\mathbf{G}(1,q,H)$.

PROOF. For purposes of this proof, call a pair $\{v,w\}$ of distinct vertices of H *thick* if $\{v, w\}$ is contained in at least $2q+1$ edges of H. Let G be the graph with vertex set $V(H)$ and edges all thick (unordered) pairs.

Case 1. There is a vertex v of degree at least $q+1$ in G.

Maker chooses v in the first round and an unchosen neighbor w in the second. Breaker cannot prevent Maker from choosing a vertex x in the third round such that $v,\ w,\ x \in E(H)$ since there are at least $2q+1$ edges of H containing v and w.

Case 2. The maximum degree of G is at most q.

Let H' be the hypergraph with vertex set $V(H)$ and with edges only those edges of H which contain at least one thick pair. Then $\chi(H') \leqq q+1$ since $\chi(G) \leqq q+1$ and any good coloring of G is a good coloring of H'.

Let H'' have vertex set $V(H)$ and edge set $E(H'') = E(H) - E(H')$. Then

$$\chi(H'') \geqq \frac{\chi(H)}{\chi(H')} \geqq \frac{(6q^3+1)(q+1)}{q+1} = 6q^3+1.$$

There is a subhypergraph S of H'' such that $\chi(S) = \chi(H'')$ and

$$3\frac{e(S)}{v(S)} \geqq \chi(S) - 1 \geqq 6q^3. \tag{14}$$

Now, specialize Equation (1) to a 3-uniform hypergraph S with $p=1$: provided that

$$e(S) > q^2 d(S) v(S), \tag{15}$$

Maker wins $\mathbf{G}(1,q,S)$. But $d(S) \leqq 2q$, so Equations (14) and (15) show that Maker has a winning strategy for $\mathbf{G}(1,q,S)$ and, hence, for $\mathbf{G}(1,q,H)$. □

Note that we give something up in our argument: one can do better than the inequality in (14) and decrease the lower bound $(6q^3+1)(q+1)$ in Theorem 3 somewhat.

In the same vein as Galvin's problems relating chromatic number of a hypergraph H to the game $\mathbf{G}(1,q,H)$, we ask about the following game.

Given $q \geqq 2$, $k \geqq 3$, and a hypergraph H, let $\mathbf{Chr}_k(1,q,H)$ in which Maker and Breaker alternate selection of as yet unselected vertices, Maker first choosing one vertex, followed by Breaker selecting q vertices. Maker wins if the hypergraph induced on its vertices has chromatic number at least k; Breaker wins if it has a strategy to prevent this.

PROBLEM 2. Given integers $k \geqq 3$ and $r, q \geqq 2$, is there some K such that for all r-uniform hypergraphs H with $\chi(H) \geqq K$, Maker has a winning strategy for $\mathbf{Chr}_k(1,q,H)$?

There is a particularly enticing special case of Problem 2.

PROBLEM 3. Is there an integer K such that for all graphs G with $\chi(G) \geqq K$, Maker has a strategy to choose an odd cycle in the game, where Maker chooses one vertex and Breaker two, in each round?

REFERENCES

[1] BECK, J., Remarks on positional games. I, *Acta Math. Acad. Sci. Hungar.* **40** (1982), 65–71. *MR* **85c**:90121

[2] CHVÁTAL, V. and ERDŐS, P., Biased positional games, *Ann. Discrete Math.* **2** (1978), 221–229. *MR* **81a**:05107

[3] CSMIRAZ[!], L., On a combinatorial game with an application to Go-moku, *Discrete Math.* **29** (1980), 19–23. *MR* **81a**:05013

[4] ERDŐS, P. and SELFRIDGE, J., On a combinatorial game, *J. Combinatorial Theory Ser. A* **14** (1973), 298–301. *MR* **48** #5655

[5] GALVIN, F., Hypergraph games and the chromatic number, *A tribute to Paul Erdős* (A. Baker, B. Bollobás, and A. Hajnal, Eds.), Cambridge Univ. Press, Cambridge, 1990, 201–206. *MR* **92i**:05093

[6] WILSON, R. M., An existence theory for pairwise balanced designs. II. The structure of PBD-closed sets and the existence conjectures, *J. Combinatorial Theory Ser. A* **13** (1972), 246–273. *MR* **46** #3339

(Received November 17, 1997)

MATHEMATICS AND COMPUTER SCIENCE DEPARTMENT
EMORY UNIVERSITY
ATLANTA, GA 30322
U.S.A.

dwight@mathcs.emory.edu

DEPARTMENT OF DISCRETE MATHEMATICS
ADAM MICKIEWICZ UNIVERSITY
POZNAN
POLAND

and

MATHEMATICS AND COMPUTER SCIENCE DEPARTMENT
EMORY UNIVERSITY
ATLANTA, GA 30322
U.S.A.

tomasz@mathcs.emory.edu

MATHEMATICS AND COMPUTER SCIENCE DEPARTMENT
EMORY UNIVERSITY
ATLANTA, GA 30322
U.S.A.

rodl@mathcs.emory.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 151–164

STATIONARY STATES OF INTERACTING BROWNIAN MOTIONS

J. FRITZ, S. ROELLY and H. ZESSIN

In Memoriam Alfréd Rényi

Abstract

We are interested in a description of stationary states of gradient dynamics of interacting Brownian particles. In contrast to lattice models, this problem does not seem to be solvable at a formal level of the stationary Kolmogorov equation. We can only study stationary states of a well controlled Markov process. In space dimensions four or less, for smooth and superstable pair potentials of finite range the non-equilibrium dynamics of interacting Brownian particles can be constructed in an explicitly defined deterministic set of locally finite configurations, see [Fr2]. This set is of full measure with respect to any canonical Gibbs state for the interaction, and every canonical state is a stationary one. Assuming translation invariance of a stationary measure, and also the finiteness of its specific entropy with respect to an equilibrium Gibbs state, it is shown that this stationary state is canonical Gibbs. Related ideas of Alfréd Rényi and some of their consequences are also reviewed.

1. Introduction

The main purpose of this paper is to identify a class of stationary states of the following system of interacting particles as the set of translation invariant canonical Gibbs states with interaction U. The evolution law is given by an infinite system of stochastic differential equations,

$$d\omega_k = -\frac{1}{2}\sum_{j\neq k} \operatorname{grad} U(\omega_k-\omega_j)\,dt + dw_k, \quad \omega_k(0)=\sigma_k, \quad k\in S, \tag{1.1}$$

where S is a countable index set, $w=(w_k)_{k\in S}$ is a family of independent standard d-dimensional Wiener processes, and each $\omega_k=\omega_k(t), t\geqq 0$ is assumed to be a continuous trajectory in $\mathbb{R}^d$. The potential $U:\mathbb{R}^d\mapsto\mathbb{R}$ is symmetric and superstable with finite range, that is $U(x)=U(-x)$, there is an $R>0$ such that $U(x)=0$ if $|x|>R$, and we have constants $A\geqq 0$, $B>0$ such that

1991 *Mathematics Subject Classification.* Primary 60K35, 60J35.
Key words and phrases. Gibbs measures, stochastic gradient systems, reversible Markov processes, relative entropy.

This work was supported in parts by the French-Hungarian project Balaton F-27/96, and by Hungarian National Science Foundation Grant T-16665.

for any finite sequence $q_1, q_2, \ldots, q_n$ of not necessarily distinct points from $\mathbb{R}^d$

$$nA + \sum_{k=1}^{n} \sum_{j \neq k} U(q_k - q_j) \geqq BN, \tag{1.2}$$

where N is the number of pairs $\{j, k\}$ such that $|q_k - q_j| \leqq R$, see [Ru1]. Let Ω denote the set of configurations $\omega = (\omega_k)_{k \in S}$ having no limit points. Although the right-hand side of (1.1) is certainly well defined for such, locally finite configurations, to develop a satisfactory existence theory we have to restrict the configuration space in a much more radical way. On the other hand, the set of allowed configurations should be large enough to support a possibly wide set of probability measures including Gibbs states with various interactions.

The first mathematical results concerning this model go back to R. Lang, see [La1] and [La2], where the existence of equilibrium dynamics, and also the canonical Gibbs property of reversible measures is proven. These dynamics are defined almost surely with respect to a Gibbs state with interaction U, see also the more sophisticated argument of [Os]. For a study of stationary measures in general, we need a more direct construction involving explicit bounds on the rate of convergence of solutions to finite subsystems (partial dynamics) when the number of active particles tends to infinity, see Section 3 below. Indeed, the problem of stationary measures cannot be solved at a formal level of the stationary Kolmogorov equation because a full Hille–Yoshida theory is not available in the present context. Since we do not know any Banach space in which the underlying Markov semigroup is strongly continuous, we have to materialize our arguments at a level of finite dimensional approximations, see [FFL] and [FLO] for a discussion of related questions.

For a generic, locally finite configuration $\omega = (\omega_k)_{k \in S}$ let $H(\omega, m, r)$ denote total energy in the ball $B(m, r)$ of center $m \in \mathbb{R}^d$ and radius $r \geqq 1$, and for $\alpha \geqq 0$ define

$$\overline{H}_\alpha(\omega) := \sup_{m \in \mathbb{Z}^d} \sup_{r \in \mathbb{N}} \frac{H(\omega, m, r g_\alpha^{1/d}(m))}{r^d g_\alpha(m)} \quad \text{where}$$

(1.3)

$$H(\omega, m, r) := \frac{1}{2} \sum_{k : \omega_k \in B(m,r)} \quad \sum_{j \neq k : \omega_j \in B(m,r)} \quad U(\omega_j - \omega_k)$$

and $g_\alpha(u) := 1 + |u|^\alpha \log(1 + |u|)$ for $u \in \mathbb{R}, \mathbb{R}^d$. The set of allowed configurations is now specified as $\overline{\Omega}_\alpha := \{\omega \in \Omega : \overline{H}_\alpha(\omega) < +\infty\}$; we shall see that for an effective a priori bound we need $\alpha \leqq 2 - d/2$, thus $d \leqq 4$. Let $C_0(\mathbb{R}^d)$ denote the space of continuous $\varphi : \mathbb{R}^d \mapsto \mathbb{R}$ of compact support. Spaces of k times continuously differentiable functions with compact supports are marked by a superscript k, while a subscript b in place of 0 refers to bounded functions without any support condition. For an open and bounded domain $\Lambda \subset \mathbb{R}^d$ the σ-field $\mathcal{F}_\Lambda$ is generated by the variables $\omega(\varphi) := \sum_{k\in S} \varphi(\omega_k)$ such that the support of $\varphi \in C_0(\mathbb{R}^d)$ is contained in Λ; the number of points in Λ will be denoted by $\omega(\Lambda)$. This means that configurations are interpreted as nonnegative, integer valued measures, and $\overline{\Omega}_\alpha$ is equipped with the associated weak topology and Borel structure. Observe that, due to superstability (1.2), the level sets $\overline{\Omega}_{\alpha,h} := [H_\alpha(\omega) \leq h]$ are compact if h is large enough. The restriction of $\omega \in \overline{\Omega}_\alpha$ to Λ is ω_Λ, and Λ^c denotes the complement of Λ.

For any bounded domain $\Lambda \subset \mathbb{R}^d$, $\sigma \in \overline{\Omega}_\alpha$ and $n \in \mathbb{N}$ let $\Omega_\Lambda(n|\sigma)$ denote the set of $\omega \in \overline{\Omega}_\alpha$ such that $\omega(\Lambda) = n$ and $\omega_{\Lambda^c} = \sigma_{\Lambda^c}$. A probability measure λ is a canonical Gibbs state (with unit temperature) for U if its conditional distribution $\lambda[d\omega_\Lambda|\omega_{\Lambda^c}=\sigma_{\Lambda^c}, \omega(\Lambda)=n]$, given the configuration outside of Λ and the number of points in Λ, admits an nd-dimensional Lebesgue density $f_{\Lambda,n}$,

$$
\begin{aligned}
&f_{\Lambda,n}(\omega|\sigma) := \frac{\exp(-H_{\Lambda,n}(\omega|\sigma))}{Z_{\Lambda,n}(\sigma)} \quad \text{if } \omega \in \Omega_\Lambda(n|\sigma), \text{ where} \\
&H_{\Lambda,n}(\omega|\sigma) := \frac{1}{2} \sum_{k:\omega_k\in\Lambda} \sum_{j\neq k:\omega_j\in\Lambda} U(\omega_k - \omega_j) + \sum_{\omega_k\in\Lambda} \sum_{\sigma_j\in\Lambda^c} U(\omega_k - \sigma_j)
\end{aligned}
\tag{1.4}
$$

and Z is the canonical partition function (normalization). Gibbs states are the extremal canonical measures, see e.g. [Geo]. In view of the superstability estimates of [R2], there exists at least one translation invariant Gibbs state λ such that $\lambda(\overline{\Omega}_0) = 1$, of course $\overline{\Omega}_\alpha \subset \overline{\Omega}_\beta$ if $\alpha < \beta$.

The unique strong solution $\boldsymbol{\omega} = \omega(t,\sigma)$ to the infinite system (1.1) with initial configuration $\sigma \in \overline{\Omega}_\alpha$ is constructed as the a.s limit of partial solutions $\boldsymbol{\omega}^\theta = \omega^\theta(t,\sigma)$ when a spatial cutoff θ is removed. Partial (approximate) solutions are constructed in such a way that particles are frozen outside of a bounded region, they follow (1.1) in the central part of this domain, and there is a continuous transition from a full activity to a vanishing one at the boundary. It is relevant that partial dynamics preserve any canonical Gibbs measure. More precisely, for any $\theta \in C_0^2(\mathbb{R}^d)$ with $0 \leq \theta \leq 1$ there is a differential operator $\mathcal{L}_\theta$,

$$
\mathcal{L}_\theta \phi := \frac{1}{2} \sum_{k\in S} \sum_{i=1}^{d} e^{H_k(\omega)} \partial_{k,i}(\theta(\omega_k) e^{-H_k(\omega)} \partial_{k,i}\phi(\omega)), \tag{1.5}
$$

where $\partial_{k,i}$ denotes differentiation with respect to the i coordinate of ω_k and

$$
H_k(\omega) := \sum_{j\neq k} U(\omega_j - \omega_k). \tag{1.6}
$$

We consider $\mathcal{L}_\theta$ as the (formal) generator of partial dynamics with cutoff θ, the infinite system (1.1) corresponds to $\theta \equiv 1$. All generators of this kind are certainly well defined on $C_0^2(\Omega)$, where $C_0^k(\Omega)$ is the space of test functions

$$(1.7)\quad \phi(\omega) = \psi(\omega(\varphi_1), \omega(\varphi_2), \ldots, \omega(\varphi_l)), \quad \psi \in C_b^k(\mathbb{R}^l),\ \varphi_j \in C_0^k(\mathbb{R}^d),\ l \in \mathbb{N}.$$

The stochastic equations for cutoff θ read as

$$(1.8)\qquad d\omega_k = \tfrac{1}{2}\, e^{H_k}\, \partial_k(\theta(\omega_k)e^{-H_k})\, dt + \sqrt{\theta(\omega_k)}\, dw_k,$$

they have a unique strong solution $\boldsymbol{\omega}^\theta = \omega^\theta(t,\sigma)$ for each initial configuration $\sigma \in \Omega$. If $\Lambda \supset \operatorname{supp}\theta$ then the particle number in Λ is a constant of motion, that is $\omega^\theta(t,\sigma)(\Lambda) \equiv \sigma(\Lambda)$. Therefore (1.8) defines a fairly regular diffusion in each $\Omega_\Lambda(n|\sigma)$, and it is easy to verify that realizations of the canonical conditional distribution $\lambda[d\omega_\Lambda | \omega_{\Lambda^c} = \sigma_{\Lambda^c}, \omega(\Lambda) = n]$, $n \in \mathbb{N}$, $\sigma \in \Omega$ are all reversible measures of the associated (nd-dimensional) diffusion process. The associated Markov semigroup will be denoted as $\mathcal{P}_\theta^t$, it is strongly continuous in the Banach space $C_b(\overline{\Omega}_\alpha)$ of continuous and bounded $\phi : \overline{\Omega}_\alpha \mapsto \mathbb{R}$, and also in $L^2(\overline{\Omega}_\alpha, \lambda)$ whenever λ is a canonical Gibbs state.

In the paper [Fr2] it is shown that if $d \leqq 4$ then for every initial configuration $\sigma \in \overline{\Omega}_0$ the sequence of partial solutions $\omega^\theta(t,\sigma)$ converges almost surely to a strong solution $\boldsymbol{\omega} = \omega(t,\sigma)$ of (1.1) as $\theta \to 1$ in a clever way. This limiting solution is distinguished by an a priori bound: $\overline{H}_0(\omega(t,\sigma))$ is bounded on finite intervals of time, and there is no other solution having this property. Following the lines of the proof we see that the result extends immediately to all $\alpha \leqq 2 - d/2$, see also Proposition 1 in Section 3. Since the rate of convergence of partial solutions does depend on $\overline{H}_\alpha(\sigma)$, the limiting semigroup, $\mathcal{P}^t$, is not strongly continuous in $C_b(\overline{\Omega}_\alpha)$, thus the Hille–Yoshida theory is available in a restricted form only.

As a general reference measure we choose a translation invariant Gibbs state λ with interaction U and unit temperature, it is also a reversible measure of each partial dynamics. Introduce $F_\lambda(\phi) := \log \lambda(e^\phi)$, then entropy of another probability measure μ relative to λ is just its convex conjugate $I[\mu|\lambda]$,

$$(1.9)\qquad I[\mu|\lambda] := \sup_\phi \{\mu(\phi) - F_\lambda(\phi) : \phi \in C_0(\Omega)\} = \int \log \frac{d\mu}{d\lambda}\, d\mu$$

if $\mu \ll \lambda$; $I[\mu|\lambda] = +\infty$ otherwise. It is easy to verify that $\mu(\phi) \leqq I[\mu|\lambda] + F_\lambda(\phi)$ whenever $\phi : \overline{\Omega}_\alpha \mapsto \mathbb{R}$ is measurable and $\mu(\phi) < +\infty$. The entropy of μ in $\Lambda \subset \mathbb{R}^d$ is

$$(1.10)\quad I_\Lambda[\mu|\lambda] := I[\mu_\Lambda|\lambda_\Lambda] = I[\mu_\Lambda\lambda|\lambda] = \sup_\phi \{\mu(\phi) - F_\lambda(\phi) : \phi \in \mathcal{F}_\Lambda \cap C_0(\Omega)\},$$

where μ_Λ is the restriction of μ to $\mathcal{F}_\Lambda$ and $\mu_\Lambda\lambda$ is the measure obtained by extending μ_Λ to the whole space by means of the conditional distribution of λ, that is $(\mu_\Lambda\lambda)(d\omega) := \lambda(d\omega_{\Lambda^c}|\omega_\Lambda)\mu_\Lambda(d\omega_\Lambda)$. If μ is translation invariant and Λ_n denotes the centered cubic box of side $2n$ then

$$\begin{aligned} \bar{I}[\mu|\lambda] &:= \lim_{n\to\infty} \frac{I_{\Lambda_n}[\mu|\lambda]}{|\Lambda_n|} = \sup_\phi\{\mu(\phi) - \overline{F}_\lambda(\phi) : \phi \in C_0(\Omega)\} \\ \overline{F}_\lambda(\phi) &:= \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \log \int \exp\Big(\sum_{m\in\Lambda_n\cap\mathbb{Z}^d} \mathrm{s}^m\phi\Big)\, d\lambda, \end{aligned} \tag{1.11}$$

denotes the (relative) specific entropy of μ, see Section 5 in [OVY]. Here and also later on, s^m is the shift by $m\in\mathbb{R}^d$, i.e. $\mathrm{s}^m\phi(\omega)\equiv\phi(\mathrm{s}^m\omega)$. Observe that $\bar{I}[\mu|\lambda] < +\infty$ implies $\mu(\overline{\Omega}_d) = 1$ by the ergodic theorem. Our main result is the following:

THEOREM 1. *Suppose that μ^* is a translation invariant stationary distribution of the infinite system* (1.1), *that is $\mu^*(\overline{\Omega}_\alpha) = 1$ for some $0 \leq \alpha \leq 2 - 2/d$ by assumption. If $\bar{I}[\mu^*|\lambda] < +\infty$ then μ^* is a canonical Gibbs state of unit temperature with interaction U.*

The starting point of the argument is a quite general entropy inequality for Markov processes in such situations when the initial distribution has finite entropy relative to a stationary reference measure, see e.g. [FLO]. This inequality and some of its first consequences are discussed in the next section. In Section 3 we develop some uniform estimates on the rate of convergence of partial dynamics to the full (infinite) one. These bounds are then used in Section 4 to extend the basic entropy inequality to the infinite system, which completes the proof.

2. An entropy inequality and its consequences

The idea that relative entropy with respect to a stationary measure is a nice and effective tool of the study of ergodic properties of Markov processes goes back to A. Rényi [Re1, Re2], where ergodicity of irreducible Markov chains in a finite state space is shown by using entropy as a Liapunov function to show the convergence of the evolved measure. Let us first review this argument in a general context of discrete time Markov processes in a probability space $(X, \mathcal{X}, \lambda)$; see e.g. [Fo] for basic notions and results. Let $\mathcal{P}$ denote a positive contraction of $L^\infty(\lambda)$ into itself, it is interpreted as the operator of conditional expectation of the underlying Markov process. If $\mu \ll \lambda$ is a probability measure on $(X, \mathcal{X})$ then $\mu\mathcal{P}$ is defined by $\mu\mathcal{P}(\varphi) = \mu(\mathcal{P}\varphi)$ for $\varphi \in L^\infty(\lambda)$; given an initial distribution $\mu_0 = \mu \ll \lambda$, the evolved measure at time $t \in \mathbb{N}$ is denoted as $\mu_t = \mu\mathcal{P}^t$; $c\mathcal{P}^t\varphi := \mathcal{P}^{t-1}\psi$ if $\psi = \mathcal{P}\varphi$. We are assuming that $\lambda = \lambda\mathcal{P}$ is a stationary measure, then $\mathcal{P}$ is a contraction of each

$L^p(\lambda)$ space with $1 \leqq p \leqq +\infty$, and $I[\mu\mathcal{P}|\lambda] \leqq I[\mu|\lambda]$ by convexity. Moreover, as noticed by I. Csiszár [Cs], the difference is again a relative entropy:

$$I[\mu|\lambda] - I[\mu_t|\lambda] = I[\mu \circ \mathcal{P}^t | \mathcal{Q}^t \circ \mu_t], \tag{2.1}$$

where $\mu \circ \mathcal{P}$ and $\mathcal{Q} \circ \mu$ are probability measures on $X \times X$ characterized by

$$\begin{aligned}(\mu \circ \mathcal{P})(\phi) &= \int \mu(dx)\varphi(x)\mathcal{P}\psi(x) \quad \text{and} \\ (\mathcal{Q} \circ \mu)(\phi) &= \int \mu(dy)\psi(y)\mathcal{Q}\varphi(y)\end{aligned} \tag{2.2}$$

for $\phi(x,y) = \varphi(x)\psi(y)$ with $\varphi, \phi \in L^\infty(\lambda)$. Here and later on, $\mathcal{Q}$ denotes the transition operator of the backward process reversed with respect to λ; it is just the adjoint of $\mathcal{P}$ in $L^2(\lambda)$, i.e. $\lambda(\varphi\mathcal{P}\psi) = \lambda(\psi\mathcal{Q}\varphi)$ for $\varphi, \psi \in L^2(\lambda)$. In general we do not know that $\mathcal{Q}$ is given by a transition probability, but it is again a positive contraction of $L^2(\lambda)$, thus $\mu\mathcal{Q} := \mu(\mathcal{Q}\varphi), \varphi \in L^2(\lambda)$ is a probability measure if μ is so, $\lambda = \lambda\mathcal{Q}$. In view of (2.2), $I[\mu|\lambda] = I[\mu\mathcal{P}|\lambda] < +\infty$ implies $\mu \circ \mathcal{P} = \mathcal{Q} \circ \mu\mathcal{P}$, thus μ is a stationary and reversible measure of the composed, reversible process $\mathcal{R} := \mathcal{P}\mathcal{Q}$, see [Fr1]. Of course, $\mathcal{R}^t \neq \mathcal{P}^t\mathcal{Q}^t$ in general, because $\mathcal{P}$ and $\mathcal{Q}$ need not commute. Nevertheless, the following reformulation of results by Rényi and Csiszár demonstrates an intrinsic relationship of the notions of entropy and reversibility.

THEOREM 2. *Every absolutely continuous stationary measure $\bar{\mu} \ll \lambda$, $\bar{\mu} = \bar{\mu}\mathcal{P}$, is reversible with respect to $\mathcal{R}$. If $\mu \ll \lambda$ then so is $\mu\mathcal{P}^t$, and the sequence of densities, $f_t := d\mu\mathcal{P}^t/d\lambda$ is uniformly integrable with respect to λ. Moreover, if $\mu\mathcal{P}^{t_n}(\varphi) \to \bar{\mu}(\varphi)$ for all $\varphi \in L^\infty(\lambda)$ as $t_n \to +\infty$ then $\bar{\mu}$ is a reversible measure of $\mathcal{R}$, that is we have a weak convergence of the evolved state to the set of $\mathcal{R}$-reversible measures.*

PROOF. Suppose first that $I[\mu|\lambda] < +\infty$, then $I[\mu\mathcal{P}^t|\lambda] \leqq I[\mu|\lambda]$ implies the uniform integrability of $f_t, t \in \mathbb{N}$, thus the Dunford–Pettis Theorem applies. We have to show that every weak limit point $\bar{\mu}$ satisfies $I[\bar{\mu}|\lambda] = I[\bar{\mu}\mathcal{P}|\lambda]$.

If $\bar{\mu}(\varphi) = \lim_n \mu\mathcal{P}^{t_n}(\varphi)$ for all $\varphi \in L^\infty(\lambda)$ and $\phi: X \times X \mapsto \mathbb{R}$ is measurable and bounded, then

$$\begin{aligned}(\bar{\mu} \circ \mathcal{P})(\phi) - \log(\mathcal{Q} \circ \bar{\mu})(e^\phi) &= \lim_{n\to\infty} (\mu_{t_n} \circ \mathcal{P})(\phi) - \log(\mathcal{Q} \circ \mu_{t_n+1})(e^\phi) \\ &\leqq \lim_{n\to\infty} (I[\mu_{t_n}|\lambda] - I[\mu_{t_n+1}|\lambda]) = 0.\end{aligned} \tag{2.3}$$

Taking the supremum on the left-hand side we get $I[\bar{\mu} \circ \mathcal{P} | \mathcal{Q} \circ \bar{\mu}\mathcal{P}] = 0$, whence $\bar{\mu} \circ \mathcal{P} = \mathcal{Q} \circ \bar{\mu}\mathcal{P}$, i.e. $\bar{\mu} = \bar{\mu}\mathcal{R}$. Replacing $\mathcal{P}$ by $\mathcal{R}$ in the argument above, we get $\bar{\mu} \circ \mathcal{R} = \mathcal{R} \circ \bar{\mu}$, the condition of reversibility of $\bar{\mu}$ with respect to $\mathcal{R} = \mathcal{P}\mathcal{Q}$.

The general case of $\mu \ll \lambda$ follows by a direct approximation procedure. For each $\varepsilon > 0$ we have some μ^ε such that $I[\mu^\varepsilon|\lambda] < +\infty$ and $|\mu - \mu^\varepsilon|_1 < \varepsilon$, where $|\cdot|_1$ denotes the variational distance. Set $f_t^\varepsilon := d\mu^\varepsilon\mathcal{P}^t/d\lambda$ and $|x|_+ := \max\{0, x\}$; since $\mathcal{P}$ is a contraction of $L^1(\lambda)$,

$$\int |f_t - a|_+ \, d\lambda \leqq \int |f_t^\varepsilon - a|_+ \, d\lambda + \int |f_t - f_t^\varepsilon| \, d\lambda$$
$$\leqq \int |f_0^\varepsilon - a|_+ \, d\lambda + \int |f_0 - f_0^\varepsilon| \, d\lambda \leqq 2\varepsilon$$

if a is large enough, thus f_t is still a uniformly integrable sequence. Consider now a weak limit point $\bar{\mu}$ of $\mu\mathcal{P}^t$, $t_n \to +\infty$ is the subsequence along which $\mu\mathcal{P}^t$ converges to $\bar{\mu}$, and let $\bar{\mu}^\varepsilon$ denote a limit point of $\mu^\varepsilon\mathcal{P}^{t_n}$. We have a subsequence $\{t_n'\} \subset \{t_n\}$ such that for any $\varphi \in L^\infty(\lambda)$

$$|\bar{\mu}(\varphi) - \bar{\mu}^\varepsilon(\varphi)| = \lim_{n\to\infty} |\mu\mathcal{P}^{t_n'}(\varphi) - \mu^\varepsilon\mathcal{P}^{t_n'}(\varphi)| \leqq \varepsilon \sup_{x\in X} |\varphi(x)|,$$

so that $|\bar{\mu} - \bar{\mu}^\varepsilon|_1 \leqq \varepsilon$ implying $\bar{\mu}(\varphi\mathcal{R}\psi) = \bar{\mu}(\psi\mathcal{R}\varphi)$ for $\varphi, \psi \in L^\infty(\lambda)$. □

This result is useful because usually it is easier to identify the reversible measures than the stationary ones. Of course, the set of reversible measures of $\mathcal{R} = \mathcal{P}\mathcal{Q}$ can be much larger than the set of stationary measures of $\mathcal{P}$. Anyway, Theorem 2 yields some preliminary information for further, more specific investigations.

For example, if X is a countable set then $\mathcal{P}$ is given by a stochastic matrix $p = p(x, y)$, i.e. $\mathcal{P}\varphi(x) = \sum_y p(x, y)\varphi(y)$, then the associated backward transition probability is just $q(y, x) := \lambda(x)p(x, y)/\lambda(y)$; $\lambda(x) > 0$ for all $x \in X$ may be assumed. From (2.3) with $\mathcal{P}^t$ in place of $\mathcal{P}$ we get $\bar{\mu} \circ \mathcal{P}^t = \mathcal{Q}^t \circ \bar{\mu}\mathcal{P}^t$ for any limit distribution $\bar{\mu}$, which reads as

$$\frac{\bar{\mu}(x)}{\lambda(x)} p^t(x, y) = p^t(x, y) \frac{\bar{\mu}_t(y)}{\lambda(y)} \tag{2.4}$$

in the present context. Therefore if the chain is aperiodic in the sense that for each $x \in X$ there exists an integer $t(x) > 0$ such that $p^t(x, x) > 0$ whenever $t \geqq t(x)$, then $\bar{\mu}(x) = \bar{\mu}_t(x)$ for $t \geqq t(x)$. Similarly, $\bar{\mu}_1(x) = \bar{\mu}_{t+1}(x)$ if $t \geqq t(x)$, consequently $\bar{\mu}(x) = \bar{\mu}_1(x)$ for all $x \in X$, i.e. $\bar{\mu} = \bar{\mu}\mathcal{P}$. The uniqueness of the stationary measure follows immediately from a condition of irreducibility: if for each pair $x, y \in X$ we have some $t = t(x, y)$ such that $p^t(x, y) > 0$ then we get $\bar{\mu}(x)/\lambda(x) = \bar{\mu}(y)/\lambda(y)$, whence $\bar{\mu}(x) = \lambda(x)$ for all $x \in X$, consequently we have $\mu_t(x) \to \lambda(x)$ for all $x \in X$ as $t \to \infty$.

In the case of continuous time it is natural to assume that X is a complete and separable metric space, and both $\mathcal{P}^t$ and its adjoint $\mathcal{Q}^t$ form strongly continuous contraction semigroups in $L^2(\lambda)$ and also in $C_b(X)$; basic notations are the same as above. To obtain a lower bound for $I[\mu|\lambda] - I[\mu\mathcal{P}^t|\lambda]$, consider an auxiliary distribution $\nu \ll \lambda$ such that $\psi := d\nu/d\lambda > 0$; then $\mu \ll \nu$

and $I[\mu|\nu] = I[\mu|\lambda] - \nu(\log\psi)$, while $I[\mu\mathcal{P}^t|\nu\mathcal{P}^t] = I[\mu\mathcal{P}^t|\lambda] - \mu(\log\mathcal{Q}^t\psi)$ as $d\nu\mathcal{P}^t/d\lambda = \mathcal{Q}^t\psi$. Since $I[\mu\mathcal{P}^t|\nu\mathcal{P}^t] \leqq I[\mu|\nu]$ by convexity,

$$I[\mu|\lambda] - I[\mu\mathcal{P}^t|\lambda] \geqq \mu(\log\psi) - \mu\mathcal{P}^t(\log\mathcal{Q}^t\psi) \geqq \mu(\log\psi) - \mu(\log\mathcal{R}^t\psi) \geqq \int \frac{\psi - \mathcal{R}^t\psi}{\psi}\,d\mu \tag{2.5}$$

as $\log x - \log y \geqq (x-y)/x$. Observe that $\lambda\circ\mathcal{R}^t$ is a symmetric measure, thus with $f = d\mu/d\lambda$ we get

$$\int \frac{\mathcal{R}^t\psi}{\psi}\,d\mu = \frac{1}{2}\iint (\lambda\circ\mathcal{R}^t)(dx,dy)\Big(\frac{f(x)\psi(y)}{\psi(x)} + \frac{f(y)\psi(x)}{\psi(y)}\Big) \geqq \iint (\lambda\circ\mathcal{R}^t)(dx,dy)\sqrt{f(x)}\sqrt{f(y)}. \tag{2.6}$$

This means that the right-hand side is maximal if $\psi = \sqrt{f}$, consequently

$$\sup_{\psi>0}\Big\{\int \frac{\psi - \mathcal{R}^t\psi}{\psi}\,d\mu : \psi\in L^2(\lambda)\Big\} = \int \sqrt{f(x)}\big(\sqrt{f(x)} - \mathcal{R}^t\sqrt{f}(x)\big)\,\lambda(dx) \tag{2.7}$$

whenever $\mu \ll \lambda$ and $d\mu = f\,d\lambda$.

Consider now the Donsker–Varadhan rate function D, it is obtained by differentiating (2.7) with respect to time:

$$D[\mu|\mathcal{G}] := \sup_{\psi}\Big\{-\int \frac{\mathcal{G}\psi}{\psi}\,d\mu : \psi\in \mathrm{Dom}\ \mathcal{G},\ \inf\psi > 0\Big\}, \tag{2.8}$$

where $\mathcal{G}$ is any semigroup generator. Remember that Dom $\mathcal{G}$ in the definition of D can be replaced by any core of $\mathcal{G}$ in $C_b(X)$. Moreover, if $\mathcal{G}$ is self-adjoint in $L^2(\lambda)$ and $f = d\mu/d\lambda$, then $D[\mu|\mathcal{G}] < +\infty$ implies $\sqrt{f}\in \mathrm{Dom}\ (-\mathcal{G})^{1/2}$ and

$$D[\mu|\mathcal{G}] = \int (\sqrt{-\mathcal{G}}\sqrt{f})^2\,d\lambda \tag{2.9}$$

see (2.7) and Theorem 5 in [DV].

Let $\mathcal{L}$ and $\mathcal{L}^*$ denote the generators of $\mathcal{P}^t$ and $\mathcal{Q}^t$ in $L^2(\lambda)$, respectively. Although $\mathcal{R}^t = \mathcal{P}^t\mathcal{Q}^t$ does not form a semigroup, it is self-adjoint, and $\mathcal{R}^t\phi = t\mathcal{G}\phi + o(t)$ for small t by a formal calculation, where $\mathcal{G} = \mathcal{L} + \mathcal{L}^*$. Therefore, under some natural conditions the right-hand side of (2.4) becomes $-t\mu(\mathcal{G}\psi/\psi) + o(t)$, thus we have

PROPOSITION 1. *Suppose that $\mathcal{G} = \mathcal{L} + \mathcal{L}^*$ is self-adjoint in $L^2(\lambda)$, and we have a dense linear space $C^* \subset C_b(X)$ such that $\mathcal{P}^tC^* \subset C^*$ and $\mathcal{Q}^tC^* \subset C^*$ for all $t > 0$, then*

$$I[\mu\mathcal{P}^t|\lambda] + 2tD[\bar\mu_t|\mathcal{G}] \leqq I[\mu|\lambda] \quad \textit{where } \bar\mu_t := \frac{1}{t}\int_0^t \mu\mathcal{P}^s\,ds.$$

The argument above can be made rigorous by exploiting our conditions postulating a duplicated semigroup structure, see [FLO] for details. Therefore $I[\mu|\lambda] < +\infty$ implies $D[\mu|\mathcal{G}] = 0$ in a stationary regime, see (2.9) for an analytic way of solving this stationary Kolmogorov equation. To get a converse statement, set $\bar{\mathcal{R}}^t := e^{t\mathcal{G}}$; it is also a semigroup of self-adjoint contractions in $L^2(\lambda)$, its Markov property follows immediately by the Trotter product formula:

$$\bar{\mathcal{R}}^t = \lim_{n\to+\infty} (\mathcal{P}^{t/n}\mathcal{Q}^{t/n})^n. \tag{2.10}$$

The proof of Theorem 1 will be reduced to

LEMMA 1. *If* $D[\mu|\mathcal{G}] = 0$ *then, under conditions of Proposition 1,* $\mu(\phi\bar{\mathcal{R}}^t\psi) = \mu(\psi\bar{\mathcal{R}}^t\phi)$ *for all* $\phi, \psi \in L^2(\lambda)$ *and* $t > 0$.

PROOF. Observe first that $\partial_t \log(\bar{\mathcal{R}}^t\psi/\psi) = \mathcal{G}\bar{\mathcal{R}}^t\psi/\bar{\mathcal{R}}^t\psi$, whence by $D[\mu\,|\,\mathcal{G}] = 0$

$$0 \leqq \int \log\frac{\bar{\mathcal{R}}^t\psi}{\psi}d\mu \leqq \int \frac{\bar{\mathcal{R}}^t\psi - \psi}{\psi}d\mu$$
$$\leqq -1 + \int \frac{(\bar{\mathcal{R}}^t\psi^2)^{1/2}}{\psi}d\mu \leqq -1 + (\mu(\bar{\mathcal{R}}^t\psi^2/\psi^2))^{1/2}.$$

Choosing $\psi^2 = f := d\mu/d\lambda$ we get $1 \leqq \mu(\bar{\mathcal{R}}^t\sqrt{f}/\sqrt{f}) \leqq (\lambda(\bar{\mathcal{R}}^t f))^{1/2} = 1$, whence $f = \bar{\mathcal{R}}^t f$ λ-a.s., i.e. $\mu = \mu\bar{\mathcal{R}}^t$ for all $t > 0$, which implies also the equation of reversibility, see [F1]. □

It is not rare that $\mathcal{G}$ is heavily degenerated, even $\mathcal{G} \equiv 0$ is possible as it is for Hamiltonian dynamics, when λ is an equilibrium Gibbs state, $\mathcal{L}^* = -\mathcal{L}$, and entropy is a constant of motion. The opposite extreme situation is that of reversible diffusion processes. In that case $\mathcal{Q}^t = \mathcal{P}^t$, i.e. $\mathcal{G} = 2\mathcal{L}$, and the verification of the conditions of Proposition 1 amounts to establishing smooth dependence of solutions on initial values. Assuming the smoothness of the coefficients of the underlying stochastic equations, a standard argument shows that twice continuously differentiable functions with compact supports form a core of the generator. If the diffusion matrix is positive then $D[\mu|\mathcal{G}] = 0$ yields $\mu = \lambda$, thus $\mu_0\mathcal{P}^t \to \lambda$ as $t \to \infty$ for all $\mu_0 \ll \lambda$.

Our next task is to extend these results to infinite volumes, this is done by means of a familiar argument of Holley [Ho]; in translation invariant situations we can pass to the thermodynamic limit. This procedure cannot be carried out in a general framework, see e.g. [FFL] and [FLO]; technical requirements are summarized in the next section.

3. On locality of dynamics

Results of [Fr2] are not directly applicable in the present situation, that is why we review some parts of the argument. A convenient collection Θ of cutoff functions is defined for $m \in \mathbb{R}^d$ and $l \geqq 1$ by $\theta^l_m = \theta^l_m(x) := \theta_0(||x-m| - l|_+)$, where $\theta_0 \in C^3_0(\mathbb{R})$ satisfies $0 \leqq \theta_0(u) \leqq 1$ $\forall u \in \mathbb{R}$, while $\theta_0(u) = 1$ if $u \leqq 1$ and $\theta_0(u) = 0$ if $u \geqq 2$; $\overline{\Theta}$ is obtained by joining $\theta \equiv 1$ to Θ. In case of the full (infinite) dynamics the mark $\theta = 1$ is usually omitted, the limiting solution, the associated semigroup and its generator will be denoted as $\boldsymbol{\omega} = \omega(t,\sigma)$, $\mathcal{P}^t$ and $\mathcal{L}$, respectively. The basic a priori bound of [Fr2] can be reformulated as follows, see Proposition 2 and (3.18) there. Let $N_k(\omega)$ denote the number of points of ω in $B(\omega_k, 1)$ and

$$\overline{N}_\theta(t,\sigma) := 1 + \sup_{k\in S}\max_{s\leqq t} \frac{N_k(\omega^\theta(s))}{\sqrt{g_\alpha(\omega^\theta_k(s))}}. \tag{3.1}$$

Exploiting the superstability of the interaction, by means of the argument of Proposition 2 in [Fr2] we get

PROPOSITION 2. *If $\alpha \leqq 2 - d/2$ then for each $t > 0$ and $h > 0$*

$$\lim_{\rho\to\infty}\sup_{\theta\in\overline{\Theta}}\sup_{\sigma\in\overline{\Omega}_{\alpha,h}} P[\overline{N}_\theta(t,\sigma) > \rho] = 0.$$

First we derive a uniform bound on the localization of particles. From the stochastic equations

$$\begin{aligned} |\omega^\theta_k(t,\sigma) - \sigma_k| \leqq K_1 \overline{N}_\theta(t,\sigma) \int_0^t \sqrt{g_\alpha(\omega^\theta_k(s,\sigma))}\, ds \\ + \left|\int_0^t \sqrt{\theta(\omega^\theta_k(s,\sigma))}\, dw_k\right|. \end{aligned} \tag{3.2}$$

Let $g_*(u) := (1+|u|)^{4/5}$, by a direct calculation

$$\begin{gathered} \delta_{\theta,k}(t,\sigma) := \max_{s\leqq t} |\omega^\theta_k(s,\sigma) - \sigma_k| \leqq \xi_\theta(t,\sigma)(g_*(\sigma_k) + g_*(\delta_{\theta,k}(t,\sigma))), \\ \xi_\theta(t,\sigma) := K_2 \int_0^t \overline{N}_\theta(s,\sigma)\, ds + \sup_{k\in S}\max_{s\leqq t} \frac{K_2}{g_*(\sigma_k)} \left|\int_0^t \sqrt{\theta(\omega^\theta_k(s,\sigma))}\, dw_k\right| \end{gathered} \tag{3.3}$$

including $\theta \equiv 1$, whence by assuming $\delta_{\theta,k} \geqq g_*(\sigma_k)$ we get

$$\delta_{\theta,k}(t,\sigma) \leqq \eta_\theta(t,\sigma) g_*(\sigma_k), \tag{3.4}$$

where the explicit form $\eta = K_3\xi^5$ is not relevant, we only need

$$\lim_{y\to\infty}\sup_{\theta\in\overline{\Theta}}\sup_{\sigma\in\overline{\Omega}_{\alpha,h}} P[\eta_\theta(t,\sigma) > y] = 0 \tag{3.5}$$

for all $t,h>0$, which is a direct consequence of the definition of ξ.

Now we are in a position to estimate the rate of convergence of partial dynamics $\boldsymbol{\omega}^\theta$ to its limit $\boldsymbol{\omega}$ as $\theta\to 1$. For any initial configuration $\sigma\in\overline{\Omega}_\alpha$ let $S(m,r,\sigma)$ denote the set of $k\in S$ such that $|\sigma_k-m|\leqq r$, and consider

$$\Delta^l_m(t,r,\sigma):=\max_{k\in S(m,r,\sigma)}\max_{s\leqq t}|\omega^\theta_k(s,\sigma)-\omega_k(s,\sigma)|\quad\text{with }\theta=\theta^l_m. \tag{3.6}$$

For any fixed $T>0$ and $r_0,l\geqq 1$ define r_κ, $\kappa=0,1,\ldots,\chi,\ldots$ by

$$\begin{aligned} r_{\kappa+1}&=r_\kappa+2g_*(|m|+l)\,\tilde{\eta}^l_m(T,\sigma)+R+1,\quad\text{where}\\ \tilde{\eta}^l_m(T,\sigma)&:=\max\{\eta_\theta(T,\sigma),\eta_1(T,\sigma)\},\quad\theta=\theta^l_m. \end{aligned} \tag{3.7}$$

In view of (3.4) this means that before time T the particles starting from $B(m,r_\kappa)$ cannot interact with those starting from outside of $B(m,r_{\kappa+1})$, therefore

$$\begin{aligned} \Delta^l_m(t,r_\kappa,\sigma)&\leqq Lg(|m|+l)\,\tilde{N}^l_m(t,\sigma)\int_0^t\Delta^l_m(s,r_{\kappa+1},\sigma)\,ds,\quad\text{where}\\ \tilde{N}^l_m(t,\sigma)&:=\max\{\overline{N}_{\theta^l_m}(t,\sigma),\overline{N}_1(t,\sigma)\}, \end{aligned} \tag{3.8}$$

provided that $r_{\kappa+1}+R\leqq l$.

Suppose that (3.8) can χ times be iterated, then for $t<T$

$$\Delta^l_m(t,r_0,\sigma)\leqq 2(l+1)\frac{(Lt)^\chi}{\chi!}\left(g_*(|m|+l)\tilde{N}^l_m(t,\sigma)\right)^\chi, \tag{3.9}$$

where $\chi=O(l(|m|+l)^{-4/5})$ is a random number. Of course, this inequality implies the a.s. convergence of partial solutions; this was shown in [Fr2] when $m=0$ and $l\to+\infty$. Here we need a more delicate result: $\boldsymbol{\omega}^\theta$ with $\theta=\theta^l_m$ converges even if $|m|$ increases together with l. More precisely, for any $r_0,t,h,\varepsilon>0$ we have

$$\lim_{n\to\infty}\sup_{\sigma,m}\{P[\tilde{N}^l_m(t,\sigma)\Delta^l_m(t,r_0,\sigma)>\varepsilon]:\sigma\in\Omega_{\alpha,h},\ m,l\in M_n\}=0, \tag{3.10}$$

where $M_n:=\{m,l:|m|+l+R<n,\ l>n^{5/6}\}$. Indeed, in this situation χ of (3.9) goes a.s. to $+\infty$ as $n\to\infty$.

In the next section the following consequence of (3.10) will be needed. Suppose that we are given a translation invariant probability measure μ such that $\mu(\overline{\Omega}_\alpha)=1$, and set $\hat{\mu}_n:=\mu_{\Lambda_n}\lambda$. The above calculations are summarized in

LEMMA 2. *For any* $\phi \in C_0^1(\Omega)$ *and* $t > 0$ *we have*

$$\lim_{n\to\infty} \sup_{s,m}\{|\hat{\mu}_n \mathcal{P}^s \mathrm{s}^m \phi - \mu \mathcal{P}^s \phi| : s \le t,\ |m| < n - n^{5/6}\} = 0\,.$$

PROOF. Since $\hat{\mu}_n(s^m\varphi) = \mu(\varphi)$ if $\varphi, \mathrm{s}^m\varphi \in \mathcal{F}_{\Lambda_n}$, it is natural to approximate $\mathcal{P}^s\phi$ by $\varphi_l = \mathcal{P}_\theta^s\phi$ with $\theta = \theta_0^l$; remember that $\mathrm{s}^m\varphi_l = \mathcal{P}_{m,l}^s\phi$. Since ϕ is Lipschitz continuous by assumption, we can compare $\mathrm{s}^m\varphi_l$ and $\mathcal{P}^s\mathrm{s}^m\phi$ via (3.10), at least if $|m| < n - n^{5/6}$. The missing part of the argument, namely

$$\lim_{h\to\infty} \sup_{n\in\mathbb{N}} \hat{\mu}_n(\overline{\Omega}_{\alpha,h}) = 0 \tag{3.11}$$

follows from the basic superstability estimate of Ruelle [Ru2]. Indeed, for any box Λ of given shape and size we have $\lambda[\omega(\Lambda) > \nu|\mathcal{F}_{\Lambda^c}] \le Ce^{-c\nu^2}$, where c and C do not depend on ω. In view of (1.2) this yields $\lambda(\overline{\Omega}_\alpha) = 1$ by the Borel Cantelli lemma. Since $\mu(\overline{\Omega}_\alpha) = 1$ by assumption, estimating the contribution of particles from Λ_n^c to H_α via superstability, we get (3.11) by a similar computation. □

REMARK. Since the level sets of $\overline{H}$ are compact, the Stone Weierstrass theorem allows us to extend Lemma 2 to continuous and bounded local functions.

4. Passage to the thermodynamic limit

Now we are in a position to prove Theorem 1 by extending Proposition 1 to infinite volumes. Using the notation $\hat{\mu}_n^* = \mu_{\Lambda_n}^* \lambda$ of Lemma 2, we have

$$I[\hat{\mu}_n^*\mathcal{P}_\theta^t|\lambda] + 2tD[\bar{\mu}_{n,\theta,t}^*|\mathcal{L}_\theta] \le I[\hat{\mu}_n^*|\lambda] = I_{\Lambda_n}[\mu^*|\lambda] \tag{4.1}$$

for any smooth cutoff θ, where $\bar{\mu}_{n,\theta,t}^*$ is the time average of the evolved measures $\hat{\mu}_n^*\mathcal{P}_\theta^s$ from $s = 0$ through $s = t$. In view of (2.6), D is subadditive in the following sense. Suppose that $J_\theta^l(n) \subset \Lambda_n$ satisfies $\theta \ge \theta_m^l$ and $\theta_m^l\theta_k^l = 0$ for $m, k \in J_\theta^l(n)$, $k \ne m$ then

$$D[\bar{\mu}_{n,\theta,t}^*|\mathcal{L}_\theta] \ge \sum_{m\in J_\theta^l(n)} D[\bar{\mu}_{n,\theta,t}^*|\mathcal{L}_{\theta_m^l}] \ge \sum_{m\in J_\theta^l(n)} \int \frac{-\mathcal{L}_{\theta_m^l}\mathrm{s}^m\psi}{\mathrm{s}^m\psi}\, d\bar{\mu}_{n,\theta,t}^* \tag{4.2}$$

for smooth $\psi > 0$. Similarly, for all $\varphi \in C_0(\Omega)$

$$I[\hat{\mu}_n^*\mathcal{P}_\theta^t|\lambda] \ge \int S_n(\mathcal{P}_\theta^t(\varphi))\, d\hat{\mu}_n^* - F_\lambda(S_n(\varphi)), \quad \text{with} \qquad S_n(\varphi) := \sum_{m\in\Lambda_n\cap\mathbb{Z}^d} \mathrm{s}^m\varphi. \tag{4.3}$$

Now we can remove the cutoff of dynamics. Keeping $J_0^l(n) = J^l(n) \subset \Lambda_{n-n^{5/6}}$ fixed during this procedure we get

(4.4)

$$\sum_{m\in J^l(n)} \int_0^t ds \int \frac{-\mathcal{L}_{\theta_m^l} \mathrm{s}^m\psi}{\mathrm{s}^m\psi}\, d\hat{\mu}_n^* \mathcal{P}^s \leqq I[\hat{\mu}_n^*|\lambda] + F_\lambda(S_n(\varphi)) - \int S_n(\mathcal{P}^t\varphi)\, d\hat{\mu}_n^*.$$

As far as l is fixed, we may assume that Card $J^l(n) \geqq c_l|\Lambda_n|$ with some $c_l > 0$; thus dividing both sides by $|\Lambda_n|$ we can pass to a thermodynamic limit. Indeed, in view of Lemma 2 all terms of $\hat{\mu}_n^*\mathcal{P}^t(S_n(\varphi))$ become asymptotically identical when $n\to\infty$. Since $\mathcal{L}_{\theta_m^l} = \mathrm{s}^m\mathcal{L}_{\theta_0^l}$, the same holds true on the left-hand side, thus for all $\theta\in\Theta$ with compact support we have some $c_\theta > 0$ such that

$$tc_\theta \int \frac{-\mathcal{L}_\theta\psi}{\psi}\, d\mu^* \leqq I[\mu^*|\lambda] + F_\lambda(\varphi) - \mu^*(\varphi), \tag{4.6}$$

where $\varphi\in C_0(\Omega)$ is arbitrary, consequently $D[\mu^*|\mathcal{L}_\theta] = 0$ for all $\theta\in\Theta$. In this way we have managed to decouple (localize) the equations of stationarity. In fact, if $\Lambda := \operatorname{supp}\theta$ then $\mathcal{L}_\theta$ generates a reversible diffusion with a nonsingular diffusion matrix in each layer $\Omega_\Lambda(n|\sigma)$. Such diffusions have a unique stationary measure, which completes the proof of Theorem 1 by Lemma 1.

REFERENCES

[CRZ] Cattiaux, P., Roelly, S. and Zessin, H., Une approche gibbsienne des diffusions browniennes infini-dimensionnelles, *Probab. Theory Related Fields* **104** (1996), 147–179. *MR* **97b**:60128

[Cs] Csiszár, I., Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8** (1963), 85–108. *MR* **29** #1671

[D] Dobrushin, R.L., Description of a random field by means of its conditional probabilities and conditions of its regularity, *Theor. Probab. Appl.* **13** (1968), 197–224. Russian original in *Teor. Verojatnost. i Primenen.* **13** (1968), 201–229. *MR* **37** #6989

[DV] 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

[EK] Ethier, S. N. and Kurtz, T. G., *Markov processes. Characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley and Sons, Inc., New York, 1986. *MR* **88a**:60130

[Fo] Foguel, S. R., *The ergodic theory of Markov processes*, Van Nostrand Mathematical Studies, No. 21, Van Nostrand Reinhold Co., New York – Toronto – London, 1969. *MR* **41** #6299

[Fr1] Fritz, J., An information-theoretical proof of limit theorems for reversible Markov processes, *Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* (Tech. Univ., Prague, 1971; dedicated to the memory of Antonin Spacek), Academia, Prague, 1973, 183–197. *MR* **51** #6992

[FR2] FRITZ, J., Gradient dynamics of infinite point systems, *Ann. Probab.* **15** (1987), 478–514. *MR* **89a**:60239

[FFL] FRITZ, J., FUNAKI, T. and LEBOWITZ, J. L., Stationary states of random Hamiltonian systems, *Probab. Theory Related Fields* **99** (1994), 211–236. *MR* **95g**:82003

[FLO] FRITZ, J., LIVERANI, C. and OLLA, S., Reversibility in infinite Hamiltonian systems with conservative noise, *Comm. Math. Phys.* **189** (1997), 481–496.

[GEO] GEORGII, H. O., *Canonical Gibbs measures. Some extensions of de Finetti's representation theorem for interacting particle systems*, Lecture Notes in Mathematics, **760**, Springer, Berlin, 1979. *MR* **81b**:60108

[HO] HOLLEY, R., Free energy in a Markovian model of a lattice spin system, *Comm. Math. Phys.* **23** (1971), 87–99. *MR* **45** #1535

[LA1] LANG, R., Unendlich-dimensionale Wienerprozesse mit Wechselwirkung. I. Existenz, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** (1977), 55–72. *MR* **55** #4434

[LA2] LANG, R., Unendlich-dimensionale Wienerprozesse mit Wechselwirkung. II. Die reversiblen Masse sind kanonische Gibbs–Masse, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **39** (1977), 277–299. *MR* **56** #13401

[OVY] OLLA, S., VARADHAN, S. R. S. and YAU, H. T., Hydrodynamical limit for a Hamiltonian system with weak noise, *Comm. Math. Phys.* **155** (1993), 523–560. *MR* **94k**:60158

[OS] OSADA, H., Dirichlet form approach to infinite-dimensional Wiener processes with singular interactions, *Comm. Math. Phys.* **176** (1996), 117–131. *MR* **96m**:60178

[RE1] RÉNYI, A., On measures of entropy and information, *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.*, Vol. I, Univ. California Press, Berkeley, Calif., 1961, 547–561. *MR* **24** #A2410

[RE2] RÉNYI, A., *Probability theory*, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10, North-Holland Publishing Co., Amsterdam–London, 1970. *MR* **47** #4296

[RU1] RUELLE, D., *Statistical mechanics: Rigorous results*, W. A. Benjamin, Inc., New York – Amsterdam, 1969. *MR* **44** #6279

[RU2] RUELLE, D., Superstable interactions in classical statistical mechanics, *Comm. Math. Phys.* **18** (1970), 127–159. *MR* **42** #1468

(Received January 20, 1998)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS STATISZTIKAI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY
jofri@math.elte.hu

LABORATOIRE DE STATISTIQUE ET PROBABILITÉS
URA CNRS D–751
DÉPARTEMENT D'ANALYSE ET GEOMÉTRIE
UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE
CITÉ SCIENTIFIQUE
F–59655 VILLENEUVE D'ASQ Cedex
FRANCE
sylvie.roelly@univ-lille1.fr

LABORATOIRE DE STATISTIQUE ET PROBABILITÉS
UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE
CITÉ SCIENTIFIQUE
F–59655 VILLENEUVE D'ASQ Cedex
FRANCE
hans.zessin@univ-lille1.fr

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 165–181

ON THE MAXIMAL GAIN OVER HEAD RUNS

A. FROLOV, A. MARTIKAINEN and J. STEINEBACH

To the memory of A. Rényi

Abstract

Let (X_i, Y_i) be a sequence of i.i.d. random vectors where $\{X_i\}$ are gains and $\{Y_i\}$ are indicators of successes in repetitions of a game of heads and tails. Put $S_0 = 0$, $S_k = X_1 + \cdots + X_k$, and let $I\{\cdot\}$ denote the indicator function of the event in brackets. Then $M_N = \max_{0 \leq l < m \leq N}(S_m - S_l)I\{Y_{l+1} = \cdots = Y_m = 1\}$ is the maximal gain over sequences of successes without interruptions ("head runs"). We derive necessary and sufficient conditions for strong laws of large numbers for M_N and find rates of convergence in these laws.

1. Introduction and results

Consider a sequence $\{(X_i, Y_i)\}_{i=1,2,\ldots}$ of independent, identically distributed (i.i.d.) random vectors, where

$$\mathbf{P}(Y_1 = 1) = p = 1 - \mathbf{P}(Y_1 = 0) \in (0, 1). \tag{1.1}$$

Let $S_k = \sum_{i=1}^{k} X_i$, $S_0 = 0$, and

$$M_N = \max_{0 \leq l < m \leq N} (S_m - S_l) I\{Y_{l+1} = \cdots = Y_m = 1\},$$

where $I\{\cdot\}$ denotes the indicator function of the event in brackets.

If $p = 1$, the random variable M_N has been studied in various contexts (cf. e.g. Dembo and Karlin [7], Karlin and Dembo [10], and the work mentioned therein). Typically, if the random walk $\{S_k\}$ has a negative drift, an almost sure limiting behaviour of $\{M_N\}$ requires a logarithmic normalization.

This phenomenon (of logarithmic normalization) has earlier been observed by Erdős–Rényi [8] and Shepp [12] for maxima of increments over subintervals of logarithmic length in a large interval. Csörgő–Steinebach [3] studied maxima over increments of *at most* logarithmic length and obtained a first convergence rate result. For the precise rates of convergence, and

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F10.

Key words and phrases. Head run, increment of random walk, strong law of large numbers, large deviations, convergence rate.

0081–6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

for various extensions and improvements of the Erdős–Rényi–Shepp laws we refer to Deheuvels–Devroye–Lynch [5] and Deheuvels–Devroye [6].

In this paper, we focus our attention on the case $p<1$, i.e. we are interested in the maximal exceedances of a random walk $\{S_k\}$ over "runs" of a companion sequence $\{Y_i\}$. Since head runs are of logarithmic order (cf. Corollary 1 below), one can expect that, if the drift is positive, the maximal gain over head runs is of logarithmic order, too. This, however, may fail when the drift is negative. In the latter case, the head run corresponding to the maximal gain has *at most* logarithmic length. It turns out that still an Erdős–Rényi–Shepp type phenomenon holds for the maximal gain over head runs whatever drift the random walk may have.

Note that the limit in our main result does not depend only on $\mathbf{E}X_1$ and/or the marginal distribution of Y_1. It is a function of the full distribution of X_1 given $Y_1=1$, and will be introduced next.

Let

$$\varphi(h)=\mathbf{E}\{e^{hX_1}|Y_1=1\},$$

and assume that

$$h_0=\sup\{h:\varphi(h)<\infty\}>0, \tag{1.2}$$

$$\mathbf{E}\{|X_1||Y_1=1\}<\infty. \tag{1.3}$$

Define

$$h^*=\sup\left\{h\geqq 0:\varphi(h)\leqq\frac{1}{p}\right\}. \tag{1.4}$$

Note that $0<h^*\leqq\infty$, and, if $h^*<h_0$, then h^* is the unique positive solution of the equation

$$\varphi(h)=\frac{1}{p}.$$

If $h^*=h_0$, then $\varphi(h^*)\leqq 1/p$ by the monotone convergence theorem, but the inequality may be strict.

Theorem 1. *Let $\{(X_i,Y_i)\}_{i=1,2,\dots}$ be a sequence of i.i.d. random vectors satisfying* (1.1) *to* (1.3). *Then*

$$\lim_{N\to\infty}\frac{M_N}{\log N}=\frac{1}{h^*}\qquad a.s.. \tag{1.5}$$

When $p=1$, $\mathbf{E}X_1<0$, and the X_i's are bounded, a corresponding result has been proved by Karlin and Dembo [10].

On choosing $X_i=Y_i$, we obtain the Erdős–Rényi strong law for the length of the longest head run:

COROLLARY 1. *Let Z_N be the length of the longest head run in N independent tosses of a coin with* $\mathbf{P}(\text{"head"}) = p \in (0,1)$. *Then*

$$\lim_{N\to\infty} \frac{Z_N}{\log N} = \frac{1}{\log(1/p)} \qquad a.s..$$

Refined and extended versions of the latter statement have been proved by Erdős–Révész [9] and Deheuvels [4].

REMARK 1. Condition (1.2) is also necessary for (1.5). Indeed, if $\limsup M_N/\log N < \infty$ a.s., then, for some $h > 0$,

$$\limsup_{N\to\infty} \frac{X_N I\{Y_N = 1\}}{\log N} < \frac{1}{h} \qquad \text{a.s.}.$$

So, by the Borel–Cantelli lemma,

$$\sum_{N=1}^{\infty} \mathbf{P}(X_1 I\{Y_1 = 1\} \geqq \frac{1}{h}\log N) < \infty.$$

Hence

$$\sum_{N=1}^{\infty} \mathbf{P}(hX_1 \geqq \log N | Y_1 = 1) < \infty,$$

which implies (1.2).

We also obtain a convergence rate result for (1.5) provided $h^* < h_0$. It is an analogue of the Deheuvels–Devroye–Lynch [5] improvements of the Erdős–Rényi–Shepp strong laws of large numbers.

We restrict our attention to the case

$$\mathbf{P}(X_1 = x | Y_1 = 1) < 1 \quad \forall x. \tag{1.6}$$

THEOREM 2. *Let $\{(X_i, Y_i)\}_{i=1,2,\ldots}$ be a sequence of i.i.d. random vectors satisfying* (1.1) *to* (1.3) *and* (1.6). *Assume that $h^* < h_0$. Then*

$$\frac{h^* M_N - \log N}{\log\log N} \to -\frac{1}{2} \qquad \text{in probability},$$

$$\limsup_{N\to\infty} \frac{h^* M_N - \log N}{\log\log N} = \frac{1}{2} \qquad a.s.,$$

$$\liminf_{N\to\infty} \frac{h^* M_N - \log N}{\log\log N} = -\frac{1}{2} \qquad a.s..$$

The case $h^* = h_0$ is excluded from Theorem 2 because it involves large deviation probabilities in a wider zone than before. A different asymptotic of these probabilities may result in a different convergence rate in (1.5).

If (1.6) fails, the behaviour of M_N can also be different. Confer, e.g. Theorem 2 in Deheuvels [4].

2. Large deviation results

Let $Z_1, Z_2, \ldots$ be a sequence of i.i.d. random variables with $\mathbf{E}|Z_1| < \infty$, $\mathbf{P}(Z_1 = x) < 1$ for all x,

$$t_0 = \sup\{t \geqq 0 : \phi(t) = \mathbf{E}e^{tZ_1} < \infty\} > 0.$$

Define $T_n = \sum_{i=1}^{n} Z_i$, $T_0 = 0$, and

$$\mu(t) = \frac{\phi'(t)}{\phi(t)} \quad , \quad \rho(\alpha) = \inf_t \phi(t) e^{-t\alpha}.$$

Put $A = \lim_{t \uparrow t_0} \mu(t)$. Then, for all $\alpha \in (\mathbf{E}Z_1, A)$, there exists a unique $t^* \in (0, t_0)$ such that $\mu(t^*) = \alpha$. Moreover we have

$$\rho(\alpha) = \phi(t^*) e^{-\alpha t^*}.$$

THEOREM 3 (Petrov [11]). *Under the assumptions above,*

$$\mathbf{P}(T_n \geqq n\alpha) \sim \frac{\psi(t^*)}{\sqrt{n}} \rho^n(\alpha)$$

uniformly for $\alpha \in [\mathbf{E}Z_1 + \varepsilon, \min\{A - \varepsilon, 1/\varepsilon\}]$, *where* $\varepsilon > 0$ *is arbitrary, and* $\psi(t^*) > 0$ *is a finite constant depending upon* t^* *and the distribution of* Z_1 *only.*

For nonlattice distributions, $\psi(t^*) = 1/(t^* \sigma(t^*) \sqrt{2\pi})$, *while for lattice distributions with span* H, $\psi(t^*) = H/(\{1 - e^{-Ht^*}\} \sigma(t^*) \sqrt{2\pi})$, *where* $\sigma(t) = \mu'(t)$.

We will use the following corollary of Petrov's theorem.

LEMMA 1. *Let* $\alpha \in (\mathbf{E}Z_1, A)$ *and let* y_n *be a sequence of numbers satisfying* $ny_n^3 \to 0$ *as* $n \to \infty$. *Then, uniformly over* z *with* $|z| \leqq |y_n|$, *we have*

$$\mathbf{P}(T_n \geqq n(\alpha + z)) \sim \frac{\psi(t^*)}{\sqrt{n}} \rho^n(\alpha) \exp\left\{-nzt^* - \frac{nz^2}{2\sigma(t^*)}\right\}.$$

PROOF. We have

$$(\log \rho(\alpha))' = (\log \phi(t^*) - \alpha t^*)' = \mu(t^*)(t^*(\alpha))' - t^*(\alpha) - \alpha(t^*(\alpha))' = -t^*(\alpha).$$

It follows from $\mu(t^*(\alpha)) = \alpha$ that $\sigma(t^*(\alpha))(t^*(\alpha))' = 1$. Hence,

$$(\log \rho(\alpha))'' = -(t^*(\alpha))' = -\frac{1}{\sigma(t^*(\alpha))}.$$

Since $t^*(\alpha)$ is a continuous function of α, $\psi(t^*(\alpha))$ is continuous. Applying Taylor's expansion we get the conclusion of the lemma.

3. Proofs

First we prove some auxiliary results under the conditions (1.1) to (1.3) and (1.6).

Put

$$m(h) = \frac{\varphi'(h)}{\varphi(h)},$$

$$I_{l,m} = I\{Y_{l+1} = \cdots = Y_m = 1\},$$

and

$$m^* = m(h^*), \quad m' = m'(h^*), \quad C = \frac{1}{h^* m^*}. \tag{3.1}$$

Let $\bar{X}_j$ be a sequence of independent random variables with cumulative distribution function $F(x) = \mathbf{P}(X_1 < x | Y_1 = 1)$, and $\bar{S}_j = \sum_{i=1}^{j} \bar{X}_i$, $\bar{S}_0 = 0$. Note that

$$\mathbf{P}(S_j > x | I_{0,j} = 1) = \mathbf{P}(\bar{S}_j > x). \tag{3.2}$$

The proofs below make use of similar techniques as developed in Deheuvels–Devroye–Lynch [5] and Deheuvels–Devroye [6].

LEMMA 2.

$$\limsup_{N\to\infty} \frac{h^* M_N - \log N}{\log\log N} \leqq \frac{1}{2} \quad a.s..$$

PROOF. Put

$$x = x_N = \frac{1}{h^*} \log N + (1+\varepsilon)\frac{1}{2h^*} \log\log N,$$

$$K_1 = [C \log N - A\sqrt{\log N \log\log N}], \quad K_2 = [C \log N + A\sqrt{\log N \log\log N}],$$

$$K_3 = [C_1 \log N],$$

where $\varepsilon > 0$, and $[x]$ denotes the integer part of x. The positive constants A and C_1 will be specified below. Note that $x > 0$ for $N > e^e$.

We have

$$\mathbf{P}(M_N > x) \leqq P(N) + Q(N), \tag{3.3}$$

where

$$P(N) = \mathbf{P}(M_N > x, I_{l,m} = 0 \text{ for all } l, m \text{ such that } m \geqq l + K_3), \tag{3.4}$$

$$Q(N) = \mathbf{P}(I_{l,m} = 1 \text{ for some } l, m \text{ such that } m \geqq l + K_3). \tag{3.5}$$

Then, for any $\delta > 0$.

$$Q(N) \leqq N^2 p^{K_3} \leqq \frac{1}{p} N^{2+C_1 \log p} = \frac{1}{p} N^{-\delta}, \tag{3.6}$$

provided $C_1 = (2+\delta)/\log(1/p)$. From the definition of h^* and m^* in (1.4) and (3.1), it is easy to check that $C \leqq 1/\log(1/p)$, so $C < C_1$.

We further get

$$\begin{aligned} P(N) &\leqq N\mathbf{P}(\max_{1\leqq j\leqq K_3} S_j I_{0,j} > x) \leqq N \sum_{j=1}^{K_3} \mathbf{P}(S_j I_{0,j} > x) \\ &= N \sum_{j=1}^{K_3} p^j \mathbf{P}(S_j > x | I_{0,j} = 1). \end{aligned} \tag{3.7}$$

Hence, by (3.2),

$$P(N) \leqq N \sum_{j=1}^{K_3} p^j \mathbf{P}(S_j > x) = N(P_1(N) + P_2(N) + P_3(N)), \tag{3.8}$$

where

$$P_1(N) = \sum_{j=K_2+1}^{K_3} (\cdot), \quad P_2(N) = \sum_{j=K_1+1}^{K_2} (\cdot), \quad P_3(N) = \sum_{j=1}^{K_1} (\cdot).$$

Put

$$\varepsilon_N = A\sqrt{\frac{\log\log N}{\log N}},$$

and let t_N be the unique solution of the equation

$$m(t_N) = \frac{1}{h^*(C+\varepsilon_N)}.$$

Note that the function $m(t)$ is strictly increasing and therefore $t_N \uparrow h^*$. By Markov's inequality

$$P_1(N) \leqq \sum_{j=K_2+1}^{K_3} (p\varphi(t_N))^j e^{-xt_N} \leqq C_1 \log N (p\varphi(t_N))^{K_2+1} e^{-xt_N}, \tag{3.9}$$

because, by the strict convexity of φ together with $\varphi(0) = 1$, $\varphi(h^*) = 1/p > 1$, one has $\varphi(t_N) < \varphi(h^*)$. By the definitions of K_2, x and ε_N,

$$\log N (p\varphi(t_N))^{K_2+1} e^{-xt_N} \leqq (\log N)^{1-(t_N/(2h^*))(1+\varepsilon)} N^y, \tag{3.10}$$

where

$$y = -\frac{t_N}{h^*} + (C + \varepsilon_N)\log(p\varphi(t_N)).$$

Since

$$\log(p\varphi(t_N)) = m^*(t_N - h^*) + \frac{m'}{2}(t_N - h^*)^2 + o((t_N - h^*)^2),$$

we get, by the definitions of C, t_N and ε_N,

$$\begin{aligned} y &= -1 + (C+\varepsilon_N)m^*(t_N - h^*) - \frac{1}{h^*}(t_N - h^*) + C\frac{m'}{2}(t_N - h^*)^2 + o((t_N - h^*)^2) \\ &= -1 + \frac{1}{h^*}\Big(\frac{m^*}{m(t_N)} - 1\Big)(t_N - h^*) + C\frac{m'}{2}(t_N - h^*)^2 + o((t_N - h^*)^2) \\ &= -1 + \frac{1}{h^*}(t_N - h^*)g(t_N) + o((t_N - h^*)^2), \end{aligned}$$

where

$$g(t) = \frac{m^*}{m(t)} - 1 + C\frac{h^* m'}{2}(t - h^*).$$

Note that

$$g'(h^*) = -\frac{m'}{2m^*} < 0.$$

Since $g(h^*) = 0$, we have

$$g(t_N) = g'(h^*)(t_N - h^*) + o((t_N - h^*)).$$

Thus,

$$y = -1 - C\frac{m'}{2}(t_N - h^*)^2 + o((t_N - h^*)^2).$$

Moreover, by the definition of t_N,

$$\varepsilon_N = C\Big(\frac{m^*}{m(t_N)} - 1\Big) \sim -C\frac{m'}{m^*}(t_N - h^*).$$

Here we have used the same argument as for $g(t)$. It follows that

$$C\frac{m'}{2}(t_N - h^*)^2 \sim \frac{3}{2}\frac{\log\log N}{\log N},$$

provided $A = (3m'/h^*(m^*)^3)^{1/2}$. Then, for any $\delta > 0$, the inequality

$$y < -1 - \Big(\frac{3}{2} - \delta\Big)\frac{\log\log N}{\log N}$$

holds for all large N. By (3.9), (3.10) and the last inequality,

$$P_1(N) \leqq C_1 N^{-1}(\log N)^{-(1+\tau)} \tag{3.11}$$

for all large N, if $\tau > 0$ is chosen small enough.

Let t'_N be the solution of the equation

$$m(t'_N) = \frac{1}{h^*(C-\varepsilon_N)}.$$

Then $t'_N \downarrow h^*$. As in (3.9) and (3.10), we have

$$P_3(N) \leqq C \log N (p\varphi(t'_N))^{K_1} e^{-xt'_N} = C(\log N)^{1-(t'_N/(2h^*))(1+\varepsilon)} N^z, \tag{3.12}$$

where

$$z = -\frac{t'_N}{h^*} + (C-\varepsilon_N)\log(p\varphi(t'_N)).$$

In the same way as before,

$$\begin{aligned} z &= -1 + \frac{1}{h^*}(t'_N - h^*)g(t'_N) + o((t'_N - h^*)^2) \\ &= -1 - C\frac{m'}{2}(t'_N - h^*)^2 + o((t'_N - h^*)^2) < -1 - \left(\frac{3}{2} - \delta\right)\frac{\log\log N}{\log N} \end{aligned}$$

for any $\delta > 0$ when N is large enough. Hence

$$P_3(N) \leqq C N^{-1}(\log N)^{-(1+\tau)} \tag{3.13}$$

for all large N, if $\tau > 0$ is chosen small enough.

Finally, put $z_j = z_{N,j} = (x - m^* j)/j$. By the definitions of x and m^*, we conclude that $z_j = O(j^{-1/2}(\log j)^{1/2})$ if $K_1 \leqq j \leqq K_2$. By Lemma 1, uniformly in $K_1 \leqq j \leqq K_2$,

$$\begin{aligned} \mathbf{P}(S_j > x) &= \mathbf{P}(S_j > j(m^* + z_j)) \\ &\sim \frac{\psi(h^*)}{\sqrt{j}}(\varphi(h^*)e^{-h^* m^*})^j \exp\left\{-(x - m^* j)h^* - \frac{(x - m^* j)^2}{2jm'}\right\} \\ &= \frac{\psi(h^*)}{\sqrt{j}} p^{-j} e^{-xh^*} \exp\left\{-\frac{(x - m^* j)^2}{2jm'}\right\}. \end{aligned}$$

It follows that

$$\begin{aligned} P_2(N) &\leqq \frac{\psi(h^*)}{\sqrt{K_1}} e^{-xh^* + x(m^*/m')} \sum_{j=K_1+1}^{K_2} e^{-(x^2/(2jm')) + (j(m^*)^2/(2m'))} \\ &\leqq \frac{\psi(h^*)}{\sqrt{K_1}} e^{-xh^* + x(m^*/m') - x^2/(2K_1 m')} \times \end{aligned}$$

$$\times (e^{-(K_1+1)(m^*)^2/(2m')} - e^{-(K_2+1)(m^*)^2/(2m')})(1 - e^{-(m^*)^2/(2m')})^{-1}$$
$$\leqq B\frac{K_2 - K_1}{\sqrt{K_1}} e^{-xh^* - (x - m^* K_1)^2/(2K_1 m')} \leqq B_1 N^{-1}(\log N)^{-(1+\tau)},$$

where the constants B and B_1 depend on h^* and the distribution of X_1 only. Hence

$$P_2(N) \leqq B_1 N^{-1}(\log N)^{-(1+\tau)} \tag{3.14}$$

for all large N. Via (3.11), (3.13) and (3.14), the inequality (3.8) gives

$$P(N) \leqq B_2(\log N)^{-(1+\tau)} \tag{3.15}$$

for all large N, $\tau > 0$ small enough.

Set $N_k = \min\{N : \log N \geqq k\}$. Then, for $N_k \leqq N < N_{k+1}$, one has $k \leqq \log N < k + 1$, and $M_{N_k} \leqq M_N \leqq M_{N_{k+1}}$. It follows from (3.3), (3.6) and (3.15) that the series

$$\sum_k \mathbf{P}(M_{N_k} > x_{N_k})$$

converges. Since M_N, $\log N$ and $\log\log N$ are non-decreasing and $\log N_{k+1} - \log N_k \leqq 2$, $\log\log N_k \sim \log k \sim \log\log N_{k+1}$ as $k \to \infty$, Lemma 2 follows via the Borel–Cantelli lemma.

LEMMA 3. *Put* $b = m^*k + (-u - \varepsilon)(\log k/h^*)$. *If* $\varepsilon > 0$, $\theta > 0$, $s > 0$, $u + s + \varepsilon \geqq 0$, *then*

$$\mathbf{P}(\bar{S}_k \geqq b, \bar{S}_{v+k} - \bar{S}_v \geqq b) \leqq p^{-k}e^{-k/C}v^{-s} + \mathbf{P}(\bar{S}_k \geqq b)k^{u+s+\varepsilon}v^{-\theta}$$

for $k \geqq v \geqq v_1 = v_1(\theta)$.

PROOF. Set $S' = \bar{S}_v, S'' = \bar{S}_k - \bar{S}_v, S''' = \bar{S}_{v+k} - \bar{S}_k$. The random variables $S' + S''$ and S''' are independent, and therefore

$$\begin{aligned}\mathbf{P}(\bar{S}_k \geqq b, \bar{S}_{v+k} - \bar{S}_v \geqq b) &\leqq \mathbf{P}(S'' \geqq q) + \mathbf{P}(S' + S'' \geqq b, S'' + S''' \geqq b, S'' < q)\\ &\leqq \mathbf{P}(S'' \geqq q) + \mathbf{P}(S' + S'' \geqq b)\mathbf{P}(S''' \geqq b - q)\\ &\leqq (\varphi(t))^{k-v}e^{-tq} + \mathbf{P}(\bar{S}_k \geqq b)(\varphi(t_1))^v e^{-t_1(b-q)}\end{aligned}$$

for any q and any positive t, t_1.

Now choose

$$t = h^*, \quad t_1 < t, \quad q = m^*k - \frac{v}{t}\log\varphi(t) + \frac{s}{t}\log v.$$

We have

$$\mathbf{P}(\bar{S}_k \geqq b, \bar{S}_{v+k} - \bar{S}_v \geqq b)$$

$$\leqq p^{-k}e^{-k/C}v^{-s}+\mathbf{P}(\bar{S}_k\geqq b)k^{t_1(u+\varepsilon)/t}v^{st_1/t}\exp\Big\{-t_1v\Big(\frac{1}{t}\log\varphi(t)-\frac{1}{t_1}\log\varphi(t_1)\Big)\Big\}$$

$$\leqq p^{-k}e^{-k/C}v^{-s}+\mathbf{P}(\bar{S}_k\geqq b)k^{u+\varepsilon+s}\exp\Big\{-t_1v\Big(\frac{1}{t}\log\varphi(t)-\frac{1}{t_1}\log\varphi(t_1)\Big)\Big\}.$$

Here we used the definitions of h^*, C in (1.4), (3.1), and the inequalities $t_1<t$, $v\leqq k$. By the strict convexity of $\log\varphi$ together with $\log\varphi(0)=0$,

$$\frac{1}{t}\log\varphi(t)-\frac{1}{t_1}\log\varphi(t_1)>0.$$

This gives Lemma 3.

LEMMA 4.
$$\limsup_{N\to\infty}\frac{h^*M_N-\log N}{\log\log N}\geqq\frac{1}{2}\quad a.s..$$

PROOF. For $j=1,2,\dots$, put

$$b_j=\frac{j}{h^*}+\frac{1}{h^*}\Big(\frac{1}{2}-\varepsilon\Big)\log j\quad,\quad K=[Cj],$$

with C specified in (3.1), and define

$$\bar{M}_j=\max_{[e^{j-1}]<l\leqq[e^j]-K}(S_{l+K}-S_l)I_{l,l+K}.$$

Note that the $\bar{M}_j$ are independent.

We prove that $\mathbf{P}(\bar{M}_j>b_j \text{ i.o.})=1$. To do so, it is enough to show that the series

$$\sum_j P_j=\sum_j\mathbf{P}(\bar{M}_j>b_j)$$

diverges, and to apply the Borel–Cantelli lemma.

Let

$$A_l=\{(S_{l+K}-S_l)I_{l,l+K}\geqq b_j\},\quad l=0,1,\dots,N,$$

where $N=N_j=[e^j]-[e^{j-1}]-K$. By the Chung–Erdős lemma (cf. Chung–Erdős [2]),

$$P_j=P\Big(\bigcup_{l=0}^{N}A_l\Big)\geqq\frac{(NP(A_0))^2}{NP(A_0)+(NP(A_0))^2+2N\sum_{k=1}^{K}P(A_0A_k)}.$$

By (3.2) and Lemma 1,

$$P(A_0)=\mathbf{P}(S_KI_{0,K}\geqq b_j)=p^K\mathbf{P}(\bar{S}_K\geqq b_j)$$
$$\sim p^K\frac{\psi(h^*)}{\sqrt{K}}\exp\Big\{-K(h^*m^*-\log\varphi(h^*))-\Big(\frac{1}{2}-\varepsilon\Big)\log j\Big\}.$$

It follows that

$$D_1 j^{-1+\varepsilon} \leqq NP(A_0) \leqq D_2 j^{-1+\varepsilon}. \tag{3.16}$$

Here and in the sequel, $D_1, D_2, \ldots$ are positive constants not depending on N.

Let

$$\tilde{A}_l = \{\bar{S}_{l+K} - \bar{S}_l \geqq b_j\}, \quad l = 0, 1, \ldots, N.$$

It is not difficult to check that

$$P(A_0 A_k) = p^{k+K} P(\tilde{A}_0 \tilde{A}_k) \quad (k = 0, 1, \ldots, K). \tag{3.17}$$

An application of Lemma 3 with $k = K, v = k, u = -1/2, s = 1 + 2/\varepsilon$, gives

$$P(\tilde{A}_0 \tilde{A}_k) \leqq p^{-K} e^{-K/C} k^{-(1+2/\varepsilon)} + P(\tilde{A}_0) K^{1/2+2/\varepsilon+\varepsilon} k^{-\theta},$$

for large k, where θ is an arbitrary positive constant.

Put $l = [K^{\varepsilon/2}]$. We have

$$N \sum_{k=1}^{K} p^K P(\tilde{A}_0 \tilde{A}_k) = N \sum_{k=1}^{l-1} p^K P(\tilde{A}_0 \tilde{A}_k) + N \sum_{k=l}^{K} p^K P(\tilde{A}_0 \tilde{A}_k)$$

$$\leqq N l p^K P(\tilde{A}_0) + N e^{-K/C} \sum_{k=l}^{K} k^{-(1+2/\varepsilon)} + N p^K P(\tilde{A}_0) K^{3/2+2/\varepsilon+\varepsilon} l^{-\theta}$$

$$\leqq D_3 K^{-1+3\varepsilon/2} + N e^{-K/C} l^{-2/\varepsilon} \sum_{k=l}^{K} k^{-1} + D_4 K^{-1+\varepsilon} K^{3/2+2/\varepsilon+\varepsilon} K^{-\theta\varepsilon/2}$$

$$\leqq D_3 K^{-1+3\varepsilon/2} + D_5 K^{-1} \log K + D_4 K^{-1+\varepsilon} K^{3/2+2/\varepsilon+\varepsilon} K^{-\theta\varepsilon/2} \leqq D_6 K^{-1+3\varepsilon/2},$$

since θ can be chosen arbitrarily large. Here we used the definitions of l, N, (3.16), (3.17) and Lemma 3. It follows that

$$N \sum_{k=1}^{K} P(A_0 A_k) \leqq N \sum_{k=1}^{K} p^K P(\tilde{A}_0 \tilde{A}_k) \leqq D_7 j^{-1+3\varepsilon/2} \tag{3.18}$$

for large j. By the last inequality and (3.16), we get

$$P_j \geqq D_8 \frac{j^{-2+2\varepsilon}}{j^{-1+3\varepsilon/2}} = D_8 j^{-1+\varepsilon/2},$$

which implies Lemma 4. Note that $\log N_j = j + o(1)$ as $j \to \infty$.

LEMMA 5. *For any* $\varepsilon > 0$,

$$\mathbf{P}\Big(\frac{h^* M_N - \log N}{\log\log N} \leqq -\frac{1}{2} + \varepsilon\Big) \to 1.$$

PROOF. This follows from the proof of Lemma 2 via formally replacing $1/(2h^*)$ by $-1/(2h^*)$.

LEMMA 6.

$$\liminf_{N\to\infty} \frac{h^* M_N - \log N}{\log\log N} \leqq -\frac{1}{2} \quad a.s..$$

PROOF. Lemma 6 is an immediate consequence of Lemma 5.

LEMMA 7. *For any* $\varepsilon > 0$,

$$\mathbf{P}\Big(\frac{h^* M_N - \log N}{\log\log N} \geqq -\frac{1}{2} - \varepsilon\Big) \to 1.$$

PROOF. Put $k = [C \log N]$ with C as defined in (3.1), and set

$$\tilde{M}_N = \max_{0 \leqq l \leqq N-k} (S_{l+k} - S_l) I_{l,l+k},$$

$$b_k = m^* k - \frac{1}{h^*}\Big(\frac{1}{2} + \varepsilon\Big) \log k.$$

Evidently, $M_N \geqq \tilde{M}_N$ and

$$\{\tilde{M}_N \geqq b_k\} = \bigcup_{l=0}^{N-k} A_l', \quad \text{where } A_l' = \{(S_{l+k} - S_l) I_{l,l+k} \geqq b_k\}.$$

$\mathbf{P}(\tilde{M}_N \geqq b_k)$ can be estimated as in Lemma 4. Note that the current definition of b_k differs from the one in Lemma 4. We obtain

$$D_1 j^{\varepsilon} \leqq NP(A_0') \leqq D_2 j^{\varepsilon}$$

instead of (3.16), and

$$N \sum_{k=1}^{K} P(A_0' A_k') \leqq D_8 j^{3\varepsilon/2}$$

instead of (3.18). This also implies

$$N \sum_{k=1}^{K} P(A_0' A_k') = o((NP(A_0'))^2).$$

An application of the Chung–Erdős lemma yields

$$\mathbf{P}(M_N \geqq b_k) \geqq \frac{(NP(A_0'))^2}{NP(A_0') + (NP(A_0'))^2 + 2N\sum_{k=1}^{K} P(A_0'A_k')} \to 1.$$

LEMMA 8.

$$\liminf_{N\to\infty} \frac{h^* M_N - \log N}{\log\log N} \geqq -\frac{1}{2} \quad a.s..$$

PROOF. Put $n_j = \inf\{n : k = j\}$ with k as defined in Lemma 7. Via the Borel–Cantelli lemma it is enough to prove that, for any $\varepsilon > 0$,

$$\sum_j \mathbf{P}\Big(M_{n_j} < m^* j - \frac{1}{h^*}\Big(\frac{1}{2}+\varepsilon\Big)\log j\Big) < \infty. \tag{3.19}$$

Set $J_l = \{i : 2lj \leqq i < (2l+1)j, i = r[j^{\varepsilon/2}], r = 0, 1, \dots\}$, and define independent random variables $Q_0, Q_1, \dots$ as

$$Q_l = \sup_{i\in J_l}(S_{i+j} - S_i)I_{i,j}.$$

Then

$$M_{n_j} \geqq \sup_{0\leqq l\leqq L} Q_l,$$

where L is the largest integer such that $(2L+1)j - 1 \leqq n_j - j$. Now,

$$P = \mathbf{P}\Big(M_{n_j} < m^* j - \frac{1}{h^*}\Big(\frac{1}{2}+\varepsilon\Big)\log j\Big) \leqq \prod_{l=0}^{L} \mathbf{P}\Big(Q_l < -m^* j - \frac{1}{h^*}\Big(\frac{1}{2}+\varepsilon\Big)\log j\Big).$$

Putting

$$A_i = \Big\{(S_{i+j} - S_i)I_{i,i+j} \geqq m^* j - \frac{1}{h^*}\Big(\frac{1}{2}+\varepsilon\Big)\log j\Big\},$$

we get

$$\mathbf{P}\Big(Q_l \geqq m^* j - \frac{1}{h^*}\Big(\frac{1}{2}+\varepsilon\Big)\log j\Big) = \mathbf{P}\Big(\bigcup_{i\in J_l} A_i\Big) \geqq \Sigma_1 - \Sigma_2,$$

where

$$\Sigma_1 = \sum_{i\in J_l} \mathbf{P}(A_i), \qquad \Sigma_2 = \sum_{i\neq m; i,m\in J_l} \mathbf{P}(A_i A_m).$$

Denote the cardinality of J_l by $|J_l|$. Then $|J_l| \sim j^{1-\varepsilon/2}$ as $j \to \infty$ for all l.

Recall that

$$\mathbf{P}(A_i)=p^j\mathbf{P}(\hat{A}_i), \quad \mathbf{P}(A_iA_m)=p^{m-i+j}\mathbf{P}(\hat{A}_i\hat{A}_m) \quad (i\leqq m\leqq i+j),$$

where

$$\hat{A}_i=\left\{\bar{S}_{i+j}-\bar{S}_i\geqq m^*j-\frac{1}{h^*}\left(\frac{1}{2}+\varepsilon\right)\log j\right\}.$$

By Lemma 1,

$$\Sigma_1=|J_l|\mathbf{P}(A_0)\geqq|J_l|D_9e^{-j/C}j^{\varepsilon}\geqq D_{10}e^{-j/C}j^{1+\varepsilon/2}$$

for large j. We further have

$$\Sigma_2\leqq|J_l|p^j\sum_{r=1}^{|J_l|-1}\mathbf{P}(\hat{A}_i\hat{A}_{i+r[j^{\varepsilon/2}]}).$$

Applying Lemma 3 with $k=j$, $v=r[j^{\varepsilon/2}]$, $u=1/2$, $s=2$, we conclude that

$$\mathbf{P}(\hat{A}_i\hat{A}_{i+r[j^{\varepsilon/2}]})\leqq D_{11}(p^je^{-j/C}r^{-2}j^{-\varepsilon}+\mathbf{P}(\hat{A}_i)j^{5/2+\varepsilon}r^{-\theta}j^{-\theta\varepsilon/2})$$

for large j, where θ is an arbitrary positive constant. By Lemma 1,

$$\sum_{r=1}^{|J_l|-1}\mathbf{P}(\hat{A}_i\hat{A}_{i+r[j^{\varepsilon/2}]})\leqq D_{12}p^je^{-j/C}j^{-\varepsilon}+D_{13}|J_l|p^je^{-j/C}j^{5/2+2\varepsilon}j^{-\theta\varepsilon/2}$$
$$=o(p^jj^{\varepsilon}e^{-j/C})$$

as $j\to\infty$, since θ can be chosen arbitrarily large.

Thus

$$\Sigma_2=o(j^{1+\varepsilon/2}e^{-j/C})=o(\Sigma_1) \text{ as } j\to\infty,$$

and, for large j,

$$\mathbf{P}(Q_l\geqq m^*j-\frac{1}{h^*}\left(\frac{1}{2}+\varepsilon\right)\log j)\geqq D_{14}j^{1+\varepsilon/2}e^{-j/C}.$$

So,

$$\text{(3.20)}\qquad \begin{aligned} P&\leqq\left(1-D_{14}j^{1+\varepsilon/2}e^{-j/C}\right)^{L+1}\\ &\leqq\exp\left\{-D_{14}(L+1)j^{1+\varepsilon/2}e^{-j/C}\right\}\leqq\exp\left\{-D_{15}j^{\varepsilon/2}\right\}\end{aligned}$$

for large j, where we used the inequalities $1-x\leqq\exp(-x)$ and

$$L\geqq\frac{n_j}{2j}-2\geqq D_{16}j^{-1}e^{j/C}$$

for large j. By (3.20), the series in (3.19) converges, which proves Lemma 8.

PROOFS OF THEOREMS. Theorem 2 follows from Lemmas 2 and 4 to 8. Under (1.1) to (1.3), (1.6) and $h^* < h_0$, Theorem 2 implies Theorem 1.

Now we assume that (1.1) to (1.3) and (1.6) hold, but $h^* = h_0$. Put $x = b \log n$, $K = [a \log n]$, where $a > 0$, $0 < b < 1/h^*$. We have

$$\text{(3.21)} \quad \mathbf{P}(M_N < x) \leqq \Big(\mathbf{P}(S_K I_{0,K} < x)\Big)^{[N/K]} \leqq \exp\Big\{-\Big[\frac{N}{K}\Big]\mathbf{P}(S_K I_{0,K} \geqq x)\Big\}.$$

By Chernoff's [1] large deviation theorem,

$$\text{(3.22)} \qquad \frac{1}{K}\log \mathbf{P}(S_K \geqq x)\} \to \log \rho(b/a),$$

where $\rho(\alpha) = \inf_h\{\varphi(h)e^{-h\alpha}\}$ denotes the Chernoff function.

Assume first that $h^* = h_0 < \infty$. For $\alpha \geqq \alpha_c = \mathbf{E}\{X_1|Y_1 = 1\}$ we have $\varphi(h)e^{-h\alpha} \geqq 1$, if $h < 0$, and $\varphi(h) = +\infty$, if $h > h^*$. This implies

$$\rho(\alpha) = \inf_{0 \leqq h \leqq h^*}\{\varphi(h)e^{-h\alpha}\} \geqq e^{-h^*\alpha} \inf_{0\leqq h\leqq h^*} \varphi(h) > 0$$

for all $\alpha \geqq \alpha_c$. By (3.2) and (3.22), we have

$$\log \mathbf{P}(S_K I_{0,K} \geqq x) \sim K \log p\rho(b/a) = K \inf_{0\leqq h\leqq h^*}\{\log p\varphi(h) - hb/a\}$$

$$\geqq (a \inf_{0\leqq h\leqq h^*}\{\log p\varphi(h)\} - h^* b)\log N > (-1+\delta)\log N,$$

if a is chosen small enough. Hence

$$\text{(3.23)} \qquad \sum_{N=1}^{\infty} \mathbf{P}(M_N < x) < \infty.$$

Assume now that $h^* = h_0 = +\infty$. We have, for $x > 0$,

$$\mathbf{P}(X_1 \geqq x) \leqq \varphi(h)e^{-hx} \leqq \frac{1}{p}e^{-hx} \to 0$$

as $h \to +\infty$. It follows that $\mathbf{P}(X_1 \leqq 0|Y_1 = 1) = 1$. So the limit in (1.5) is zero.

If (1.6) does not hold, say $\mathbf{P}(X_1 = d|Y_1 = 1) = 1$, and $d \leqq 0$, then the limit in (1.5) is zero. If $d > 0$, then $1/h^* = -d/\log p$ and

$$\mathbf{P}(S_K I_{0,K} \geqq x) = \mathbf{P}(I_{0,K} = 1) = p^K,$$

if $b < ad$. On choosing a such that $b/d < a < -1/\log p$, we get

$$\log \mathbf{P}(S_K I_{0,K} \geqq x) \sim a \log p \log N.$$

This gives again (3.23).

By the Borel–Cantelli lemma,

$$\mathbf{P}(M_N < b \log N \text{ i.o.}) = 0$$

for all $b < (h^*)^{-1}$, i.e.

$$\liminf_{N\to\infty} \frac{M_N}{\log N} \geqq \frac{1}{h^*} \quad \text{a.s.}. \tag{3.24}$$

To complete the proof of Theorem 1 we need to show that

$$\limsup_{N\to\infty} \frac{M_N}{\log N} \leqq \frac{1}{h^*} \quad \text{a.s.}. \tag{3.25}$$

Put $x = [a \log N]$, $K = [C_2 \log N]$. In the same way as in (3.3) to (3.7), we get

$$\mathbf{P}(M_N > x) \leqq \frac{1}{p} N^{2+C_2 \log p} + N \sum_{j=1}^{K} p^j \mathbf{P}(\bar{S}_j > x).$$

By Markov's inequality and the definition of h^*,

$$\mathbf{P}(\bar{S}_j > x) \leqq e^{-h^* x}(\varphi(h^*))^j \leqq e^{-h^* x} p^{-j}.$$

Hence, for any subsequence $N_k = [\alpha^k], \alpha > 1$ fixed,

$$\sum_{k=1}^{\infty} \mathbf{P}(M_{N_k} > a \log N_k) < \infty,$$

provided $C_2 \log p < -2$ and $ah^* > 1$. By the Borel–Cantelli lemma, this implies

$$\limsup_{k\to\infty} \frac{M_{N_k}}{\log N_k} \leqq \frac{1}{h^*} \quad \text{a.s.}. \tag{3.26}$$

Since M_N is a non-decreasing sequence, (3.26) proves (3.25). A combination of (3.24) and (3.25) completes the proof of Theorem 1.

ACKNOWLEDGEMENTS. This research was carried out when the first two authors visited Marburg. They gratefully acknowledge partial support by a special research grant from the University of Marburg, which made this exchange possible, and by Grant 96-01-00547 from RFFI.

REFERENCES

[1] CHERNOFF, H., A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statist.* **23** (1952), 493–507. *MR* **15**, 241c

[2] 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

[3] 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

[4] DEHEUVELS, P., On the Erdős–Rényi theorem for random fields and sequences and its relationships with the theory of runs and spacings, *Z. Wahrsch. Verw. Gebiete* **70** (1985), 91–115. *MR* **86k**:60045

[5] 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

[6] DEHEUVELS, P. and DEVROYE, L., Limit laws of Erdős–Rényi–Shepp type, *Ann. Probab.* **15** (1987), 1363–1386. *MR* **88f**:60055

[7] DEMBO, A. and KARLIN, S., Strong limit theorems of empirical functionals for large exceedances of partial sums of i.i.d. variables, *Ann. Probab.* **19** (1991), 1737–1755. *MR* **92h**:60044

[8] ERDŐS, P. and RÉNYI, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42** #6907

[9] ERDŐS, P. and RÉVÉSZ, P., On the length of the longest head-run, *Topics in Information Theory* (Second Colloq., Keszthely, 1975), Csiszár, I., Elias, P., eds., Colloq. Math. Soc. J. Bolyai, 16, North-Holland, Amsterdam, 1977, 218–228. *MR* **57** #17788

[10] KARLIN, S. and DEMBO, A., Limit distributions of maximal segmental score among Markov-dependent partial sums, *Adv. in Appl. Probab.* **24** (1992), 113–140. *MR* **93b**:60042

[11] PETROV, V. V., On the probabilities of large deviations for sums of independent random variables, *Teor. Verojatnost. i Primenen.* **10** (1965), 310–322 (in Russian). *MR* **32** #3107. Translation: *Theor. Probab. Appl.* **10** (1965), 287–298.

[12] SHEPP, L. A., A limit law concerning moving averages, *Ann. Math. Statist.* **35** (1964), 424–428. *MR* **29** #4091

(Received April 14, 1997)

FACULTY OF MATHEMATICS AND MECHANICS
ST. PETERSBURG UNIVERSITY
BIBLIOTECHNAYA PL. 2
ST. PETERHOF
RU–198904 ST. PETERSBURG
RUSSIA

Andrei.Frolov@pobox.spbu.ru
Alexander.Martikainen@pobox.spbu.ru

FACHBEREICH MATHEMATIK UND INFORMATIK
PHILIPPS-UNIVERSITÄT MARBURG
HANS-MEERWEIN-STRASSE
D–35032 MARBURG/LAHN
GERMANY

jost@mathematik.uni-marburg.de

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 183–189

A LIMIT LAW RELATED TO THE LAW OF THE ITERATED LOGARITHM

K. GRILL

Dedicated to the memory of Alfréd Rényi

Abstract

We study the upper limiting behaviour of the time the Wiener process spends above a given lower class function.

1. Introduction

Let $(W(t), t \geqq 0)$ be a standard Wiener process. For a function $f(.) > 0$ let

$$U(f,t) = \int_0^t 1_{\{W(t)>f(t)\}}(t)dt. \tag{1}$$

The case where

$$f(t) = \sqrt{2\gamma t \log\log t}$$

with $0 < \gamma \leqq 1$ was studied by Chan [1].

If $\gamma < 1$, Strassen's [3] law implies that

$$\limsup_{t\to\infty} t^{-1}U(f,t) = 1 - \exp\left(-4\left(\frac{1}{\gamma} - 1\right)\right).$$

Furthermore, it is obvious that this result remains true if we only have

$$\frac{f(t)}{\sqrt{2t\log\log t}} \to \sqrt{\gamma}. \tag{2}$$

Chan presents a large deviation law that is closely related to this question. For related results, see Uchiyama [4, 5].

If $\gamma = 1$, Strassen's law implies that the lim sup above is zero; thus, we are left with the question of the right rate of convergence. In this direction, Chan proves that (for $f(t) = \sqrt{2t\log\log t}$)

$$\limsup_{t\to\infty} (\log\log t)^{2/3} t^{-1} U(f,t) < \infty \tag{3}$$

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F17.

Key words and phrases. Functional laws, strong approximation, sums of independent random variables.

and

(4) $$\limsup_{t\to\infty}(\log\log t)t^{-1}U(f,t)=\infty.$$

It is the purpose of the present paper to give the exact limiting behaviour for the case $\gamma=1$. Observe first that we no longer have the same limiting rate for all functions that satisfy

$$\frac{f(t)}{\sqrt{2t\log\log t}}\to 1$$

as in (2) above. This should be clear from the fact that there are functions f satisfying this relation for which we have $W(t)<f(t)$ eventually with probability one. Thus, we are led to have a look at the functions for which $W(t)>f(t)$ infinitely often (here we commit a slight abuse of notation; actually, this is meant as "there is a sequence $t_n\to\infty$ such that $W(t_n)>f(t_n)$"; throughout this paper "infinitely often" will be understood in this sense). This is the subject of the famous Erdős–Feller–Kolmogorov–Petrovski integral test (cf. Feller [2]) which states as follows:

THEOREM A. *Let*

(5) $$f(t)=\sqrt{t}\psi(t),$$

where

(6) $$\psi(t)\uparrow\infty$$

and

(7) $$t^{-1/2}\psi(t)\downarrow 0.$$

Then we have

$$W(t)>f(t)\quad i.o.$$

iff

(8) $$\Lambda(t)=\int_0^t\frac{\psi(u)}{u}\exp\Big(-\frac{\psi(u)^2}{2}\Big)du\to\infty.$$

Thus, from now on, we assume that $f(.)$ satisfies equations (5) to (8). Furthermore, we assume that

(9) $$\frac{\psi(t)}{\sqrt{2\log\log t}}\to 1.$$

Under these assumptions, we have

THEOREM 1. *If f satisfies* (5) *to* (9), *then*

$$\limsup_{t\to\infty} \rho(t)^{-1}U(f,t) = 1,$$

where

$$\rho(t) = \frac{8t\log\Lambda(t)}{\psi(t)^2} \approx \frac{4t\log\Lambda(t)}{\log\log t}$$

(we use $f(t)\approx g(t)$ to signify that $f(t)/g(t)\to 1$ as $t\to\infty$).

REMARK 1. For Chan's original question, i.e., for $\psi(t) = \sqrt{2\log\log t}$, we get

$$\rho(t) \approx \frac{6t\log\log\log t}{\log\log t}.$$

REMARK 2. As $\Lambda(t)\to\infty$, we see that (3) remains true in the general case; the rate of going to infinity there, however, may be arbitrarily slow. On the other hand, there are functions f for which (4) fails.

REMARK 3. If the limit in (9) is less than one, our theorem still gives the right rate but the wrong constant. It would not be too hard to combine Theorem 1 and (1) into one result that covers both cases.

2. Proof of Theorem 1 — upper part

We have to prove that, for any $\epsilon > 0$,

$$\mathbf{P}(U(f,t) > (1+\epsilon)\rho(t) \text{ i.o.}) = 0.$$

Observe that $\rho(t)$ is nondecreasing and, at least for large t, greater than $t/\log\log t$. Thus, if

$$U(f,t) > (1+\epsilon)\rho(t),$$

then there are u and v with

$$\frac{\epsilon t}{3\log\log t} \leq u \leq v \leq t,$$

$$v - u \geqq (1+\frac{2\epsilon}{3})\rho(t),$$

and

$$W(u) > f(u), \quad W(v) > f(v).$$

This implies that

$$v \geqq u + \left(1+\frac{2\epsilon}{3}\right)\rho(u)$$

and

$$v \leqq \frac{4}{\epsilon} u \log\log u$$

if u is large enough. Now, let

$$t_n = \exp\Big(\frac{n}{\log n}\Big).$$

If we have u and v as above, then we can find n such that

$$t_{n-1} < u \leqq t_n.$$

This implies that

$$t_{n-1} + \Big(1 + \frac{2\epsilon}{3}\Big)\rho(t_{n-1}) \leqq v \leqq \frac{4}{\epsilon} t_n \log\log t_n.$$

Now, let

$$A_n = \{\exists u, v : t_{n-1} \leqq u \leqq t_n, t_{n-1} + \Big(1 + \frac{2\epsilon}{3}\Big)\rho(t_{n-1}) \leqq v \leqq \frac{4}{\epsilon} t_n \log\log t_n,$$
$$W(v) \geqq f(v), W(u) \geqq f(u)\}.$$

We have to prove that

$$\mathbf{P}(A_n \text{ i.o.}) = 0.$$

To this end, define

$$\theta = \frac{t_n}{t_{n-1}}$$

and

$$B_{nm} = \{\exists u, v : t_{n-1} \leqq u \leqq t_n, \theta^m t_n \leqq v \leqq \theta^{m+1} t_n,$$
$$W(u) \geqq f(t_{n-1}), W(v) \geqq f(\theta^m t_n)\}.$$

Obviously,

$$\mathbf{P}(A_n) \leqq \sum_{m=m_1}^{m_2} \mathbf{P}(B_{nm}),$$

where

$$m_1 = \frac{\log\Big(1 + \big(1 + \frac{2\epsilon}{3}\big)\frac{\varrho(t_{n-1})}{t_{n-1}}\Big)}{\log\theta}$$

and

$$m_2 = (\log n)^2.$$

By a simple reflection principle argument,

$$\mathbf{P}(B_{nm}) \leqq 4\mathbf{P}(W(t_n) \geqq f(t_{n-1}), W(\theta^{m+1} t_n) \geqq f(\theta^m t_n)) =$$
$$4\mathbf{P}(W(t_{n-1}) \geqq (\theta)^{-1/2} f(t_{n-1}), W(\theta^m t_n) \geqq (\theta)^{-1/2} f(\theta^m t_n)).$$

The latter probability can be estimated above using the following lemma which we state without proof:

LEMMA 1. *If X_1 and X_2 have a joint normal distribution with $\mathbf{E}X_i=0$, $\mathbf{E}X_i^2=1$ $(i=1,2)$ and $\mathbf{E}X_1X_2=\rho>0$, then*

$$(1-\Phi(a))\Big(1-\Phi\Big(\frac{b-\rho a}{\sqrt{1-\rho^2}}\Big)\Big)\leqq \mathbf{P}(X_1\geqq a, X_2\geqq b)$$
$$\leqq C(1-\Phi(a))\Big(1-\Phi\Big(\frac{b-\rho a}{\sqrt{1-\rho^2}}\Big)\Big).$$

(Here and in what follows, C will denote an absolute constant whose actual value may change from one occurrence to the other.)

As a consequence, if $a>0$ and $b>\rho a$,

$$\mathbf{P}(X_1\geqq a, X_2\geqq b)\leqq \frac{C}{a}\exp\Big(-\frac{a^2}{2}\Big)\exp\Big(-\frac{(b-\rho a)^2}{2(1-\rho^2)}\Big).$$

For our purposes, this implies that

$$\mathbf{P}(B_{nm})\leqq C\sqrt{\theta}\frac{1}{\psi(t_{n-1})}\exp\Big(-\frac{\psi(t_{n-1})^2}{2\theta}\Big)\exp\Big(-\frac{\psi(t_{n-1})^2(1-\theta^{-m/2})}{2(1+\theta^{-m/2})\theta}\Big).$$

If n is large enough and ϵ is small enough, we obtain

$$\mathbf{P}(B_{nm})\leqq C\frac{1}{\psi(t_{n-1})}\exp\Big(-\frac{\psi(t_{n-1})^2}{2}\Big)\exp\Big(-\frac{(1-\epsilon/6)m}{4}\Big),$$

if $m<\epsilon\log n$, and

$$\mathbf{P}(B_{nm})\leqq C\frac{1}{\psi\Big(t_{n-1})}\exp(-\frac{\psi(t_{n-1})^2}{2}\Big)n^{-\epsilon/8},$$

if $m\geqq\epsilon\log n$.

This, in turn, implies that

$$\mathbf{P}(A_n)\leqq C\frac{1}{\psi(t_{n-1})}\exp\Big(-\frac{\psi(t_{n-1})^2}{2}\Big)\exp\Big(-\Big(1+\frac{\epsilon}{8}\Big)\log\Lambda(t_{n-1})\Big).$$

Thus, the series

$$\sum\mathbf{P}(A_n)$$

is dominated by the integral

$$C\int\psi(t)^2\exp\Big(-\frac{\psi(t)^2}{2}\Big)(\Lambda(t))^{-(1+\epsilon/8)}dt=C\int(\Lambda(t))^{-(1+\epsilon/8)}d\Lambda(t),$$

so it is convergent, and the Borel–Cantelli lemma implies the upper half of our theorem.

3. Proof of Theorem 1 — lower part

Now, it is our goal to prove that

$$\mathbf{P}(U(f,t) > (1-\epsilon)\rho(t) \text{ i.o.}) = 1.$$

Again, let

$$t_n = \exp\Big(\frac{n}{\log n}\Big).$$

Furthermore, let

$$\tilde{\psi}(t) = \psi(t)\Big(1 + \frac{\rho(t)^2}{t^2}\Big),$$

and

$$\tilde{f}(t) = \sqrt{t}\tilde{\psi}(t).$$

Define events

$$A_n = \{W(t_n) > \sqrt{t_n}\tilde{\psi}(t_n), W(t_n - \Big(1-\frac{\epsilon}{2}\Big)\rho(t_n)) > \sqrt{t_n - \Big(1-\frac{\epsilon}{2}\Big)\rho(t_n)}\tilde{\psi}(t_n)\}.$$

By Levy's arc sine law, we get that

$$\mathbf{P}(B_n) \geqq C\mathbf{P}(A_n),$$

where

$$B_n = A_n \cap \{U(f,t_n) > (1-\epsilon)\rho(t_n)\}.$$

By a lower estimate derived from Lemma 1, we get

$$\mathbf{P}(B_n) > C\frac{1}{\psi(t_n)}\exp\Big(-\frac{\psi(t_n)^2}{2}\Big)(\Lambda(t_n))^{-(1-\epsilon/4)}.$$

Thus, the series $\sum \mathbf{P}(B_n)$ is divergent. In order to prove that $\mathbf{P}(B_n \text{ i.o.}) = 1$, we employ the following version of the Borel–Cantelli lemma:

LEMMA 2. *Let $(A_k, k \in \mathbb{N})$ be a sequence of events satisfying the following conditions*

(i) $\displaystyle\sum_{k=1}^{\infty} \mathbf{P}(A_k) = \infty$

(ii) $\displaystyle\lim_{n\to\infty} \frac{\sum_{k=1}^{n}\sum_{r=1}^{n}\mathbf{P}(A_kA_r)}{\left(\sum_{k=1}^{n}\mathbf{P}(A_k)\right)^2} \leqq M < \infty.$

Then it holds:

$$\mathbf{P}(A_k \text{ i.o.}) \geqq M^{-1}.$$

So, we need to find upper bounds to the probabilities

$$\mathbf{P}(B_n B_{n+k}).$$

First,

$$\mathbf{P}(B_n B_{n+k}) \leqq \mathbf{P}(A_n A_{n+k}).$$

The following estimates are obtained in a similar way as the ones in the proof of the upper half of Theorem 1, so we can forego the somewhat intricate details:

$$\mathbf{P}(A_n A_{n+k}) \leqq \mathbf{P}(A_n)\exp(-C_1 k)$$

if $k < 2\log n$

$$\mathbf{P}(A_n A_{n+k}) \leqq \mathbf{P}(A_n) n^{-C_2}$$

if $2\log n \leqq k \leqq (\log n)^3$, and

$$\mathbf{P}(A_n A_{n+k}) \leqq C_3 \mathbf{P}(A_n)\mathbf{P}(A_{n+k})$$

if $k > (\log n)^3$ (the constants C_1, C_2 and C_3 above may of course depend on ε).

Putting everything together, we find that the events B_n satisfy the hypothesis of Lemma 2, so we find that

$$\mathbf{P}(B_n \text{ i.o.}) > 0.$$

This, together with the zero-one law for the Wiener process, proves the lower part of our theorem.

REFERENCES

[1] Chan, T., Some asymptotic results related to the law of the iterated logarithm for Brownian motion, *J. Theoret. Probab.* **8** (1995), 643–667. *MR* **96j**:60058

[2] Feller, W., The general form of the so-called law of the iterated logarithm, *Trans. Amer. Math. Soc.* **54** (1943), 373–402. *MR* **5**, 125c

[3] Strassen, V., An invariance principle for the law of iterated logarithm, *Z. Wahrsch. Verw. Gebiete* **3** (1964), 211–226. *MR* **30** #5379

[4] Uchiyama, K., Brownian first exit from and sojourn over one sided moving boundary and application, *Z. Wahrsch. Verw. Gebiete* **54** (1980), 75–116. *MR* **82c**:60143

[5] Uchiyama, K., The proportion of Brownian sojourn outside a moving boundary, *Ann. Probab.* **10** (1982), 220–233. *MR* **83g**:60099

(Received May 30, 1997)

INSTITUT FÜR STATISTIK UND
WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
WIEDNER HAUPTSTRASSE 8–10
A–1040 WIEN
AUSTRIA

grill@ci.tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 191–210

ON ESTIMATION OF ANALYTIC FUNCTIONS

I. IBRAGIMOV

To the memory of A. Rényi

1. Introduction and main results

The aim of this paper is to present some results about nonparametric estimation of analytic functions. We consider the following three problems.

PROBLEM I. An observed signal $X_\varepsilon(t)$ on the interval $[a,b]$ is of the form

$$dX_\varepsilon(t) = f(t)dt + \varepsilon dw(t), \quad a \leqq t \leqq b. \tag{1.1}$$

Here $w(t)$ is the Wiener process, ε is a known small parameter and the unknown signal f belongs to a known class F of functions on $[a,b]$. Denote $|\cdot|_p$ the norm in $L_p([a,b])$. Put

$$\Delta_p(\varepsilon; F) = \Delta_p(\varepsilon) = \inf\sup \mathbf{E}_f|\hat{f} - f|_p \tag{1.2}$$

where sup is taken over all $f \in F$ and inf over all possible estimators $\hat{f}$ of f. We are interested in the asymptotic behaviour of $\Delta_p(\varepsilon; F)$ when the level of noise ε goes to zero. The rate of convergence of Δ depends on f. Recall some known results (see [1], ch. 7; [2]).

1. Let F consist of all periodic functions with uniformly bounded in L_p fractional derivative of order β. Then

$$\begin{aligned} \Delta_p(\varepsilon) &\asymp \varepsilon^{2\beta/1+2\beta}, \qquad 2 \leqq p < \infty, \\ \Delta_\infty(\varepsilon) &\asymp \varepsilon^{2\beta/1+2\beta}(\ln 1/\varepsilon)^{2\beta/1+2\beta}. \end{aligned} \tag{1.3}$$

2. Let F consist of all periodic functions f analytic and uniformly bounded inside a strip $|\operatorname{Im} z| < c$, $z = t + is$. Then

$$\begin{aligned} \Delta_p(\varepsilon) &\asymp \varepsilon\sqrt{\ln(1/\varepsilon)}, \qquad 2 \leqq p < \infty, \\ \Delta_\infty(\varepsilon &\asymp \sqrt{\ln(1/\varepsilon)}\sqrt{\ln\ln(1/\varepsilon)}. \end{aligned} \tag{1.4}$$

1991 *Mathematics Subject Classification.* Primary 62G07.

Key words and phrases. Curve estimation, nonparametric estimation, white noise model, density estimation, regression.

Partly supported by the Russian Foundation for Basic Research Grants 96-01-00684, 96-01-00096 and INTAS Grant 950099.

Below we study the following problem. How does $\Delta_p(\varepsilon; F)$ behave when ε goes to zero if F consists of all functions f analytic and uniformly bounded inside some region G of the complex plane, $[a,b] \subset G$?

PROBLEM II. Assume that one observes iid random variables $X_1, X_2, \dots, X_n$ taking values in $[a,b]$ and having a density function $f(x)$ with respect to the Lebesgue measure. As above we define the minimax risk

$$\Delta_p(n; F) = \Delta_p(n) = \inf \sup \mathbf{E}_f |\hat{f} - f|_p, \tag{1.5}$$

where inf is taken over all estimators $\hat{f}$ and sup over all $f \in F$ and study the behaviour of $\Delta_p(n)$ when n goes to infinity. Recall that (see [3]):

1. If F consists of all densities with uniformly bounded in $L_p(R^1)$ fractional derivative of order β, then

$$\begin{aligned} \Delta_p(n) &\asymp n^{-\beta/2\beta+1}, \qquad 1 \le p < \infty, \\ \Delta_\infty(n) &\asymp n^{-\beta/2\beta+1} (\ln n)^{\beta/2\beta+1}. \end{aligned} \tag{1.6}$$

2. If F consists of all densities $f(x)$ analytic and uniformly bounded inside a strip $|\operatorname{Im} z| < c$, $z = x + iy$, then

$$\begin{aligned} \Delta_p(n) &\asymp n^{-1/2} \sqrt{\ln n}, \qquad 1 \le p < \infty, \\ \Delta_\infty(n) &\asymp n^{-1/2} \sqrt{\ln n} \sqrt{\ln \ln n}. \end{aligned} \tag{1.7}$$

In this paper we are interested in the behaviour of $\Delta_p(n)$ when F consists of all functions analytic and uniformly bounded inside a region G of the complex plane, $[a,b] \in G$.

PROBLEM III. Let $f(x)$ be an unknown function on $[a,b]$ belonging to a given class of functions F on $[a,b]$. To estimate f one makes n observations of the function f at the points $X_1, X_2, \dots, X_n$ and observes

$$Y_j = f(X_j) + G_j(X_j, \omega),$$

where $\mathbf{E}(G_j(X_j, \omega) \mid X_1, \dots, X_{j-1}) = 0$ and the noise variables G_j are conditionally independent under a given observation plan (see details in [4]).

Let

$$\Delta_p(n; F) = \Delta_p(n) = \inf \sup \mathbf{E}_f |\hat{f} - f|_p, \tag{1.8}$$

where inf is taken over all admissible observation plans and all possible estimators and sup is taken over all $f \in F$ (see [4]). It has been proved in [4], [5] that

1. If F consists of all periodic functions with uniformly bounded in L_p fractional derivatives of order β, then

$$\Delta_p(n) \asymp n^{-\beta/2\beta+1}, \qquad p < \infty,$$
$$\Delta_\infty(n) \asymp n^{-\beta/2\beta+1}(\ln n)^{\beta/2\beta+1}. \tag{1.9}$$

2. If f consists of all periodic functions analytic and uniformly bounded inside a strip $|\operatorname{Im} z| < c$, then

$$\Delta_p(n) \asymp n^{-1/2}\sqrt{\ln n}, \qquad p < \infty,$$
$$\Delta_\infty(n) \asymp n^{-1/2}\sqrt{\ln n}\sqrt{\ln \ln n}. \tag{1.10}$$

We are again interested in the behaviour of $\Delta_p(n)$ when F consists of functions analytic inside a bounded region G, $[a,b] \in G$.

Notice that in all the results which have been cited above the function f satisfies proper regularity conditions at all real line. All the results (1.3), (1.4), (1.6), (1.7), (1.9), (1.10) are absolutely similar. In this paper we impose regularity conditions on the behaviour of f in a vicinity of the interval $[a,b]$ only. Of course, functions smooth on the interval $[a,b]$ can be smoothly extended onto R^1 and nothing will happen with the formulas (1.3), (1.6), (1.9). The situation with analytic functions is different, functions analytic in G cannot necessarily be extended analytically into a vicinity of the real line and the formulas (1.4), (1.7), (1.10) may change. The results below show that they really change. Moreover, the behaviour of Δ is not similar to the problems I and III.

THEOREM 1. *Let the expression $\Delta_p(\varepsilon; F)$ be defined by (1.2), where the set F consists of all functions f analytic in some bounded region G, $[a,b] \subset G$, and bounded there by a common constant M. Then*

$$\Delta_p(\varepsilon) \asymp \varepsilon\sqrt{\ln 1/\varepsilon}, \qquad 2 \leqq p < 4,$$
$$\Delta_4(\varepsilon) \asymp \sqrt{\ln(1/\varepsilon)}(\ln\ln(1/\varepsilon))^{1/4}, \tag{1.11}$$
$$\Delta_p(\varepsilon) \asymp \varepsilon(\ln(1/\varepsilon))^{1/2/p}, \qquad p > 4.$$

THEOREM 2. *Let the expression $\Delta_p(n; F)$ be defined by (1.5), where the set F consists of all functions f analytic in some bounded region G, $[a,b] \subset G$, and bounded there by a common constant M. Then*

$$\Delta_p(n) \asymp n^{-1/2}\sqrt{\ln n}, \qquad 1 \leqq p < 4,$$
$$\Delta_4(n) \asymp n^{-1/2}\sqrt{\ln n}(\ln\ln n)^{1/4}, \tag{1.12}$$
$$\Delta_p(n) \asymp n^{-1/2}(\ln n)^{1-2/p}, \qquad p > 4.$$

THEOREM 3. *Let the expression $\Delta_p(n; F)$ be defined by (1.8), where the set F consists of all functions f analytic in some bounded region G, $[a,b] \subset G$,* *and bounded there by a common constant M. Suppose also that G_j have finite moments of all orders. Then*

$$\Delta_p(n) \asymp n^{-1/2}\sqrt{\ln n}, \qquad p < \infty, \tag{1.13}$$
$$\Delta_\infty(n) \asymp n^{-1/2}\sqrt{\ln n}(\ln\ln n)^{1/2}.$$

Thus the results are really different from those one usually has in such kind of problems. At first, in the problems I, II even the order of convergence Δ_p depends on p; further, the exceptional value of p now is not ∞ but 4; and at last, and it is unexpected, the behaviour of Δ_p is different for problems I and III.

The rest of the paper is devoted to the proofs of the theorems. The analysis of Cases I and II is very similar and we give a sufficiently detailed proof only of Theorem 1 and do not go into details in Case II. Without loss of generality we may and will suppose that $[a, b] = [-1, 1]$.

Below we denote constants by C or c, i.e. quantities which do not depend on parameters under consideration; they may be different even inside the same formula. But sometimes we supply these constants by indices.

If G is a region in the complex plane we denote by $A(G)$ a class of functions analytic in G and uniformly bounded there by a common constant M.

2. Proof of Theorem 1

2.1. Upper bounds. The set G contains inside itself an ellipse E with the foci at the points ± 1 and the sum of halfaxes equal to $R > 1$. Any function f from F belongs also to $A(E)$ and can be represented by Fourier series with respect to the orthonormal Legendre polynomials in the form

$$f(t) = \sum_0^\infty a_j P_j(t). \tag{2.1}$$

The value of the best approximation of the function f in the L_2-norm by polynomials Q of degree n is equal to

$$\sqrt{\sum_{n+1}^\infty a_j{}^2} = \left(\inf_Q \int_{-1}^1 |f(t) - Q(t)|^2 dt\right)^{1/2}$$
$$\le \left(\inf_Q \int_{-1}^1 |f(t) - Q(t)|^2 (1 - t^2)^{-1/2} dt\right)^{1/2}$$
$$\le \left(\frac{\pi}{R^2 - 1}\right)^{1/2} \frac{M}{R^n},$$

where $M = \max(|f(z)| : z \in E)$ (see, e.g., [6]). Hence

$$|a_k| \leqq \left(\frac{\pi}{2(R-1)}\right)^{1/2} MR^{-k} = cR^{-k} = ce^{-\gamma k}. \tag{2.2}$$

Consider now the estimators $f_N(t)$ for f defined by the formula

$$f_N(t) = \sum_0^N \hat{a}_k P_k(t),$$

where

$$\hat{a}_k = \int_{-1}^{1} P_k(t) dX_\varepsilon(t) = a_k + \varepsilon z_k.$$

The random variables $z_k = \int_{-1}^{1} P_k(t) dw(t)$ are iid standard Gaussian variables.

The Legendre polynomials satisfy on the interval $[-1, 1]$ the inequalities

$$\begin{aligned} |P_k(t)| &\leqq P_k(1) = \sqrt{\frac{2k+1}{2}}, \\ |P_k(t)| &\leqq \sqrt{\frac{2}{\pi}}\sqrt{\frac{2k+1}{2k}}(1-t^2)^{-1/4} \leqq \sqrt{\frac{3}{\pi}}(1-t^2)^{-1/4} \end{aligned} \tag{2.3}$$

(see [7], Theorems 7.3.1, 7.3.3). It follows from (2.2) and (2.3) that

$$\mathbf{E}|\hat{f}_N - f|_p \leqq c\sqrt{N}e^{-\gamma N} + \varepsilon |Z_N|_p, \tag{2.4}$$

where $Z_N(t)$ is the random polynomial of degree N

$$Z_N(t) = \sum_0^N z_k P_k(t).$$

It follows from (2.3) that

$$\mathbf{E}{Z_N}^2(t) \leqq \begin{cases} 3N\pi^{-1}(1-t^2)^{-1/2}, & \\ N^2, & \text{if } N > 2. \end{cases} \tag{2.5}$$

Hence for $p < 4$

$$\mathbf{E}|Z_N|_p^p \leqq c_p \int_{-1}^{1} \left(\sum {P_k}^2(t)\right)^{p/2} dt \leqq c_p N^{p/2}. \tag{2.6}$$

If $p = 4$, then

$$(2.7)\quad \mathbf{E}|Z_n|_p^p \leqq c_p\Big(N^4\int_{-1}^{-1+N^{-2}} dt + N^4\int_{1-N^{-2}}^{1} dt + N^2\int_{N^{-2}}^{1} t^{-1}dt\Big) \leqq cN^2\ln N.$$

If $p > 4$, then

$$(2.8)\qquad \mathbf{E}|Z_N|_p^p \leqq c_p\Big(N^{p-2} + N^{p/2}\int_{N^{-2}}^{1} t^{-p/4}dt\Big) \leqq c_pN^{p-2}.$$

Combining these inequalities with (2.4) we find that

$$(2.9)\qquad \begin{aligned} &\mathbf{E}|f_N - f|_p \leqq c_p(\sqrt{N}e^{-\gamma N} + \varepsilon\sqrt{N}), \quad p < 4,\\ &\mathbf{E}|f_N - f|_4 \leqq c_4\left(\sqrt{N}e^{-\gamma N} + \varepsilon(\ln N)^{1/4}\sqrt{N}\right),\\ &\mathbf{E}|f_N - f|_p \leqq c_p\left(\sqrt{N}e^{-\gamma N} + \varepsilon N^{1-2/p}\right), \qquad 4 < p < \infty. \end{aligned}$$

Take here $N = [(\ln 1/\varepsilon)/\gamma]$ and denote f_ε the estimator F_N with such N. For these estimators

$$(2.10)\qquad \begin{aligned} &\mathbf{E}_f|f_\varepsilon - f|_p \leqq c_p\varepsilon(\ln(1/\varepsilon))^{1/2}, && p < 4,\\ &\mathbf{E}_f|f_\varepsilon - f| \leqq c_p\varepsilon(\ln(1/\varepsilon)^{1/2}(\ln\ln(1/\varepsilon))^{1/4}, && p = 4,\\ &\mathbf{E}_f|f_\varepsilon - f|_p \leqq c_p\varepsilon(\ln(1/\varepsilon))^{1-2/p}, && 4 < p < \infty. \end{aligned}$$

To treat the case $p = \infty$ we need the following result.

LEMMA 2.1 (see [8]). *Let Q be an algebraic polynomial of degree n. Then*

$$(2.11)\qquad |Q|_\infty \leqq (p+1)^{1/p}n^{2/p}|Q|_p.$$

The last inequality together with (2.8) gives that

$$\mathbf{E}|Z_N|_\infty \leqq c_pN^{2/p}\mathbf{E}|Z_N|_p \leqq c_pN$$

and hence

$$(2.12)\qquad \mathbf{E}_f|f_\varepsilon - f| \leqq c\varepsilon\ \ln(1/\varepsilon).$$

The inequalities (2.10), (2.11) prove the part of Theorem 1 concerning the upper bounds.

2.2. Lower bounds. We proceed in the following way. Evidently for any $\delta > 0$

$$\Delta_p(\varepsilon) \geqq \delta\cdot\inf_T\sup_{f\in F} P_f\{|T - f|_p \geqq \delta\}.$$

Let now $S=\{f_j; j=1,\dots,M\}$ be a family of functions $f\in F$ such that $|f_i-f_j|_p>2\delta$ for any $i\neq j$. Then

$$\sup_{f\in F}\mathbf{P}_f\{|T-f|_p\geqq\delta\}\geqq\frac{1}{M}\sum\mathbf{P}_{f_i}\{|T-f_i|_p\geqq\delta\}.$$

The right-hand side can be bounded by Fano lemma (see [1]) and we find that

$$\Delta_p(\varepsilon)\geqq\sup_{d>0}\delta\left(1-\frac{C_\varepsilon(S)+1}{\ln M-1}\right), \tag{2.13}$$

where the Shannon's capacity

$$C_\varepsilon(S)\leqq\sup_{f\in S}\left(\frac{|f|_2^2}{2\varepsilon^2}\right)$$

(see for details of these arguments in [1], pp. 355, 356). The construction of the set S depends on p and we consider separately three cases.

1. Estimation of $\Delta_p(\varepsilon)$ for $p<4$. Consider functions $f(t)$ of the form

$$f_a(t)=e^{-\gamma N}\sum_0^N a_jP_j(t),\qquad a_j=\pm1.$$

The polynomials $P_j(z)$ for $z\notin[-1,1]$ satisfy the inequlity (see [7])

$$|P_n(z)|\leqq c\sqrt{n}\left|z+(z^2-1)^{1/2}\right|^n.$$

Hence one can choose γ in such a way that all functions $f_a(z)$ are bounded by a constant M for $z\in E$, where $E\supset G$ is an ellipse with the foci at the points ±1. Hence all such functions $f_a\in F$. Evidently

$$|f_a|_2^2=(N+1)e^{-2\gamma N}.$$

LEMMA 2.2. *Let $A=A(d)$ be the set of vectors $a=(a_0,\dots,a_N)$ such that for any $a,a'\in A$*

$$\sum|a_j-a'_j|>2d.$$

Then for the number of elements A we have

$$\operatorname{card}(A)\geqq2^{N+1}\left[\sum_0^{d-1}\binom{N+1}{i}\right]^{-1}. \tag{2.14}$$

Inequality (2.14) is called Gilbert or Varshamov–Gilbert bound (see [9]).

LEMMA 2.3. *Let d in* (2.14) *be equal to* $[N/4]$. *Then for all $N>N_0$*

$$\text{card}\,(A) > \exp\{-N/8\}. \tag{2.15}$$

The proof of (2.15) can be achieved by elementary estimation of $\sum_0^{N/4} \binom{N+1}{i}$ or derived from some elementary results about probabilities of large deviations.

Let now $S = \{f_a, a \in A\}$, where $A = A([N/4])$. We have for any two $f_a, f_{a'}$

$$\begin{aligned}|f_a - f_{a'}|_p &\geqq 2^{1/p-1/2}|f_1 - f_{a'}|_2 \\ &= 2^{1/p-1/2}\Big(\sum |a_j - a'_j|^2\Big)^{1/2} e^{-\gamma N} \\ &\geqq 2^{1/2+1/p}(N+1)^{1/2}e^{-\gamma N} \geqq cN^{1/2}e^{-\gamma N},\end{aligned}$$

where $c > 0$. It follows from (2.13) that for $2 \leqq p < 4$

$$\begin{aligned}\Delta_p(\varepsilon) &\geqq c_1\sqrt{N}e^{-\gamma N}\left(1 - c_2 N(\varepsilon^2 e^{2\gamma N}\text{card}\,A(N/4))^{-1}\right) \\ &\geqq c_1\sqrt{N}e^{-\gamma N}(1 - c_3\varepsilon^{-2}e^{-2\gamma N}).\end{aligned}$$

Take here $N \sim c\gamma^{-1}\ln(1/\varepsilon)$. We find then that

$$\Delta_p(\varepsilon) \geqq c\varepsilon\sqrt{\ln 1/\varepsilon}, \quad c > 0. \tag{2.16}$$

Consider now the case $p > 4$. Let λ be a small positive number. Define the functions

$$\begin{aligned}f_0(t) &= \exp\{-\gamma N\} \sum_{N/2 \leqq j \leqq N} P_j(t), \\ f_1(t) &= \exp\{-\gamma N\} \sum_{N/2 \leqq j \leqq N - N\lambda} P_j(t), \\ &\dots\dots\end{aligned} \tag{2.17}$$

The collection of functions f_j will constitute the set S. The number of points in this set is close to $N/2\lambda$. Evidently

$$|f_j|_2^2 \leqq N\exp\{-2\gamma N\}.$$

For any $f_i, f_j \in S$

$$|f_i - f_j|_p \geqq \exp\{-\gamma N\}\left(\int_{-1}^{1}\Big|\sum_{k \in I(i,j)} P_k(t)\Big|^p\right)^{1/p} dt,$$

where $I(i,j)$ is the interval of the type $[xN, yN]$ and $y - x \geqq \lambda$.

The Legendre polynomials satisfy the inequality

$$\max_{-1\leqq t\leqq 1} |P_k(t)| = P(1) = \sqrt{\frac{2k+1}{2}}. \tag{2.18}$$

By V. Markov's inequality (see [8])

$$|P_k'|_\infty \leqq ck^2 |P_k|_\infty \leqq ck^{5/2}. \tag{2.19}$$

Hence one can find $\theta > 0$ such that for $t \in [1, 1-\theta N^{-2}]$ and $N/2 \leqq k \leqq N$

$$P_k(t) \geqq 2^{-1}\sqrt{N}. \tag{2.20}$$

It follows that for all $f_i, f_j \in S$

$$|f_i - f_j|_p \geqq c\lambda N^{3/2-2/p} e^{-\gamma N}.$$

Apply now (2.13). We find that

$$\Delta_p(\varepsilon) \geqq c\lambda N^{3/2-2/p} e^{-\gamma N}(1 - cN\varepsilon^{-2}e^{-2\gamma N}(\ln 1/\lambda)^{-1}).$$

Choosing here N in such a way that the expression in the brackets is close to $1/2$,

$$N = \frac{1}{\gamma}\ln(1/\varepsilon) + \frac{1}{2\gamma}\ln\ln(1/\varepsilon) + \dots,$$

we find that

$$\Delta_p(\varepsilon) \geqq c\varepsilon(\ln(1/\varepsilon))^{1-2/p}.$$

Consider now the last and the most complicated case $p = 4$ which needs a special treatment (cf. [1], Theorem 2.12.1). Denote Γ the set of functions $f(t) = \sum_0^N a_j P_j(t)$, where the vectors $a = (a_0, \dots, a_N)$ run the ball $\left\{a : \sum a_j^2 \leqq N^4\varepsilon^2\right\}$ in R^{N+1}. We will denote by Γ this a-set also. If N is chosen from the relation $\varepsilon \sim e^{-\gamma N}$ then under a proper choice of γ the set of functions $\Gamma \subset F$. Hence

$$\Delta_4(\varepsilon) = \Delta(\varepsilon) \geqq \inf_t \frac{1}{\operatorname{mes}(\Gamma)} \int_\Gamma \mathbf{E}_a \Big|\sum_0^N (a_j - t_j)P_j\Big|_4 da$$

$$= \inf \frac{1}{\operatorname{mes}(\Gamma)} \int_\Gamma \mathbf{E}_0 \Big\{ \Big|\sum_0^N (a_j - t_j)P_j\Big|_4 \times \exp\Big\{-|a|^2/2\varepsilon^2 + \varepsilon^{-2}\int_{-1}^1 f(t)\,dw(t)\Big\}\Big\} da. \tag{2.21}$$

We have from the last relation that
(2.22)
$$\Delta(\varepsilon) \geqq \mathbf{E}\Big\{\exp\Big\{\frac{1}{2\varepsilon^2}\sum_0^N \xi_j^2\Big\}\frac{1}{\operatorname{mes}(\Gamma)}\int_{\Gamma-\xi}\Big|\sum_0^N (a_j+\xi_j-t_j)P_j\Big|_4 e^{-|a|^2/2}da\Big\},$$

where $\xi = (\xi_0 \ldots \xi_N)$ denotes the standard Gaussian vector.

Consider now the set $A = \{x = (x_0, \ldots, x_n) : |x| \leqq \varepsilon N^2 - \varepsilon N\}$ and the set $\Gamma_0 = \{a : |a| \leqq \varepsilon N\}$. Then the inequality (2.22) gives that

$$\Delta(\varepsilon) \geqq (\operatorname{mes}(\Gamma))^{-1}(2\pi)^{-(N+1)/2}e^{-(N+1)} \times \int_A d\xi \int_{\Gamma_0}\Big|\sum (a_j+\xi_j-t_j)\Big|_4 e^{-|a|^2/2}da.$$

If we apply Anderson's lemma to the last integral (see [1]), we find that

$$\Delta(\varepsilon) \geqq \varepsilon(\operatorname{mes}(\Gamma))^{-1}\int_A (2\pi)^{-(N+1)/2}d\xi \int_{\varepsilon^{-1}\Gamma_0}\Big|\sum a_jP_j\Big|_4 e^{-|a|^2/2}da.$$

Denote B the complement of the set $\varepsilon^{-1}\Gamma_0$. The integral

$$(2\pi)^{-(N+1)/2}\int_B \Big|\sum a_jP_j\Big|_4 e^{-|a|^2/2}da$$
$$\leqq cN^{3/4}\int_N^\infty r^{N+1}e^{-r^2/2}dr \leqq cN^N e^{-N^2/4} < c\varepsilon^2.$$

Hence

(2.23)
$$\Delta(\varepsilon) \geqq \varepsilon\frac{\operatorname{mes}(A)}{\operatorname{mes}(\Gamma)}(2\pi)^{-(N+1)/2}\int_\infty^\infty \Big|\sum a_jP_j\Big|_4 e^{-|a|^2/2}da + O(\varepsilon^2)$$
$$\geqq 4\varepsilon\mathbf{E}\Big|\sum \xi_jP_j\Big|_4 + O(\varepsilon^2),$$

where $\xi = (\xi_0 \ldots \xi_n)$ denotes again the standard Gaussian vector.

Finally, prove that

$$\mathbf{E}\Big|\sum_0^N \xi_jP_j\Big|_4 \geqq c\sqrt{N}(\ln N)^{1/4}, \quad c > 0. \tag{2.24}$$

Let λ be a small positive number. Consider the probability

$$Q = \mathbf{P}\left\{\left|\sum_0^N \xi_j P_j\right|_4 \leqq \lambda^{1/4}\left(\int_{-1}^{1}\left(\sum P_j^2(x)\right)^2 dx\right)^{1/4}\right\}.$$

Since

$$\mathbf{E}\left|\sum \xi_j P_j\right|_4^4 = 3\int_{-1}^{1}\left(\sum P_j^2(x)\right)^2 dx,$$

this probability is

$$Q \leqq \frac{\operatorname{Var}\left(|\sum \xi_j P_j|_4^4\right)}{(3-\lambda)^2\left[\int_{-1}^{1}\left(\sum_0^N P_j^2(x)dx\right)^2\right]^2}.$$

Recall some asymptotic formulas for the Legendre polynomials (see [7], ch. 8):

The Laplace formula: uniformly on $\varepsilon \leqq \theta \leqq \pi - \varepsilon$

$$P_n(\cos\theta) = \sqrt{\frac{2n+1}{\pi n \sin\theta}}\cos\{(n+1/2)\theta - \pi/4\} + O(n^{-1/2}); \tag{2.26}$$

and the Hilb formula: uniformly on $0 \leqq \theta \leqq \pi - \varepsilon$

$$P_n(\cos\theta) = (n+1/2)^{1/2}(\theta/\sin\theta)^{1/2} J_0\{(n+1/2)\theta\} + O(n^{-1/2}), \tag{2.27}$$

where $J_0(x)$ is the Bessel function.

It follows from (2.26) that the numerator in (2.25)

$$(3-\lambda)^2\left[\int_{-1}^{1}\left(\sum P_j^2(x)dx\right)^2\right]^2 \geqq cN^4 \ln^2 N, \qquad c > 0, \tag{2.28}$$

the numerator in (2.25) is the combination of the summands of the type

$$I(a,b) = \int_{-1}^{1}\int_{-1}^{1}\left(\sum P_j(x)P_j(y)\right)^a\left(\sum P_j^2(x)\right)^b\left(\sum P_j^2(y)\right)^b dx\,dy,$$

where $(a,b) = (2,1)$ or $(4,0)$.

Applying the Hilb formula (2.27) we find that

$$I(a,b) = o(N^4 \ln^2 N).$$

Thus $Q = o(1)$ and hence

$$\mathbf{E}\left\{\frac{\left|\sum_0^N \xi_j P_j\right|_4}{\left[\int_{-1}^{1}\left(\sum P_j^2\right)^2 dx\right]^{1/4}}\right\} \geqq \lambda^{1/4}(1+o(1)).$$

The last inequality and (2.28) prove (2.24).

Since $N \sim \ln 1/\varepsilon$, it follows from (2.23) and (2.24) that

$$\Delta_4(\varepsilon) \geqq c\varepsilon(\ln 1/\varepsilon)^{1/2}(\ln\ln 1/\varepsilon)^{1/4}.$$

Putting together all upper and lower bounds for $\Delta_p(\varepsilon)$ we prove Theorem 1.

3. Proof of Theorem 2

3.1. Upper bounds. In general we follow the arguments of Section 2.1. Namely, we consider again the expansion of the density function $f(x)$ into Fourier series with respect to the orthonormal Legendre polynomials

$$f(x) = \sum_{j=0}^{\infty} a_j P_j(x). \tag{3.1}$$

All these series converge uniformly with respect to f inside some ellipse E. Estimate the coefficients a_j by the statistics

$$\hat{a}_j = \frac{1}{n}\sum_{i=1}^{n} P_j(X_i), \tag{3.2}$$

and the density $f(x)$ by

$$f_N(x) = \sum_0^N \hat{a}_j P_j(x). \tag{3.3}$$

We have

$$f_N(x) - f(x) = [\mathbf{E} f_N(x) - f(x)] + \xi_N(x),$$

where $\xi_N(x)$ is the random polynomial of degree N

$$\xi_N(x) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{N} P(x)[P_j(X_i) - \mathbf{E}P_j(X_i)]. \tag{3.4}$$

Thus

$$\mathbf{E}|f_N - f|_p \leqq |\mathbf{E}f_N - f|_p + |\xi_N|_p. \tag{3.5}$$

The first member on the right-hand side

$$|\mathbf{E}f_N - f|_p = \Big|\sum_{N+1}^{\infty} a_j P_j\Big|_p \leqq c\sqrt{N}e^{-\gamma N}, \tag{3.6}$$

where $\gamma > 0$ depends on the ellipse E (and hence on G) (see Section 2.1).

The estimation of $\mathbf{E}|\xi_N|_p$ depends on p. Note at first that

$$\mathbf{E}_f P_j(X_i) = a_j = O(e^{-\gamma j}).$$

Hence

$$\mathbf{E}\xi_N^2(x) = \frac{1}{n}\mathbf{E}\Big(\sum_0^N P_j(x)P_j(X_1)\Big)^2 + O(n^{-1}),$$

where the $O(\cdot)$ term is uniformly bounded with respect to N. Thus

$$\begin{aligned}\mathbf{E}_f|\xi_N|_2^2 &= n^{-1}\int_{-1}^{1} f(x)\sum_0^N P_j^2(x)dx + O(n^{-1}) \\ &\leqq c|f|_\infty N n^{-1}.\end{aligned} \tag{3.7}$$

It follows that for $1 \leqq p \leqq 2$

$$\mathbf{E}|f_N - f|_p \leqq c\left(e^{-\gamma N}\sqrt{N} + \sqrt{N/n}\right).$$

Put here $N \sim \ln n/\gamma$. We find then that

$$\Delta_p(n) \leqq (\ln n/n)^{1/2}, \qquad p \leqq 2. \tag{3.8}$$

LEMMA 3.1 (H. Rosenthal). *Let $Z_1, \dots, Z_n$ be independent random variables. Let $\mathbf{E}Z_k = 0$ and let $p \geqq 2$. Then*

$$\begin{aligned}&\mathbf{E}\Big|\sum_1^n Z_k\Big|^p \\ &\leqq c\Big(\sum_1^n \mathbf{E}|Z_k|^p + \Big(\sum_1^n \mathbf{E}Z_k^2\Big)^{p/2}\Big),\end{aligned} \tag{3.9}$$

where c is a positive constant depending only on p.

The proof of the Lemma can be found in [11], Section 2.3.

We now return to the estimation of $\mathbf{E}|\xi_N|_p$. It follows from (3.9) that

$$\begin{aligned}\mathbf{E}|\xi_N|_p^p \leqq cn^{-p}\Big\{ & n\int_{-1}^{1}\mathbf{E}\Big|\sum_0^N P_j(x)(P_j(X_1) - \mathbf{E}P_j(X_1))|^p dx \\ &+ n^{p/2}\int_{-1}^{1}\Big(\mathbf{E}\Big|\sum_0^N P_j(x)P_j(X_1)\Big|^2\Big)^{p/2}dx\Big\}.\end{aligned} \tag{3.10}$$

By a formula of Laplace ([7], Theorem 8.21.1) uniformly outside any region including the interval $[-1,1]$ when n goes to infinity

$$P_n(z) \sim (2\pi)^{-1/2}(z^2-1)^{-1/4}(z+(z^2-1)^{1/2})^{n+1/2}.$$

Hence by (2.2)

$$\left|\int_{-1}^{1} P_i(x)P_j(x)f(x)dx\right| \leqq c_1 e^{-c|j-i|}, \qquad c>0,$$

and

$$\int_{-1}^{1}\left(\mathbf{E}\left|\sum_{1}^{N} P_j(x)P_j(X_1)\right|^2\right)^{p/2} dx$$

$$(3.11) \qquad = \int_{-1}^{1}\Big(\sum_{i,j=0}^{N} P_i(x)P_j(x)\int_{-1}^{1} P_i(y)P_j(y)f(y)dy\Big)^{p/2}$$

$$\leqq c^p \int_{-1}^{1}\Big(\sum_{0}^{N} P_i^2(x)\Big)^{p/2} dx.$$

Finally

$$\mathbf{E}\left|\sum_{0}^{N} P_j(x)[P_j(x)-\mathbf{E}P_j(X_1)]\right|^p \leqq cN^{2p}.$$

Thus, for $p>2$

$$(3.12) \quad \mathbf{E}|\xi_N|_p \leqq \mathbf{E}^{1/p}|\xi_N|_p^p \leqq cN^2 n^{-1+1/p} + n^{-1/2}\left(\int_{-1}^{1}\Big(\sum_{0}^{N} P_i^2(x)\Big)^{p/2}dx\right)^{1/p}.$$

The integral on the right-hand side has been estimated in Section 2. Applying these estimates and taking $N \sim \ln n$ we find that

$$(3.13) \qquad \begin{aligned} \Delta_p(n) &\leqq c_p n^{-1/2}(\ln n)^{1/2}, && 2\leqq p<4,\\ \Delta_4 &\leqq cn^{-1/2}(\ln n)^{1/2}(\ln\ln n)^{1/4}, && p=4,\\ \Delta_p(n) &\leqq c_p n^{-1/2}(\ln n)^{1-2/p}, && 4<p<\infty. \end{aligned}$$

To treat the case $p=\infty$ we apply again the inequality (2.10). We have then from (3.12) for some $p>4$ and $N\sim \ln n$

$$(3.14) \qquad \begin{aligned} \mathbf{E}|\xi_N|_\infty &\leqq (p+1)^{1/p}N^{2/p}\mathbf{E}|\xi_N|_p\\ &\leqq c(N^{2+2/p}n^{-1+1/p}+n^{-1/2}N^{2/p}) \leqq cn^{-1/2}\ln n. \end{aligned}$$

2.2. Lower bounds. The proofs are similar to those of the Section 2 only instead of (2.13) we apply the following result.

LEMMA 3.2 ([3], Theorem 3.1). *Assume that there are $N(\delta)$ densities $f_{i\delta} \in \Sigma$, $i = 1, \ldots$ such that $|f_{i\delta} - f_{j\delta}|_p \geqq \delta$. Let $\{f_{0\delta}\}$ be an arbitrary family of densities, $\delta > 0$. Let the set $\delta(n, \Sigma)$ be defined as*

$$\delta(n, \Sigma) = \sup \left\{ \delta : \frac{1}{\ln N(\delta)} \max_i \left| \frac{f_{i\delta} - f_{j\delta}}{\sqrt{f_{0\delta}}} \right|_2^2 \leq \frac{1}{2n} \right\}.$$

Then for any estimator $\hat{f}_n$ of f

$$\sup_{f \in \Sigma} \mathbf{E}_f |\hat{f}_n - f|_p \geqq 1/4\, \delta(n, \Sigma). \tag{3.15}$$

The construction of the set $\{f_{i\delta}\}$ depends on p. If $p < 4$ the set $\{f_{i\delta}\}$ consists of functions

$$f_a(x) = 1/2 + e^{-\gamma N} \sum_1^N a_j P_j(x), \qquad a_j = \pm 1$$

and $f_{0\delta}(x) = 1/2$. The further arguments coincide with the arguments of Section 2.2, the case $p < 4$, and we omit them. The final result is

$$\Delta_p(n) > c_p (n^{-1} \ln n)^{1/2}, \qquad c_p > 0, \quad p < 4. \tag{3.16}$$

If $p > 4$ the set $\{f_{i\delta}\}$ consists of the functions

$$1/2 + f_j(x), \qquad j = 0, 1 \ldots$$

and the functions $f_j(x)$ are defined by (2.17). The same arguments as in the Section 2.2 and Lemma 3.2 show that

$$\Delta_p(n) \geqq c_p n^{-1/2} (\ln n)^{1 - 2/p}, \qquad c_p > 0, \quad 4 < p \leqq \infty. \tag{3.17}$$

The more complicated case $p = 4$ again as in the Section 2.2 needs special arguments (also in the spirit of the Section 2.2, $p = 4$). Namely, consider the set Γ of densities

$$f_a(x) = 1/2 + \frac{1}{\sqrt{n}} \sum_1^N a_j P_j(x),$$

where the vectors $a = (a_1, \ldots, a_N)$ run the ball $\left\{ a : \sum_1^N a_j^2 \leq N^4 \right\}$. If N is chosen from the relation $N \sim c \ln n$, the set $\Gamma \subseteqq F$. Hence

$$\Delta_4(n) \geqq \inf_t \frac{1}{\operatorname{mes}\Gamma} \int_\Gamma \mathbf{E}_a \Big| \sum_1^N (a_j - t_j) P_j \Big|_4 da$$

$$= \inf_t \frac{1}{\operatorname{mes}\Gamma} \int_\Gamma \mathbf{E}_0 \Big\{ \Big| \sum_1^N (a_j - t_j) P_j \Big|_4 \prod_{i=1}^n \Big(1 + \frac{2}{\sqrt{n}} \sum_{j=1}^N a_j P_j(X_i)\Big) \Big\} da.$$

The rest of the proof is the combination of arguments of the Section 2.2, the case $p=4$, and the arguments used to prove the Hájek–LeCam minimax theorem (see [1], Section 2.12) therefore we omit it. The final result is that, as in Section 2.2,

$$\Delta_4(n) \geqq cn^{-1/2} \mathbf{E} \Big| \sum_1^N \xi_j P_j \Big|_4,$$

and $(\xi_1, \ldots, \xi_N)$ is the standard Gaussian vector. We have seen that the last inequality implies

$$\Delta_4(n) \geqq cn^{-1/2} (\ln n)^{1/2} (\ln \ln n)^{1/4}. \tag{3.18}$$

The upper bounds (3.13) and (3.14) and the lower bounds (3.16)–(3.18) prove the theorem.

4. Proof of the Theorem 3

We need to prove the upper bounds only; the lower bounds have been proved in [4], the case $p=\infty$, and in [5], $p<\infty$ (see Section 1).

For the sake of simplicity we consider only the case of Gaussian noise. Namely, we suppose that the observations

$$Y_i = f(x_i) + \xi_i, \qquad i = 1, \ldots, n,$$

and ξ_i are iid Gaussian random variables with $\mathbf{E}\xi_i = 0$, $\xi_i^2 = 1$. The general case can be treated as in [4].

We choose the following plan of observations. Take integers N, n, $r = [n/N]$. Pick N knots $x_{kN} \in [-1,1]$ and at any knot, except maybe one, make r observations. The number of observations, at the exceptional knot is $n - r(N-1)$. For the sake of simplicity we suppose below that $n = rN$. Let

$$\overline{Y}_{kN} = f(x_{kN}) + r^{-1} \sum \xi_i = f(x_{kN}) + \eta_{kn}$$

be the arithmetic means of observations at the point x_{kN}. Evidently, η_{kn} are Gaussian with means zero and variances equal to r^{-1}. We write below $\overline{Y}_k$, x_k, etc. instead of $\overline{Y}_{kN}$, etc.

Take as an estimator for $f(x)$ the statistic

$$f_N(x) = \sum \overline{Y}_k l_k(x),$$

where $l_k(x) = l_{kN}(x)$ are Lagrange interpolation polynomials and

$$l_{kN}(x_{jN}) = l_k(x_j) = \delta_{kj}.$$

We take for the knots x_{jN} the zeros of the Chebyshev polynomial $T_N(x)$ of order N (see [7])

$$x_k = x_{kN} = \cos((2k-1)\pi(2N)^{-1}).$$

Then

$$f_N(x) - f(x) = \Big[\sum f(x_k) l_{kN}(x) - f(x)\Big] + \xi_N(x),$$

where the random polynomial

$$\xi_N(x) = \sum \eta_k l_k(x).$$

The rate of approximation of analytic functions $f \in A(G)$ by the interpolation polynomials $\sum f(x_k) l_k(x)$

$$\Big| f - \sum f(x_k) l_k \Big|_\infty \leqq cMe^{-\gamma N},$$

where $\gamma > 0$ depends on the region G (see [10]). Hence as before

$$\mathbf{E}|f - f_N|_p \leqq cMe^{-\gamma N} + \mathbf{E}|\xi_N|_p. \tag{4.1}$$

Further

$$\mathbf{E}|\xi_N|_p^p = (2\pi)^{1/2} 2^{(p+1)/2} r^{-p/2} \Gamma((p+1)/2) \int_{-1}^{1} \Big(\sum l_k^2(x)\Big)^{p/2} dx.$$

Prove that

$$\sum l_k^2(x) \leqq 2.$$

Let

$$h_k(x) = \left(1 - \frac{T_N''(x_k)}{T_N'(x_k)}(x - x_k)\right) l_k^2(x) = v_k(x) l_k^2(x)$$

be the Hermite interpolation polynomials. Then

$$\sum h_k(x) = \sum v_k(x) l_k^2(x) = 1$$

(see [7], Section 14.1). The Chebyshev polynomials T_N satisfy the equation (see [7])

$$(1-x^2)T_N''(x) - xT_N'(x) + N^2 T_N = 0.$$

Thus $T_N''(x_k)(T_N'(x_k))^{-1} = x_k(1-x_k^2)^{-1}$. Hence

$$v_k(1) = (1+x_k^2)^{-1} \geqq 1/2, \qquad v_k(-1) \geqq 1,$$

and the functions $v_k(x) \geqq 1/2$. It follows that

$$1 = \sum v_k(x) l_k^2(x) \geqq \frac{1}{2} \sum l_k^2(x).$$

We have for $p < \infty$

$$\mathbf{E}|\xi_N|_p \leqq cp^{1/2} r^{-1/2} = c(pN)^{1/2} n^{-1/2}. \tag{4.2}$$

Take here $N \sim \ln n/\gamma$. We find from (4.2) that

$$\Delta_p(n) \leqq c\sqrt{\frac{\ln n}{n}}.$$

If $p = \infty$, we apply (2.10). Then

$$\mathbf{E}|\xi_N|_\infty \leqq (p+1)^{1/p} N^{2/p} \mathbf{E}|\xi_N|_p \leqq cN^{2/p}\sqrt{p}\sqrt{N/n}.$$

Take here $N \sim \ln n/\gamma$, $p \sim \ln N$. We find

$$\Delta_\infty \leqq cn^{-1/2}(\ln n)^{1/2}(\ln \ln n)^{1/2}.$$

The proof is completed.

5. Analytic functions of several variables

Denote $\mathbf{F}_d(G, M)$ the collection of functions $f(x_1, \dots, x_d)$ defined and analytic on the closed d-dimensional cube $J = \{\mathbf{x} = (x_1, \dots, x_d) : |x_j| \leqq 1\}$. We suppose that all functions from $\mathbf{F}_d$ have analytic continuation into a region $G \supset J$ of the complex space of d complex variables $\mathbf{z} = (z_1, \dots, z_d)$, $z_k = x_k + iy_k$, and are bounded there by a constant M. We suppose that $G \subset E_1 \times E_2 \times \cdots \times E_d$, where E_k are ellipses of the complex plane $z_k = x_k + iy_k$ with foci at the points ± 1 of the real axis.

Consider multidimensional generalizations of our initial problems I–III.

Problem 5.I. An observed signal $X_\varepsilon(t)$ is of the form

$$dX_\varepsilon(t) = f(t)dt + \varepsilon dw(t),$$

where $f(t) = f(t; x_1, \ldots, x_{d-1})$ and $w(t)$ is a cylindrical Wiener process (see details in [12]). The signal $f \in \mathbf{F}_d(G, M)$ and $\Delta_p(\varepsilon)$ are defined as above

$$\Delta_p(\varepsilon) = \inf \sup \mathbf{E}_f |\hat{f} - f|_p$$

and the upper bound is taken over all $f \in \mathbf{F}_d$.

THEOREM 5.1. *The expressions $\Delta_p(\varepsilon)$ satisfy the following asymptotic relations*

$$\begin{aligned} &\Delta_p(\varepsilon) \asymp \varepsilon(\ln(1/\varepsilon))^{d/2}, \qquad 2 \leqq p < 4, \\ &\Delta_4(\varepsilon) \asymp \varepsilon(\ln(1/\varepsilon))^{d/2}(\ln\ln(1/\varepsilon))^{d/4}, \\ &\Delta_p(\varepsilon) \asymp \varepsilon(\ln(1/\varepsilon))^{d(1-2/p)}, \qquad p > 4. \end{aligned} \tag{5.1}$$

PROBLEM 5.II. Assume that one observes iid random d-dimensional random vectors $X_1, \ldots, X_d$ taking values in J and having density function $f(\mathbf{x})$. Again

$$\Delta_p(n) = \inf \sup \mathbf{E}_f |\hat{f} - f|_p$$

and the upper bound is taken over all densities $f \in \mathbf{F}_d(G, M)$.

THEOREM 5.2. *The expressions $\Delta_p(n)$ satisfy the following asymptotic relations*

$$\begin{aligned} &\Delta_p \asymp n^{-1/2}(\ln n)^{d/2}, \qquad 1 \leqq p < 4, \\ &\Delta_4(n) \asymp n^{-1/2}(\ln n)^{d/2}(\ln\ln n)^{d/4}, \\ &\Delta_p(n) \asymp n^{-1/2}(\ln n)^{d(1-2/p)}, \qquad p > 4. \end{aligned} \tag{5.2}$$

PROBLEM 5.III. Assume that as in Problem III one observes

$$Y_j = f(X_j) + G_j(X_j, \omega),$$

where now the points $X_j \in J$ and the unknown regression function $f \in \mathbf{F}_d(G, M)$. Let $\Delta_p(n)$ be defined analogously to (1.8).

THEOREM 5.3. *The expressions $\Delta_p(n)$ satisfy the following asymptotic relations*

$$\begin{aligned} &\Delta_p(n) \asymp n^{-1/2}(\ln n)^{d/2}, \qquad 1 \leqq p < \infty, \\ &\Delta_\infty(n) \asymp n^{-1/2}(\ln n \cdot \ln\ln n)^{d/2}. \end{aligned} \tag{5.3}$$

The proof of these theorems has no new moments with respect to the case $d = 1$. We expand the function f into the series

$$f(\mathbf{x}) = \sum a_{\mathbf{n}} P_{\mathbf{n}}(\mathbf{x}),$$

where the multiindex $\mathbf{n} = (n_1, \ldots, n_d)$ and the polynomials

$$P_{\mathbf{n}}(\mathbf{x}) = P_{n_1}(x_1) \ldots P_{n_d}(x_d)$$

and $P_{n_k}(x_k)$ are Legendre polynomials on $[-1, 1]$ and follow the arguments of Section 2.4. We omit these arguments (see also [3]).

REFERENCES

[1] Ibragimov, I. and Hasminskii, R., *Statistical estimation: asymptotic theory*, Applications of Mathematics, 16, Springer-Verlag, New York – Berlin, 1981. *MR* **82g**:62006

[2] Ibragimov, I. and Hasminskii, R., Estimation of infinite dimensional parameter in a Gaussian white noise, *Dokl. Akad. Nauk SSSR* **252** (1976), 1059–1063; English translation: *Soviet Math. Dokl.* **18** (1977), 1307–1309.

[3] Hasminskii, R. and Ibragimov, I., On density estimation in the view of Kolmogorov's ideas in approximation theory, *Ann. Statist.* **18** (1990), 999–1010. *MR* **92b**:62049

[4] Ibragimov, I. and Hasminskii, R., Bounds for the quality of nonparametric estimation of regression, *Teor. Verojatnost. i Primenen.* **27** (1982), 81–94 (in Russian). *MR* **83e**:62052. Translation: Bounds for the risks of nonparametric regression estimates, *Probab. Theory Appl.* **27** (1982), 84–99.

[5] Ibragimov, I. and Hasminskii, R., Asymptotic bounds for the quality of nonparametric estimation of regression in L_p, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **97** (1980), 88–101 (in Russian). *MR* **83b**:62073

[6] Bernstein, S. N., *Extremal properties of polynomials*, ONTI, Moscow, 1937 (in Russian).

[7] Szegő, G., *Orthogonal polynomials*, American Mathematical Society Colloquium Publications, Vol. 23, Revised ed., American Mathematical Society, Providence, RI, 1959. *MR* **21** #5029

[8] Timan, A., *Theory of approximation of functions of a real variable*, International Series of Monographs in Pure and Applied Mathematics, Vol. 34, A Pergamon Press book, The Macmillan Co., New York, 1963. *MR* **33** #465

[9] Gallager, R., *Information theory and reliable communications*, J. Wiley, New York, 1968.

[10] Babenko, K., *Foundations of numerical analysis*, Nauka, Moscow, 1986 (in Russian). *MR* **88g**:65001

[11] Petrov, V., *Limit theorems of probability theory*, Oxford Studies in Probability, 4, The Clarendon Press. Oxford University Press, New York, 1995. *MR* **96h**:60048

[12] Ibragimov, I. and Khasminskii, R., Some estimation problem in infinity dimensional Gaussian white noise, Festschrift for L. Le Cam (Pollard D., Yang Grace, Torgersen E., eds.), Springer, 1997.

(Received August 26, 1997)

PETERSBURG BRANCH OF
STEKLOV MATHEMATICAL INSTITUTE OF
RUSSIAN ACADEMY OF SCIENCES
FONTANKA 27
191 011 ST. PETERSBURG
RUSSIA

ibr32@pdmi.ras.ru

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 211–218

ON THE MULTIPLICATIVE FUNCTION $n^{i\tau}$

I. KÁTAI[1] and M. V. SUBBARAO

Dedicated to my teacher and friend

1

Let $\overline{\mathcal{M}}$ be the set of those completely multiplicative functions f for which $|f(n)|=1$ $(n\in\mathbf{N})$. Let S_k be the group of kth roots of unity, T be the set of complex numbers z with $|z|=1$.

In our recent paper [1] we formulated the following

CONJECTURE 1. *Let $\mathcal{A}=\{\alpha_1,\alpha_2,\ldots,\alpha_k\}$ be the set of the limit points of the sequence $\{f(n+1)\overline{f}(n)\mid n\in\mathbf{N}\}$. Then $\mathcal{A}=S_k$ and $f(n)=n^{i\tau}F(n)$, where τ is a suitable real number, $F^k(n)=1$ $(n\in\mathbf{N})$, and for each $\omega\in\mathcal{A}$ there exists a suitable subsequence n_ν such that $F(n_\nu+1)\overline{F}(n_\nu)=\omega$.*

For $k=1$ this assertion can be deduced simply from the theorem of Wirsing (see [2]) asserting that $f(n+1)\overline{f}(n)\to 1$ $(n\in\mathbf{N})$, $f\in\overline{\mathcal{M}}$ implies that $f(n)=n^{i\tau}$. In [1] we proved this conjecture for $k\leqq 3$.

The purpose of this short paper is to analyze the case $k=4$.

THEOREM. *Let $\mathcal{A}=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$, $f\in\overline{\mathcal{M}}$ be such a function for which the set of limit points of $\{f(n+1)\overline{f}(n)\mid n\in\mathbf{N}\}$ is $\mathcal{A}$. Then there is some $\tau\in\mathbf{R}$ such that $f(n)=n^{i\tau}F(n)$, and either* (A) *or* (B) *hold.*

(A) $\mathcal{A}=S_4$, $F(n)\in S_4$ $(n\in\mathbf{N})$.

(B) *$\mathcal{A}$ consists of four distinct elements of S_5, i.e., $\mathcal{A}=\{\mathcal{K}^{l_1},\mathcal{K}^{l_2},\mathcal{K}^{l_3},\mathcal{K}^{l_4}\}$ ($\mathcal{K}$ is a primitive fifth root of unity), $F(n)\in S_5$ and $F(n+1)\overline{F}(n)\in\mathcal{A}$ for every large n.*

REMARK. We think that the case (B) cannot hold, especially that if $F\in\overline{\mathcal{M}}$, $F(n)\in S_5$, $F(n)\not\equiv 1$, then for each $\alpha\in S_5$, $F(n+1)\overline{F}(n)=\alpha$ occurs infinitely often.

1991 *Mathematics Subject Classification*. Primary 11K65; Secondary 11N64.
Key words and phrases. Multiplicative functions, characterization.

[1] This paper was written during the first author's stay at the University of Alberta in 1997 as a visiting Research Professor, fully funded by the second author's NSERC Grant.

2

Let $f \in \overline{\mathcal{M}}$, $\mathcal{A} = \mathcal{A}_f = \{\alpha_1, \ldots, \alpha_k\}$ the set of limit points of $f(n+1)\overline{f}(n)$. Let $C(n)$ denote that element of $\mathcal{A}_f$ which is closest to $f(n+1)\overline{f}(n)$. Since $\mathcal{A}_f$ is a finite set, therefore $C(n)$ is uniquely determined for all large n. Since

$$f(n+1)\overline{f}(n) = \prod_{j=0}^{d-1} f(dn+(j+1))\overline{f}(dn+j),$$

therefore

$$C(n) = \prod_{j=0}^{d-1} C(dn+j), \tag{2.1}$$

valid for all $n > N_1(d)$, where $N_1(d)$ is a constant that may depend on d and f.

Furthermore, since

$$f(n^2)\overline{f}(n^2-1) = f(n)\overline{f}(n-1)f(n)\overline{f}(n+1),$$

we obtain that

$$C(n-1) = C(n)C(n^2-1) \tag{2.2}$$

whenever $n > N_2$, where N_2 is a constant.

From (2.2) we obtain that if $(C(m), C(m+1)) = (\beta, \gamma)$ occurs and $m > N_2$, then $\beta\overline{\gamma} \in \mathcal{A}$.

Similarly, if $(C(2m), C(2m+1)) = (\beta, \gamma)$ occurs for at least one $m > N_1(2)$, then $\beta\gamma \in \mathcal{A}$.

3. Lemmata

LEMMA 1. *There exists no such a $F \in \overline{\mathcal{M}}$ for which $F(\mathbf{N}) = S_6$ and either*

$$\mathcal{A}_F = \{1, (\alpha =)\mathcal{K}, \mathcal{K}^3(=-1), (\beta =)\mathcal{K}^5\} \tag{3.1}$$

or

$$\mathcal{A}_F = \{1, (\alpha =)\mathcal{K}, (\beta =)\mathcal{K}^2, \mathcal{K}^3(=-1)\}, \tag{3.2}$$

where $\mathcal{K}$ is a primitive sixth root of unity.

LEMMA 2. *Let $f, g \in \overline{\mathcal{M}}$, $f(n) = g(n)n^{i\tau}$ with some $\tau \in \mathbf{R}$. Then $\mathcal{A}_f = \mathcal{A}_g$.*

LEMMA 3. *If $f \in \overline{\mathcal{M}}$, $\mathcal{A}_f \subseteq S_k$ with some k, then there exists a $\tau \in \mathbf{R}$ such that $f(n) = n^{i\tau}F(n)$, $F(n)^k = 1$.*

Lemma 2 is obvious, since $(n+1)^{i\tau}n^{-i\tau} \to 1$ $(n \to \infty)$. Lemma 3 is a consequence of the fact that $\mathcal{A}_{f^k} = \{1\}$, and of Wirsing's theorem, which implies that $f(n) = n^{i\lambda}$, whence $(f(n)/n^{i\lambda})^k = (F(n))^k = 1$.

We shall prove Lemma 1 in Section 5.

4. Proof of the theorem

First we deduce the theorem by using Lemma 1, and after that we prove Lemma 1.

4.1. Assume that $-1, 1 \notin \mathcal{A}_f$

Let $C(n) = \xi$, $C(2n) = \eta$, $C(2n+1) = \tau$, $n > N_1(2), N_2$. Then $\xi = \eta\tau$ according to (2.1). Since $1 \notin \mathcal{A}$, therefore $\xi \neq \eta$, $\xi \neq \tau$, furthermore (2.2) implies that $\eta\overline{\tau} \in \mathcal{A}$, whence $\eta \neq \tau$. Thus each $\alpha_i \in \mathcal{A}$ can be written as the product of two distinct other elements of $\mathcal{A}$, $\alpha_i = \alpha_j\alpha_l$. Let $\alpha_1 = \alpha_2\alpha_3$. Then $\alpha_2\overline{\alpha}_3 \in \mathcal{A}$, $\alpha_2\overline{\alpha}_3 \neq \alpha_2$. Since $\alpha_2\overline{\alpha}_3 = \alpha_1$ would imply that $\alpha_3 = \overline{\alpha}_3$, i.e. that $\alpha_3 \in \{1, -1\}$, we obtain that $\alpha_2\overline{\alpha}_3 \neq \alpha_1$.

Then there are two cases:

Case A: $\alpha_2\overline{\alpha}_3 = \alpha_3$, i.e. $\alpha_2 = \alpha_3^2$,

Case B: $\alpha_2\overline{\alpha}_3 = \alpha_4$, i.e. $\alpha_2 = \alpha_3\alpha_4$.

4.1.1.1. Case A/1. Let $\alpha_2 = \alpha_u\alpha_v$, and assume that either $\alpha_u = \alpha_1$, or $\alpha_v = \alpha_1$.

4.1.1.1.1. Let $\alpha_u = \alpha_1$. Then $\alpha_v = \alpha_2\overline{\alpha}_1 = \overline{\alpha}_3 \in \mathcal{A}$. If $\overline{\alpha}_3 = \alpha_1$, then $\alpha_1 = \alpha_2\alpha_3 = \alpha_3^3$, whence $\alpha_3^4 = 1$, and so $\alpha_2 = \alpha_3^2 \in \{1, -1\}$. This is impossible.

Since $\alpha_v \neq \alpha_2$, and $\alpha_v = \alpha_3$ implies that $\alpha_3 \in \mathbf{R}$, we conclude that $\alpha_v = \alpha_4 = \overline{\alpha}_3$. Thus $\alpha_3^2 = \alpha_2 = \alpha_1\overline{\alpha}_3$, whence $\alpha_1 = \alpha_3^3$, $\alpha_4 = \overline{\alpha}_3$. Thus $\mathcal{A} = \{\alpha_1 = \alpha_3^3, \alpha_2 = \alpha_3^2, \alpha_3, \alpha_4 = \overline{\alpha}_3\}$. Let us write now α_3 as $\alpha_h\alpha_l$. We have the following possibilities

$$\text{(a)} \quad \alpha_3 = \alpha_1\alpha_2; \qquad \text{(b)} \quad \alpha_3 = \alpha_1\alpha_4; \qquad \text{(c)} \quad \alpha_3 = \alpha_2\alpha_4.$$

If (a) holds, then $\alpha_3 = \alpha_3^5$, $\alpha_3^4 = 1$, whence $\alpha_2 = \alpha_3^2 \in \{1, -1\}$, contrary to our assumption.

If (b) holds, then $\alpha_3 = \alpha_3^2$, whence $\alpha_3 = 1$.

It remains to consider the case (c). So we are in the case when:

$$\alpha_1 = \alpha_2\alpha_3, \quad \alpha_2 = \alpha_1\alpha_4, \quad \alpha_3 = \alpha_2\alpha_4,$$
$$\mathcal{A} = \{\alpha_1 = \alpha_3^3, \alpha_2 = \alpha_3^2, \alpha_3, \alpha_4 = \overline{\alpha}_3\}.$$

Now we can proceed as follows. If $\alpha_4 = \alpha_1\alpha_2$, then $\overline{\alpha}_3 = \alpha_3^5$, and so $\alpha_3^6 = 1$, consequently $\alpha_1 = \alpha_3^3 = \pm 1$. But $-1 \notin \mathcal{A}$ was assumed. If $\alpha_4 = \alpha_2\alpha_3$, then $\overline{\alpha}_3 = \alpha_3^3$, $1 = \alpha_3^4$ which leads to $\alpha_3^2 \in \{1, -1\}$.

There remains the case when $\alpha_4 = \alpha_1\alpha_3$, whence $\overline{\alpha}_3 = \alpha_3^4$, $\alpha_3^5 = 1$, i.e. $\mathcal{A} = \{\alpha_3^3, \alpha_3^2, \alpha_3, \alpha_3^4\}$, i.e. $\mathcal{A}$ is the set of the fifth primitive roots of unity. In this case $(f(n+1)\overline{f}(n))^5 \to 1$, and Wirsing's theorem asserts that $f^5(n) = n^{i\tau}$ $(n \in \mathbf{N})$, consequently $f(n) = n^{i\tau/5}F(n)$, where $F^5(n) = 1$ $(n \in \mathbf{N})$, furthermore $\mathcal{A}_f = \mathcal{A}_F$, consequently $F(n+1)\overline{F}(n) \neq 1$ if n is large enough.

Thus the theorem (with B) holds in this case.

4.1.1.1.2. Let $\alpha_v = \alpha_1$. This case can be treated as the previous one in 4.1.1.1.

4.1.1.2. Case A/2. Let $\alpha_u \neq \alpha_1$, $\alpha_v \neq \alpha_1$. Then $\alpha_2 = \alpha_3\alpha_4$. But then $\alpha_4 = \alpha_2\overline{\alpha}_3 = \alpha_3$, consequently $\text{card}(\mathcal{A}_f) < 4$.

4.1.2. Case B. Then $\alpha_1 = \alpha_2\alpha_3$, $\alpha_2 = \alpha_3\alpha_4$. Let $\alpha_3 = \alpha_u\alpha_v$. We have to distinguish the following possibilities: (a) $\alpha_3 = \alpha_1\alpha_2$; (b) $\alpha_3 = \alpha_1\alpha_4$; (c) $\alpha_3 = \alpha_2\alpha_4$.

From (a) we obtain that $\alpha_1 = \alpha_1\alpha_2\alpha_2$, whence $\alpha_2 \in \{1, -1\}$. If (c) holds then $\alpha_2 = \alpha_3\alpha_4 = \alpha_2\alpha_4^2$, i.e. $\alpha_4 \in \{1, -1\}$.

Assume now that (b) holds. Then $\alpha_3 = (\alpha_2\alpha_3)\alpha_4$, whence $\alpha_2\alpha_4 = 1$, $\alpha_4 = \overline{\alpha}_2$, and so $\alpha_2 = \alpha_3\overline{\alpha}_2$, $\alpha_3 = \alpha_2^2$, $\alpha_1 = \alpha_2^3$. Thus $\mathcal{A} = \{\alpha_2, \alpha_2^3, \alpha_2^2, \overline{\alpha}_2\}$.

Then we should discuss the following cases:

(1) $\alpha_4 = \alpha_1\alpha_2$ which leads to $\overline{\alpha}_2 = \alpha_2^3\alpha_2$, $\alpha_2^5 = 1$,

(2) $\alpha_4 = \alpha_1\alpha_3$ from which $\overline{\alpha}_2 = \alpha_2^{3+2}$, i.e. $\alpha_2^6 = 1$, whence $\alpha_1 = \alpha_2^3 \in \{1, -1\}$, which is impossible.

(3) $\alpha_4 = \alpha_2\alpha_3$, consequently $\overline{\alpha}_2 = \alpha_2\alpha_2^2$, i.e. $\alpha_2^4 = 1$, $\alpha_3 = \alpha_2^2 \in \{1, -1\}$.

Thus (2), (3) do not occur. In the case (1) we deduced that $\mathcal{A}$ is the set of the fifth primitive roots of unity, i.e. $\mathcal{A} = S_5 \setminus \{1\}$. Hence the theorem immediately follows.

4.2. Assume that $\mathcal{A}_f = \{-1, 1, \alpha, \beta\}$, $\{\alpha, \beta\} \neq \{i, -i\}$.

Then $\mathcal{A}_{f^2} = \{1, \alpha^2, \beta^2\}$. If $\alpha^2 = \beta^2$, then from our Theorem in [1], for the case card $\mathcal{A}_{f^2} = 2$ we obtain that $\alpha^2 = \beta^2 = -1$, i.e. that $\{\alpha, \beta\} = \{i, -i\}$.

Assume that card $\mathcal{A}_{f^2} = 3$. Then from our theorem in [1] for $k = 3$ we obtain that $\alpha^2 = \omega$, $\beta^2 = \overline{\omega}$, where $\omega^3 = 1$, $\omega \neq 1$.

If there is an $n > \max(N_1(2), N_3)$ for which $C(n) = 1$, $C(n+1) = \alpha$ (or β), then $\overline{\alpha}$ (or $\overline{\beta}$) belongs to $\mathcal{A}_f$, consequently $\beta = \overline{\alpha}$. In this case there exists no such an $m > N_2$ for which $C(m) = -1$, $C(m+1) \in \{\alpha, \beta\}$, since it would imply that $-\overline{\alpha} = \beta$, and this is clearly impossible.

Similarly, if $C(m) = -1$, $C(m+1) = \alpha$ (or β) is realizable for some $m > N_2$, then $-\overline{\alpha} = \beta$.

We obtain immediately that with some primitive sixth root of unity $\mathcal{K}$, either $\mathcal{A}_f = \{1, \mathcal{K}(=\alpha), \mathcal{K}^3(=-1), \mathcal{K}^5(=\beta=\overline{\alpha})\}$, or $\mathcal{A}_f = \{1, \mathcal{K}(=\alpha), \mathcal{K}^2(=\beta=-\overline{\alpha}), \mathcal{K}^{-3} = -1\}$. Since $\mathcal{A}_f \subseteq S_6$, therefore by Lemma 3 we have that $f(n) = n^{i\tau}F(n)$ $(n \in \mathbf{N})$, where $F^6(n) = 1$. Since $\mathcal{A}_f = \mathcal{A}_F$, from Lemma 1 we get that this is impossible.

4.3. The Case $\mathcal{A} = \{1, \alpha, \beta, \gamma\}$, $-1 \notin \mathcal{A}$.

Let α be such an element among α, β, γ for which $C(n) = 1$, $C(n+1) = \alpha$ occurs infinitely often. Then $C(n)\overline{C(n+1)} = \overline{\alpha} \in \mathcal{A}$. Let $\beta = \overline{\alpha}$. Since $\alpha \neq \pm 1$, therefore $\overline{\alpha} \notin \{1, \alpha\}$. If there would be a sequence n_ν such that $C(n_\nu + 1) = \gamma$, $C(n_\nu) = 1$, then it would imply that $\overline{\gamma} \in \mathcal{A}$, which is impossible.

Consequently either (α,γ) or (β,γ) occurs as $(C(n),C(n+1))$, infinitely often. In the first case (1) $\alpha\overline{\gamma}\in\mathcal{A}$, in the second (2) $\beta\overline{\gamma}\in\mathcal{A}$.

Assume that (1) holds. Then $\alpha\overline{\gamma}\neq 1,\alpha$, consequently either (a) $\alpha\overline{\gamma}=\beta$, or (b) $\alpha\overline{\gamma}=\gamma$.

If (1a) holds, then $\gamma=\alpha^2$, $\mathcal{A}=\{1,\alpha,\overline{\alpha},\alpha^2\}$.

If (1b) holds, then $\alpha=\gamma^2$, $\mathcal{A}=\{1,\gamma,\gamma^2,\overline{\gamma}^2\}$.

The case (2) can be treated by changing the values α,β.

Case 1a. Let us observe that if one of $\overline{\alpha}^2,\alpha^3,\overline{\alpha}^3,\alpha^4$ belongs to $\mathcal{A}$, then either $\mathcal{A}\subseteq S_5$, or $\mathcal{A}\subseteq S_3$ or $-1\in\mathcal{A}$. In the first case our theorem (with assertion B) holds, the two other conclusions are contradictory. So we may exclude these cases.

Let ξ,η be such a pair of elements of $\mathcal{A}$ for which $(C(2n),C(2n+1))=(\xi,\eta)$ holds for at least one $n>N_3$. We observe that if $\gamma\in\{\xi,\eta\}$, then $\xi=\gamma$ and $\eta=1$, consequently $C(n)=\gamma$. Indeed, $(\xi,\eta)\neq(1,\gamma)$, since $\overline{\gamma}=\overline{\alpha}^2\notin\mathcal{A}$, $(\xi,\eta)\neq(\alpha,\gamma)$, since $\alpha\gamma=\alpha^3\notin\mathcal{A}$, $(\xi,\eta)\neq(\beta,\gamma)$, since $\beta\overline{\gamma}=\overline{\alpha}^3\notin\mathcal{A}$, $(\xi,\eta)\neq(\gamma,\gamma)$, since $\gamma^2=\alpha^4\notin\mathcal{A}$, furthermore $(\xi,\eta)\neq(\gamma,\alpha)$ since $\gamma\alpha=\alpha^3\notin\mathcal{A}$, $(\xi,\eta)\neq(\gamma,\beta)$ since $\gamma\overline{\beta}=\alpha^3\notin\mathcal{A}$. As a consequence we obtain that if $C(m)=\gamma$, $m>2N_3$, then $2\mid m$ and $C\left(\frac{m}{2}\right)=\gamma$.

Let us write each integer n as $n=2^{s(n)}A(n)$, where $A(n)$ is the highest odd divisor of n. From our previous observation follows that if $C(n)=\gamma$, then $A(n)\leqq N_3$.

From (2.2) we obtain that $C(n)=C(n+1)C(n(n+2))$. Let $n>4N_3$ be such an integer for which $C(n)=\gamma$. Then $4\mid n$, $C(n+1)=1$, consequently $C(n(n+2))=\gamma$. Since $2\|n+2$, therefore $A(n(n+2))\geqq\frac{n+2}{2}>2N_3$, and this is a contradiction.

Case 1b. Similarly as above we can exclude the cases when one of $\gamma^3,\overline{\gamma}$, belongs to $\mathcal{A}$. This implies that if $n>N_2$ and $(C(2n),C(2n+1))=(\xi,\eta)$, then $(\xi,\eta)\neq(\gamma,\alpha)$, (γ,β), $(1,\gamma)$, (α,γ), (β,η), consequently the possible pairs are $(1,1)$, $(\alpha,1)$, $(\beta,1)$, $(1,\alpha)$, $(1,\beta)$, (γ,γ) and consequently $C(n)=1$, $\alpha,\beta,\alpha,\beta,\gamma^2=\alpha$, which means that $C(n)=\gamma$ cannot occur if $n>N_2$.

4.4. Assume that $\mathcal{A}_f=\{-1,\alpha,\beta,\gamma\}$, $1\notin\mathcal{A}_f$.

Since $1\notin\mathcal{A}_f$, therefore $C(n)\neq C(n+1)$ holds for all large n. Then $(C(2n),C(2n+1))=(-1,\xi)$ or $(\xi,-1)$ holds for some $\xi\in\mathcal{A}_f$, $\xi\neq-1$. Let α be such an element for which it holds. Then $C(n)=-\xi=-\alpha$, $-\alpha\neq\alpha,-1$, consequently it is another element from $\mathcal{A}$, let $\gamma=-\alpha$.

Let now $C(n)=\beta=C(2n)C(2n+1)=\xi\eta$. If ξ or $\eta=-1$, then $-\beta\in\mathcal{A}$, but this is impossible. Clearly, $\xi,\eta\neq\beta$, thus either $\xi=\alpha$, $\eta=\gamma$ or $\xi=\gamma$, $\eta=\alpha$, whence $\beta=-\alpha^2$ follows. Thus $\mathcal{A}_f=\{-1,\alpha,-\alpha^2,-\alpha\}$. Hence $\mathcal{A}_{f^2}=\{1,\alpha^2,\alpha^4\}$, and by our theorem for $\operatorname{card}\mathcal{A}_{f^2}\leqq 3$, we obtain that either

$$\text{(a) } \operatorname{card}\mathcal{A}_{f^2}=1,\quad\text{or (b) } \operatorname{card}\mathcal{A}_{f^2}=2,\quad\text{or (c) } \operatorname{card}\mathcal{A}_{f^2}=3.$$

The cases (a) and (b) obviously cannot occur since they imply that $\alpha^2=1,-1$, and so that $\operatorname{card}\mathcal{A}_f\leqq 3$ or $1\in\mathcal{A}_f$.

It remains to consider the case (c). Then $\overline{\alpha}^2 = \omega$, $\alpha^4 = \overline{\omega}$, where ω is one of the third primitive root of unity.

Consequently $\mathcal{A}_f = \{\alpha = \mathcal{K}, \beta = \mathcal{K}^2, -1 = \mathcal{K}^3, \gamma = \mathcal{K}^4\}$, where $\mathcal{K}$ is a primitive sixth root of unity.

Let n_ν run over the sequence of integers for which $C(n_\nu) = \beta$. Let $C(2n_\nu) = \xi$, $C(2n_\nu + 1) = \eta$ for some large n_ν. Then $\beta = \xi\eta$, $\xi\overline{\eta} \in \mathcal{A}_f$. Since $-\beta \notin \mathcal{A}_f$, therefore $\xi \neq -1$, $\eta \neq -1$, consequently $\xi, \eta \in \{\alpha, \gamma\}$, $\xi \neq \eta$. Since $\alpha\gamma = \gamma\alpha = \mathcal{K}^5 \notin \mathcal{A}_f$, therefore β cannot occur in $\mathcal{A}_f$. This is a contradiction.

4.5. Assume that $\mathcal{A}_f = S_4$.

Then the theorem (Assertion A) immediately follows from Lemma 3. The proof of the theorem is completed.

5. Proof of Lemma 1

5.1. Assume that (3.1) holds.

We use the notation $C(n) = F(n+1)\overline{F}(n)$.

Assume that n is bigger than a suitable constant. Then the following assertions are true:

(1) If $C(n) = -1$, then $C(n-1)$, $C(n+1) \in \{-1, 1\}$. This is clear, since $-C(n-1)$, $-C(n+1) \in \mathcal{A}_f$, and $-\alpha, -\beta \notin \mathcal{A}_f$.

(2) If $C(n) = \alpha$ then $C(n+1) \neq \beta$; if $C(n) = \beta$, then $C(n+1) = \alpha$. Clear, since $\alpha\overline{\beta}, \beta\overline{\alpha} \notin \mathcal{A}_f$.

(3) If $C(2n) = \alpha$, then $C(2n+1) \neq \alpha$; if $C(2n) = \beta$, then $C(2n+1) \neq \beta$. Clear, $\alpha^2, \beta^2 \notin \mathcal{A}_f$.

(4) If $C(n) = \pm 1$, then $C(2n), C(2n+1) \in \{+1, -1\}$. From (4) we obtain that if M is large, and $C(M) \in \{1, -1\}$, then

$$C(2^l M + j) \in \{1, -1\} \text{ for } j = 0, \dots, 2^l - 1. \tag{5.1}$$

This is impossible. Let $G(n) := F(n)^2$. Then $G \in \overline{\mathcal{M}}$, $G^3(n) = 1$, and

$$G(n+1)\overline{G}(n) = 1 \tag{5.2}$$

if $n \in [2^l M, 2^l M + (2^l - 1)] = \mathcal{J}_l$; $l = 1, 2, \dots$. Hence we shall deduce that $G(n) = 1$ identically, which implies that $F(n) \in \{1, -1\}$ $(n \in \mathbf{N})$, consequently $\mathcal{A}_f \subseteq \{1, -1\}$, which contradicts the assumption. If $2^l > M^5$, say, then $\mathcal{J}_l$ contains a cube, $m^3 \in \mathcal{J}_l$. Then $G(m^3) = 1$, consequently $G(n) = 1$ for all $n \in \mathcal{J}_l$, if $2^l > M^5$. Let d be an arbitrary integer. If there is an m and an l with $2^l > M^5$ such that $dm^3 \in \mathcal{J}_l$, then $1 = F(dm^3) = F(d)$ follows. Let m_1 be the largest integer for which $dm_1^3 < 2^l M$. Then $d(m_1 + 1)^3 \leqq dm_1^3 + 2dm_1^2$, the right-hand side is less than

$$2^l M + 2d\left(\frac{2^l M}{d}\right)^{2/3} < 2^l M + 2^l,$$

if l is large enough. Consequently $d(m_1+1)^3 \in \mathcal{J}_l$.

5.2. Assume the fulfilment of (3.2).

By using (2.1) and (2.2), for all large values of n we obtain immediately:

(1) If $C(n)=\mathcal{K}$, then $C(n+1)\in\{1,\mathcal{K}\}$.

(2) If n is odd, then $C(n)\neq\mathcal{K}^2$.

(3) If $C(n)=1$, then $C(n+1)\in\{1,-1\}$.

(4) If $C(n)=-1$, then $C(n+1)\in\{1,-1\}$.

Starting with some large value n_0 such that $C(n_0)=1$, $C(n_0-1)\in\{\mathcal{K},\mathcal{K}^2\}$, we write the infinite sequence

$$C(n_0)C(n_0+1)\dots C(n_0+t)\dots$$

as $\mathcal{L}_1\mathcal{R}_1\mathcal{L}_2\mathcal{R}_2\dots$, where $\mathcal{L}_h$ is such a sequence in which only $1,-1$, and $\mathcal{R}_h$ is such a sequence in which only $\mathcal{K},\mathcal{K}^2$ occur. From (3) and (4) we have that in $\mathcal{L}_h$ the first element is 1, and the last element is -1.

(5) Let $\mathcal{L}_h=C(N_0)C(N_0+1)\dots C(N_0+s)$, $M=N_0+s+1$. Then M is even, since for odd M $C\left(\frac{M-1}{2}\right)=C(M-1)C(M)=-C(M)$, but $-C(M)\notin\mathcal{A}$ if $C(M)\in\{\mathcal{K},\mathcal{K}^2\}$.

(6) The first element $C(M)$ in $\mathcal{R}_h$ is $\mathcal{K}^2$. Clear, if $C(M)=\mathcal{K}$, then $C(M-1)=C(M)C(M^2-1)$, $C(M-1)=-1$, consequently $C(M^2-1)=\mathcal{K}^2$. Since M^2-1 is odd, it is impossible. Thus $C(M)=\mathcal{K}^2$.

(7) If $C(n)=\mathcal{K}^2$, then $C(n-1)=-1$. Since $n-1$=odd, therefore $C(n-1)\neq\mathcal{K}^2$, furthermore $C(n-1)\overline{C}(n)\in\mathcal{A}$ holds only if $C(n-1)=-1$. Hence we obtain that the first element of $\mathcal{R}_h$ is $\mathcal{K}^2$ and all the others, if any, are $\mathcal{K}$'s.

(8) If $C(n)=\mathcal{K}$, then $C(2n)=\mathcal{K}$ and $C(2n+1)=1$. Clearly, $\mathcal{K}=C(n)=C(2n)C(2n+1)$, $C(2n)\overline{C}(2n+1)\in\mathcal{A}$ is satisfied only if $C(2n)=\mathcal{K}$, $C(2n+1)=1$.

(9) If $C(n)=\mathcal{K}$, then $C(n+1)\neq\mathcal{K}$. Assume, to the contrary, that $C(n+1)=\mathcal{K}$. Then, by (8)

$$C(n)C(n+1)=C(2n)C(2n+1)C(2n+2)C(2n+3)=\mathcal{K}1\mathcal{K}1,$$

but this contradicts our observation that an $\mathcal{L}_j$ sequence always contains at least two elements. This finishes the proof.

REFERENCES

[1] KÁTAI, I. and SUBBARAO, M. V., The characterization of $n^{i\tau}$ as a multiplicative function, *Acta Math. Acad. Sci. Hungar.* (to appear).

[2] WIRSING, E., TANG, YUANSHENG and SHAO, PINTSUNG, On a conjecture of Kátai for additive functions, *J. Number Theory* **56** (1996), 391–395. *MR* **97a**:11158

(Received May 17, 1997)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
KOMPUTER ALGEBRA TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

katai@compalg.elte.hu

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
T6G 2G1
CANADA

m.v.subbarao@ualberta.ca

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 219–252

RANDOM WALKS CROSSING POWER LAW BOUNDARIES

H. KESTEN and R. A. MALLER

Dedicated to the memory of A. Rényi

Abstract

We collect together some known results, and prove some new results, giving criteria for $\limsup_{n\to\infty} |S_n|/n^\kappa = \infty$ a.s. or $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s., where S_n is a random walk and $\kappa \geqq 0$. Conditions which are necessary and sufficient are given for all cases, and the conditions are quite explicit in all but one case (the case $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$ for $\limsup_{n\to\infty} S_n/n^\kappa$). The results are related to the finiteness of the first passage times of the random walk out of the regions $\{(n,y) : n \geqq 1, |y| \leqq an^\kappa\}$ and $\{(n,y) : n \geqq 1, y \leqq an^\kappa\}$, where $\kappa > 0$, $a > 0$.

1. Introduction

There are many applications in sequential analysis, finance theory, and elsewhere, of results concerning the time it takes a random walk $S_n = \sum_1^n X_i$, with the increments X_i i.i.d., to escape from a region. Here we will be concerned with two very basic questions: when are the r.v.'s defined for $\kappa \geqq 0$ by

$$T_\kappa(a) = \min\left\{n \geqq 1 : \frac{|S_n|}{n^\kappa} > a\right\} \qquad (a > 0) \tag{1.1}$$

and

$$T_\kappa^*(a) = \min\left\{n \geqq 1 : \frac{S_n}{n^\kappa} > a\right\} \qquad (a > 0) \tag{1.2}$$

a.s. finite? (We take $T_\kappa(a) = \infty$ if $|S_n| \leqq an^\kappa$ for all $n \geqq 1$, and $T_\kappa^*(a) = \infty$ if $S_n \leqq an^\kappa$ for all $n \geqq 1$.) In applications areas where the properties of the exit time of a random walk from a region with curved boundaries are important, the a.s. finiteness of the exit time often follows from strong assumptions (e.g., the presence of a nonzero drift) which are not necessary and may obscure basic aspects of the problem. On the other hand, it is of course important to

1991 *Mathematics Subject Classification.* Primary 60G40, 60J15; Secondary 62L10, 60G50.

Key words and phrases. Random walks, first-passage times, exit times, boundary crossing probabilities, limsup behaviour.

know, alternatively, circumstances in which the exit time will *not* be finite a.s., so that a remedy such as a truncated test, or an estimate of the tail of the distribution of the exit time, can be constructed. Good references to procedures of this kind are in Siegmund ([22], Ch. IV) and Gut ([9], Theorem 5.1, p. 133), for example. In view of these applications there seems to be a need for a systematic study of the a.s. finiteness of the exit time. In the present paper we give such a study for exiting from power law boundaries, for which, as we will show, quite complete results can be obtained. We make no apriori assumptions (such as the existence of the mean) on the distribution of the X_i.

It is a tautology that $T_\kappa(a) < \infty$ a.s. if and only if $\max_{j\geqq 1}(|S_j|/j^\kappa) > a$ a.s., and similarly for $T^*_\kappa(a) < \infty$, but the distributions of the extended-value random variables $\max_{j\geqq 1}(|S_j|/j^\kappa)$ and $\max_{j\geqq 1}(S_j/j^\kappa)$ are not easy to deal with, in general. However, by the Hewitt–Savage 0–1 law, the random variables $\limsup_{n\to\infty}|S_n|/n^\kappa$ and $\limsup_{n\to\infty} S_n/n^\kappa$ are constants (possibly, ∞ or $-\infty$), a.s., and are correspondingly easier to handle. Thus we are led to investigate the relationship between the values of these random variables and the finiteness or otherwise of the passage times. In Theorems 1–3 we give necessary and sufficient conditions (some known, some newly derived) for $\limsup_{n\to\infty}|S_n|/n^\kappa = \infty$ a.s. and $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s.. This is best done by considering various cases corresponding to values of κ and the finiteness or otherwise of $\mathbf{E}|X|$ and of moments such as $\mathbf{E}|X|^{1/\kappa}$. Theorem 4 investigates when we can deduce $\limsup_{n\to\infty}|S_n|/n^\kappa = \infty$ a.s. from $\limsup_{n\to\infty}|S_n|/n^\kappa > 0$ a.s., and similarly for the one-sided case. Finally in Theorems 5–6 we relate these results back to $T_\kappa(a)$ and $T^*_\kappa(a)$.

Our first theorem deals with the 'two-sided' problem, which is easy to handle by means of the Marcinkiewicz–Zygmund law (Chow and Teicher ([1], p. 125). Throughout, we let X, X_i be i.i.d. r.v.s which are not degenerate at 0 and $S_n = \sum_1^n X_i$; also $\kappa \geqq 0$.

THEOREM 1. (a) *If* $\kappa \geqq 1$ *or if* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X = 0$, *then*

$$\limsup_{n\to\infty}|S_n|/n^\kappa = \infty \text{ a.s. if and only if } \mathbf{E}|X|^{1/\kappa} = \infty. \tag{1.3}$$

(b) *In all other cases, we have*

$$\limsup_{n\to\infty}|S_n|/n^\kappa = \infty \quad a.s.. \tag{1.4}$$

Next we look at one-sided case. These are not so simple. We will need the integrals

$$(1.5) \qquad A_-(x) = \int_0^x F(-y)dy \quad (x>0) \quad \text{and} \quad J_+ = \int_{[0,\infty)} \frac{xdF(x)}{A_-(x)},$$

where F is the c.d.f. of the X_i. Note that $0 \leqq A_-(x) \leqq \mathbf{E}X^-$, where

$$X^+ = \max(0, X) \quad \text{and} \quad X^- = X^+ - X$$

(and similarly for X_i^+ and X_i^-). We will only need J_+ when $F(0-) > 0$, in which case we let $A_-(x)/x$ have its limiting value, $F(0-)$, at 0.

In the next theorem, (a) and (b) are due to Chow and Zhang [2] and Erickson [5], respectively.

THEOREM 2. *Assume* $0 < F(0-) \leqq F(0) < 1$.

(a) *If* $\kappa > 1$, *then* $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ *a.s. if and only if*

$$(1.6) \qquad \int_{[0,\infty)} \min\left(x^{1/\kappa}, \frac{x}{A_-(x)}\right) dF(x) = \infty.$$

(b) *For* $\kappa = 1$:

$$(1.7) \qquad \limsup_{n\to\infty} S_n/n = \infty \quad \textit{a.s. if and only if} \quad J_+ = \infty.$$

(c) *If* $\frac{1}{2} < \kappa < 1$ *and* $\mathbf{E}|X| = \infty$, *then*

$$(1.8) \qquad \limsup_{n\to\infty} S_n/n^\kappa = \infty \quad \textit{a.s. if and only if} \quad J_+ = \infty.$$

(d) *If* $0 \leqq \kappa \leqq \frac{1}{2}$, *then*

$$(1.9) \qquad \begin{gathered} \limsup_{n\to\infty} S_n/n^\kappa = \infty \ \textit{a.s.} \\ \textit{if and only if } J_+ = \infty \textit{ or } 0 \leqq \mathbf{E}X \leqq \mathbf{E}|X| < \infty. \end{gathered}$$

(e) *If* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, *and* $\mathbf{E}X \neq 0$, *then*

$$(1.10) \qquad \limsup_{n\to\infty} S_n/n^\kappa = \infty \quad \textit{a.s. if and only if} \quad \mathbf{E}X > 0.$$

Now keep

$$(1.11) \qquad \frac{1}{2} < \kappa < 1, \quad \mathbf{E}|X| < \infty \quad \textit{and} \quad \mathbf{E}X = 0.$$

(f) *If* $\mathbf{E}(X^+)^{1/\kappa} = \infty$ *then*

$$\limsup_{n\to\infty} S_n/n^\kappa = \infty \quad a.s.. \tag{1.12}$$

(g) *If* $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ *a.s. then*

$$\mathbf{E}|X|^{1/\kappa} = \infty. \tag{1.13}$$

(h) *It is possible to have*

$$\mathbf{E}(X^+)^{1/\kappa} < \infty = \mathbf{E}(X^-)^{1/\kappa} \quad \textit{with} \quad \limsup_{n\to\infty} S_n/n^\kappa = \infty \ a.s.. \tag{1.14}$$

(i) *If* $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ *a.s. and* $\mathbf{E}(X^+)^{1/\kappa} < \infty$ *then*

$$\int_0^\infty \Big(\int_x^\infty F(-y)dy \Big)^{\kappa/(1-\kappa)} dx = \infty. \tag{1.15}$$

REMARKS. (i) If $F(0-) = 0$, then $\limsup S_n/n^\kappa = \limsup |S_n|/n^\kappa$ a.s. and Theorem 1 can be used. If $F(0) = 1$, then $\limsup S_n/n^\kappa = \infty$ cannot occur. Thus the assumption $0 < F(0-) \leqq F(0) < 1$ in Theorem 2 is not restrictive.

(ii) A necessary condition for (1.6) is $\mathbf{E}(X^+)^{1/\kappa} = \infty$ and $J_+ = \infty$. This is immediate from the definitions. Conversely, let $\tau_-(x) = \int_0^x y^{(1/\kappa)-1} F(-y)dy/\kappa$. Then, as $x \to \infty$, $\tau_-(x) \to \mathbf{E}(X^-)^{1/\kappa}$ which is in $(0, \infty]$ when $F(0-) > 0$. When $\kappa > 1$, $\tau_-(x) \geqq x^{(1/\kappa)-1} A_-(x)/\kappa$. Thus when $\kappa > 1$, a sufficient condition for (1.6) is

$$\int_{[0,\infty)} \frac{x^{1/\kappa} dF(x)}{\tau_-(x)} = \infty.$$

(iii) (1.15) is an improvement on (1.13) when $\mathbf{E}(X^+)^{1/\kappa} < \infty$, since it follows easily that $E(X^-)^{1/\kappa} = \infty$ when (1.15) holds and $\frac{1}{2} < \kappa < 1$. On the other hand, for a distribution function with $F(-x) = 1/(x^{1/\kappa} \log x)$ (for x large), $\mathbf{E}(X^-)^{1/\kappa} = \infty$, but (1.15) fails.

We next give a not very explicit condition which is necessary and sufficient for $\limsup_{n\to\infty} S_n/n^\kappa$ to be $+\infty$ a.s. which applies in the situation of (h) of Theorem 2.

THEOREM 3. *Suppose* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, *and* $\mathbf{E}(X^+)^{1/\kappa} < \infty$. *Let*

$$\Phi(\theta) = \mathbf{P}\{X > 0\} \exp[i\theta \mathbf{E}\{X \mid X > 0\}] + \mathbf{E}\{\exp[i\theta X]; X \leqq 0\}.$$

Then $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ *a.s. if and only if, for all* $\lambda > 0$,

(1.16)
$$\int_1^\infty \exp\left\{ -\lambda x^{\kappa/(1-\kappa)} \left(\int_0^{1/x} dz \int_{-\infty}^{\infty} \frac{1-\cos\theta}{\theta^2} \frac{d\theta}{1-e^{-z}\Phi(\theta)} \right)^{1/(1-\kappa)} \right\} \frac{dx}{x} = \infty.$$

REMARK. (iv) We do not know if it is possible to give a more explicit NASC than that in Theorem 3 for $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. when $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, and $\mathbf{E}(X^+)^{1/\kappa} < \infty$. Condition (1.16) of Theorem 3 is difficult to apply, and we present it mainly to suggest that a more explicit NASC is unlikely. For practical purposes, however, the sufficient condition for $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. supplied by (1.12) is probably useful enough. (1.15) is a simply-checked necessary but not sufficient condition (see Remark (v) below).

The following tables summarise the necessary and sufficient conditions (NASC) we have for $\limsup_{n\to\infty} |S_n|/n^\kappa = \infty$ a.s. and $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s..

Table 1

Value of κ	NASC for $\limsup_{n\to\infty} \lvert S_n\rvert/n^\kappa = \infty$ a.s.
$0 \leqq \kappa \leqq \frac{1}{2}$	Always true
$\frac{1}{2} < \kappa < 1$	$0 = \mathbf{E}X < \mathbf{E}\lvert X\rvert < \infty = \mathbf{E}\lvert X\rvert^{1/\kappa}$ or $0 < \lvert\mathbf{E}X\rvert \leqq \mathbf{E}\lvert X\rvert < \infty$ or $\mathbf{E}\lvert X\rvert = \infty$
$\kappa = 1$	$\mathbf{E}\lvert X\rvert = \infty$
$\kappa > 1$	$\mathbf{E}\lvert X\rvert^{1/\kappa} = \infty$

Table 2

Value of κ	NASC for $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s.
$0 \leqq \kappa \leqq \frac{1}{2}$	$J_+ = \infty$ or $0 \leqq \mathbf{E}X \leqq \mathbf{E}\lvert X\rvert < \infty$
$\frac{1}{2} < \kappa < 1$	When $\mathbf{E}X = 0$ and $\mathbf{E}\lvert X\rvert < \infty = \mathbf{E}(X^+)^{1/\kappa}$: Always true When $\mathbf{E}X = 0$ and $\mathbf{E}\lvert X\rvert \vee \mathbf{E}(X^+)^{1/\kappa} < \infty$: See Theorem 3 When $0 < \lvert\mathbf{E}X\rvert \leqq \mathbf{E}\lvert X\rvert < \infty$: $\mathbf{E}X > 0$ When $\mathbf{E}\lvert X\rvert = \infty$: $J_+ = \infty$
$\kappa = 1$	$J_+ = \infty$
$\kappa > 1$	$\int_{[0,\infty)} \min\left(x^{1/\kappa}, \frac{x}{A_-(x)}\right) dF(x) = \infty$

Next we consider the possible values of the lim sup in Theorems 1–3. Can it lie in $(0,\infty)$? One way of phrasing this is to ask:

(1.17) When does $\limsup_{n\to\infty} |S_n|/n^\kappa > 0$ a.s. imply $\limsup_{n\to\infty} |S_n|/n^\kappa = \infty$ a.s.?

(and similarly for the one-sided version). Case (b) of Theorem 1 shows that (1.17) need not be considered in the two-sided case unless $\kappa \geqq 1$ or $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X = 0$, and (1.17) is obviously not true in the case $\kappa = 1$ by the strong law of large numbers, when the limsup behaviour of $|S_n|/n$ is obvious. We consider the remaining cases as well as the one-sided situation in the next theorem.

THEOREM 4. (a) *If* $\kappa > 1$ *or if* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X = 0$ *then*

(1.18) $$\limsup_{n\to\infty} |S_n|/n^\kappa > 0 \text{ a.s. implies } \limsup_{n\to\infty} |S_n|/n^\kappa = \infty \text{ a.s.}.$$

(b) *Except when* $\kappa=1$, $\mathbf{E}|X|<\infty$ *and* $\mathbf{E}X\neq 0$, *or when* $\frac{1}{2}<\kappa<1$, $\mathbf{E}|X|<\infty$, *and* $\mathbf{E}X = 0$, *we have*

(1.19) $$\limsup_{n\to\infty} S_n/n^\kappa > 0 \text{ a.s. implies } \limsup_{n\to\infty} S_n/n^\kappa = \infty \text{ a.s.}.$$

However, (1.19) *is not true when* $\kappa = 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X \neq 0$, *or, in general, when* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, *and* $\mathbf{E}X = 0$.

REMARK. (v) The counterexample we use in the proof of Theorem 4 to show that (1.19) does not hold in general relies on a result of Klass [18], [19], which gives a condition for $\limsup_{n\to\infty} S_n/B_n \in (0,\infty)$ a.s. for a certain norming sequence B_n. See also Pruitt ([21], Theorem 7.5, p. 26). Klass ([20]) gives an example with $\limsup_{n\to\infty} S_n/B_n \in (0,\infty)$ a.s., but it is not clear that $B_n \sim n^\kappa$ for any $\kappa > 0$ in his example. Our counterexample also shows that (1.15) is not sufficient for $\limsup_{n\to\infty} S_n/n^\kappa = \infty$. (See the remark following the proof of Theorem 4.)

We next relate the size of $\limsup_{n\to\infty} S_n/n^\kappa$ to the finiteness of the passage times $T_\kappa(a)$ and $T_\kappa^*(a)$ defined in (1.1) and (1.2). We note that it is immediate from the definitions that for *fixed* $a \geqq 0$,

(1.20) $$\limsup_{n\to\infty} \frac{|S_n|}{n^\kappa} > a \text{ a.s. implies } T_\kappa(a) < \infty \text{ a.s.}$$

and

(1.21) $$\limsup_{n\to\infty} \frac{S_n}{n^\kappa} > a \text{ a.s. implies } T_\kappa^*(a) < \infty \text{ a.s.}.$$

In the opposite direction we have that

$$T_\kappa(a) < \infty \text{ for all } a \geqq 0 \text{ a.s. implies } \limsup_{n\to\infty} \frac{|S_n|}{n^\kappa} = \infty \text{ a.s.} \tag{1.22}$$

$$T_\kappa^*(a) < \infty \text{ for all } a \geqq 0 \text{ a.s. implies } \limsup_{n\to\infty} \frac{S_n}{n^\kappa} = \infty \text{ a.s.} \tag{1.23}$$

(1.23) follows from the fact that $\{T_\kappa^*(a) < \infty\} = \{\max_{j\geqq 1}(S_j/j^\kappa) > a\}$. This shows that $T_\kappa^*(a) < \infty$ for all $a \geqq 0$ implies $\max_{j\geqq 1}(S_j/j^\kappa) = \infty$, hence $\limsup_{n\to\infty} S_n/n^\kappa = \infty$. (1.22) can be proved similarly.

The next theorems discuss also when one can have $T_\kappa(a) < \infty$ a.s. for *some* $a \geqq 0$, but not for all $a \geqq 0$, and similar statements for $T_\kappa^*(a)$. We begin with $T_\kappa(a)$.

THEOREM 5. (a) *If* $0 \leqq \kappa \leqq \frac{1}{2}$, *or if* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X \neq 0$, *or if* $\frac{1}{2} < \kappa < 1$ *and* $\mathbf{E}|X| = \infty$, *or if* $\mathbf{E}|X|^{1/\kappa} = \infty$ *and either* $\kappa \geqq 1$ *or* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X = 0$, *then* $T_\kappa(a) < \infty$ *a.s. for all* $a > 0$.

(b) *Suppose* $\kappa = 1$ *and* $0 < |\mathbf{E}X| \leqq \mathbf{E}|X| < \infty$. *Then* $T_1(a) < \infty$ *a.s. for all* $a \leqq |\mathbf{E}X|$, *but* $\lim_{x\to\infty} \mathbf{P}\{T_1(x) = \infty\} = 1$; *in particular,* $T_1(x) = \infty$ *with positive probability for all large* x.

(c) *In all other cases (that is, when* $\mathbf{E}|X|^{1/\kappa} < \infty$ *and either* $\kappa > 1$ *or* $\frac{1}{2} < \kappa \leqq 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X = 0$*), we have* $\lim_{x\to\infty} \mathbf{P}\{T_\kappa(x) = \infty\} = 1$; *in particular,* $T_\kappa(x) = \infty$ *with positive probability for all large* x.

REMARK. (vi) There need not be a sharp demarcation in values of a for which $\mathbf{P}\{T_\kappa(a) = \infty\}$ is 0 or 1, in cases (b) and (c) of Theorem 5. When X is concentrated on $[x_0, x_0 + 1]$ for some $x_0 > 0$, then for $\kappa \geqq 1$ we have

$$x_0 \leqq X_1 \leqq \max_{j\geqq 1}(|S_j|/j^\kappa) \leqq x_0 + 1 \text{ a.s..}$$

Thus $T_\kappa(a) < \infty$ a.s. for $a < x_0$ and $T_\kappa(a) = \infty$ a.s. for $a > x_0 + 1$. For values of a between x_0 and $x_0 + 1$, $T_\kappa(a)$ takes the value ∞ with probability $\mathbf{P}\{\max_{j\geqq 1}(|S_j|/j^\kappa) \leqq a\}$, which lies in $(0, 1)$ for some a, if for example X is uniform on $[x_0, x_0 + 1]$. Note that $\mathbf{E}|X|^{1/\kappa} < \infty$ in this example, so we are in cases (b) and (c) of Theorem 5.

Now we consider the one-sided case.

THEOREM 6. *Assume* $0 < F(0-) \leqq F(0) < 1$.

(a) *When* $0 \leqq \kappa \leqq 1$, $T_\kappa^*(a) < \infty$ *a.s. implies* $\limsup_{n\to\infty} S_n/n^\kappa \geqq a$ *a.s. but when* $\kappa > 1$ *it does not, in general.*

(b) *When* $0 \leqq \kappa \leqq 1$ *we have*

(1.24) $T^*_\kappa(a) < \infty$ *a.s. for some* $a > 0$ *implies* $T^*_\kappa(a) < \infty$ *a.s. for all* $a > 0$,

except possibly in the cases $\kappa = 1$ *and* $\mathbf{E}|X| < \infty$, $\mathbf{E}X \neq 0$, *or* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X = 0$. *In these cases we may have* $\limsup_{n\to\infty} S_n/n^\kappa = a \in (0, \infty)$ *a.s., and if this occurs, then* $T^*_\kappa(x) < \infty$ *a.s. for all* $x < a$ *but*

$$\lim_{x\to\infty} \mathbf{P}\{T^*_\kappa(x) = \infty\} = 1. \tag{1.25}$$

(c) *When* $0 \leqq \kappa \leqq 1/2$, *we have*

$$T^*_\kappa(a) < \infty \quad \text{a.s. for some (and hence all)} \quad a > 0 \tag{1.26}$$

if and only if $J_+ = \infty$ *or* $0 \leqq \mathbf{E}X \leqq \mathbf{E}|X| < \infty$. *When* $1/2 < \kappa < 1$ *and* $\mathbf{E}|X| = \infty$, (1.26) *holds if and only if* $J_+ = \infty$. *When* $1/2 < \kappa < 1$, $\mathbf{E}|X| < \infty$ *and* $\mathbf{E}X \neq 0$, (1.26) *holds if and only if* $\mathbf{E}X > 0$.

(d) *Keep* $\kappa = 1$. *If* $\mathbf{E}|X| < \infty$, *we have* $T^*_1(a) < \infty$ *a.s. for some* $a > 0$ *if and only if* $\mathbf{E}X > 0$, *and in this case* $T^*_1(a) < \infty$ *a.s. for all* $a \leqq \mathbf{E}X$. *If* $\mathbf{E}|X| = \infty$, *then* $T^*_1(a) < \infty$ *a.s. for some* $a > 0$ *if and only if* $J_+ = \infty$, *and in this case* $T^*_1(a) < \infty$ *a.s. for all* $a > 0$.

REMARKS. (vii) Although the power law boundary n^κ in itself is important in applications, a useful generalisation would be to replace it in our results by a more general boundary, $g(n)$, say. For many of the cases we have considered this can be done straightforwardly, but others need more care. The boundary $\sqrt{n \log\log n}$, for example, is of interest when $\mathbf{E}X^2 < \infty$ with regard to the law of the iterated logarithm, and there are various generalisations of the law of the iterated logarithm for the case $\mathbf{E}X^2 = \infty$, too.

Prof. M. Klass (private communication) has suggested a proof of a generalisation of Theorem 2(f), which says that under (1.11), $\sum \mathbf{P}\{X > b_n\} = \infty$ implies $\limsup S_n/b_n = \infty$ a.s., for a certain class of nice sequences $\{b_n\}$ which contains all sequences $b_n = n^\kappa$ with $\frac{1}{2} < \kappa < 1$, and more. His proof relies on techniques developed in Klass [17]–[19], and also on recent work of Hahn and Klass [10]. Our quite different method of proof of (1.12) may be of use in other problems, too. We will not explore the issue of more general norming sequences further here, other than to mention that Chow and Zhang [2] allow fairly general norming sequences, as do Kesten and Maller [16].

(viii) A natural question, following our discussion of $T_\kappa(a)$ and $T^*_\kappa(a)$, is to relate the *last exit times*

$$L_\kappa(a) = \max\left\{n \geqq 1 : \frac{|S_n|}{n^\kappa} \leqq a\right\} \qquad (a > 0) \tag{1.27}$$

and

$$L^*_\kappa(a) = \max\left\{n \geqq 1 : \frac{S_n}{n^\kappa} \leqq a\right\} \qquad (a > 0) \tag{1.28}$$

to the liminf behaviour of S_n. The connection here is close, because for $a > 0$

$$\{L_\kappa(a) = \infty\} = \{|S_n| \leqq an^\kappa \text{ i.o.}\},$$

so

$$\liminf_{n\to\infty} \frac{|S_n|}{n^\kappa} < a \text{ a.s. implies } L_\kappa(a) = \infty \text{ a.s. implies } \liminf_{n\to\infty} \frac{|S_n|}{n^\kappa} \leqq a \text{ a.s.},$$

and similarly for $L^*_\kappa(a)$. As relevant results, we only mention here the 'other' law of the iterated logarithm of Chung [3] (recently generalised by Kesten [12]), and the example of Erickson ([6], Theorem 5).

2. Proofs

PROOF OF THEOREM 1. We will write the Marcinkiewicz–Zygmund strong law of large numbers in the following form: for $\kappa > 1/2$

$$\limsup_{n\to\infty} \frac{|S_n - cn|}{n^\kappa} < \infty \quad \text{a.s.} \tag{2.1}$$

for some finite c implies $\mathbf{E}|X|^{1/\kappa} < \infty$, and $\mathbf{E}|X|^{1/\kappa} < \infty$ implies

$$\lim_{n\to\infty} \frac{S_n - c'n}{n^\kappa} = 0 \quad \text{a.s.}, \tag{2.2}$$

where $c' = \mathbf{E}X$ if $\frac{1}{2} < \kappa \leqq 1$ and c' is arbitrary if $\kappa > 1$. (See, e.g., Chow and Teicher [1], p. 125 and their proof.)

Now (1.3) is immediate from (2.1)–(2.2) if $\kappa \geqq 1$, so keep $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X = 0$. Then $\limsup_{n\to\infty} |S_n|/n^\kappa = \infty$ a.s. implies $\mathbf{E}|X|^{1/\kappa} = \infty$ by (2.2), and the converse follows from (2.1). Thus (1.3) is proved.

Next take $0 \leqq \kappa \leqq \frac{1}{2}$. We will show that, always,

$$\limsup_{n\to\infty} \frac{|S_n|}{n^{1/2}} = \infty \quad \text{a.s.}. \tag{2.3}$$

Indeed, if this fails then $\limsup_{n\to\infty} |X_n|/n^{\frac{1}{2}} < \infty$ a.s. so by the Borel–Cantelli lemma, $\mathbf{E}X^2 < \infty$. If $\mathbf{E}X \neq 0$ then $|S_n|/n \to |\mathbf{E}X| > 0$ a.s. by the strong law of large numbers, so (2.3) holds. If $\mathbf{E}X = 0$ then for each $x > 0$

$$\limsup_{n\to\infty} \mathbf{P}\{|S_n| > xn^{1/2}\} > 0$$

by the central limit theorem, so (2.3) holds again by the Hewitt–Savage 0–1 law. This of course implies (1.4).

Now take $\frac{1}{2} < \kappa < 1$. If $\mathbf{E}|X| < \infty$ and $\mathbf{E}X \neq 0$, then $\limsup_{n\to\infty} |S_n|/n^\kappa = \limsup_{n\to\infty}(n^{1-\kappa}|S_n|/n) = \infty$ a.s. by the strong law of large numbers, so (1.4) holds. If $\mathbf{E}|X| = \infty$, then

$$\limsup \frac{|S_n|}{n} \geqq \frac{1}{2} \limsup \frac{|X_n|}{n} = \infty$$

by the Borel–Cantelli lemma, and (1.4) again follows. □

PROOF OF THEOREM 2. Assume $0 < F(0-) \leq F(0) < 1$.

(1.6) is immediate from Theorem 1 (i) of Chow and Zhang [2], and (1.7) is Theorem 2 (a) of Erickson [5].

Next we prove (1.8). Assume $\mathbf{E}|X| = \infty$. Then $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. implies $J_+ = \infty$ by (1.21) of Kesten and Maller [14]. Conversely suppose $\limsup_{n\to\infty} S_n/n^\kappa < \infty$ a.s. for some $\frac{1}{2} < \kappa < 1$. Then for some $c < \infty$

$$\frac{\sum_{i=1}^n X_i^+}{\sum_{i=1}^n X_i^-} \leqq \frac{cn^\kappa}{\sum_{i=1}^n X_i^-} + 1 \tag{2.4}$$

for all large n, a.s.. Now $\sum_{i=1}^n X_i^-/n^\kappa \to \infty$ a.s. when $0 < \kappa < 1$ as long as $\mathbf{E}X_1^- > 0$, so we obtain from (2.4) that $\limsup_{n\to\infty} \sum_{i=1}^n X_i^+ \Big/ \sum_{i=1}^n X_i^- < \infty$ a.s.. This implies $J_+ < \infty$ by Pruitt ([21, Lemma 8.1, p. 36) or Erickson ([5], Lemma 3).

With $\kappa = 0$, (1.9) is (1.21) of [14]; call this $(1.9)_0$. Then $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. for any $\kappa \geqq 0$ implies $(1.9)_0$, so the forward direction in (1.9) is immediate. Conversely, let $J_+ = \infty$ or $0 \leqq \mathbf{E}X \leqq \mathbf{E}|X| < \infty$. We will show then that

$$\limsup_{n\to\infty} \frac{S_n}{n^{1/2}} = \infty \quad \text{a.s.} \tag{2.5}$$

from which $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. for $0 \leqq \kappa \leqq 1/2$ follows. We now prove (2.5). If $J_+ = \infty$, then we know from (1.7) that $\limsup S_n/n = \infty$ a.s., so (2.5) holds. If $J_+ < \infty$ then by assumption $0 \leqq \mathbf{E}X \leqq \mathbf{E}|X| < \infty$. When $\mathbf{E}X > 0$, (2.5) is immediate from the strong law of large numbers. Thus we only have to consider the case $\mathbf{E}X = 0$. (2.5) then follows directly from the Theorem of Stone [24]. ((2.5) can also be easily obtained from a result of Klass–Pruitt ([21], Theorem 7.5, p. 26)). This completes the proof of (1.9).

(1.10) is of course trivial from the strong law of large numbers.

Now we prove (1.12). Assume that (1.11) holds. Define

$$\rho_\kappa(x) = \frac{1}{\kappa}\int_0^x y^{(1/\kappa)-1}(1-F(y))dy \tag{2.6}$$

and

$$\lambda(x) = \int_x^\infty (1-F(y))dy \tag{2.7}$$

(this is finite, because $\mathbf{E}|X| < \infty$).

LEMMA 2.1. *If* $0 < \kappa < 1$, $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, *and* $\mathbf{E}(X^+)^{1/\kappa} = \infty$, *then*

$$\int_{[1,\infty)} \frac{dF(x)}{\rho_\kappa(x)/x^{1/\kappa} + \lambda(x)/x} = \infty. \tag{2.8}$$

PROOF OF LEMMA 2.1. Assume that $\mathbf{E}(X^+)^{1/\kappa} = \infty$. We will suppose that (2.8) fails, so

$$\int_{[1,\infty)} \frac{dF(x)}{\rho_\kappa(x)/x^{1/\kappa} + \lambda(x)/x} < \infty. \tag{2.9}$$

Our first step is to show that this implies

$$\frac{x^{1/\kappa}(1-F(x))}{\rho_\kappa(x) + x^{(1/\kappa)-1}\lambda(x)} \to 0. \tag{2.10}$$

To this end we note that

$$\frac{\rho_\kappa(x)}{x^{1/\kappa}} = \frac{1}{\kappa x^{1/\kappa}}\int_0^x y^{(1/\kappa)-1}(1-F(y))dy = \frac{1}{\kappa}\int_0^1 y^{(1/\kappa)-1}(1-F(xy))dy$$

is decreasing in $x > 0$. Also $\lambda(x)$, and hence $\lambda(x)/x$, is decreasing in $x > 0$. Therefore, for any $z > 0$,

$$\frac{z^{1/\kappa}(1-F(z))}{\rho_\kappa(z) + z^{(1/\kappa)-1}\lambda(z)} \leq \int_{(z,\infty)} \frac{dF(x)}{\rho_\kappa(x)/x^{1/\kappa} + \lambda(x)/x}. \tag{2.11}$$

Thus (2.10) is an immediate consequence of (2.9).

The next step is to show that either

$$\frac{\lambda(x)/x}{\rho_\kappa(x)/x^{1/\kappa}} \to 0 \tag{2.12}$$

or

$$\liminf_{x\to\infty} \frac{\lambda(x)/x}{\rho_\kappa(x)/x^{1/\kappa}} > 0. \tag{2.13}$$

Suppose in fact that neither of these holds, so for some $\varepsilon > 0$

$$\limsup_{x\to\infty} \frac{\lambda(x)/x}{\rho_\kappa(x)/x^{1/\kappa}} \geqq \varepsilon \tag{2.14}$$

and

$$\liminf_{x\to\infty} \frac{\lambda(x)/x}{\rho_\kappa(x)/x^{1/\kappa}} = 0. \tag{2.15}$$

Our proof now is a minor modification of Proposition 3.1 of Griffin and McConnell ([8], p. 2029). Just as in their proof we can find sequences $s_k \geqq r_k \to \infty$ such that

$$\frac{\lambda(s_k)/s_k}{\rho_\kappa(s_k)/s_k^{1/\kappa}} \leqq \varepsilon/2, \tag{2.16}$$

$$\frac{\lambda(r_k)/r_k}{\rho_\kappa(r_k)/r_k^{1/\kappa}} \to \varepsilon, \tag{2.17}$$

$$\frac{\lambda(u)/u}{\rho_\kappa(u)/u^{1/\kappa}} \leqq \varepsilon \quad \text{for } r_k \leqq u \leqq s_k \tag{2.18}$$

and

$$\frac{s_k}{r_k} \to \gamma \in [1, \infty]. \tag{2.19}$$

Now fix $D > 1$ and let $u \in [r_k, Dr_k]$. Given $\eta > 0$, we can, by (2.10), choose k large enough for

$$r_k(1 - F(r_k)) \leqq (\eta\varepsilon/D^{1/\kappa})(r_k^{1-(1/\kappa)}\rho_\kappa(r_k) + \lambda(r_k)). \tag{2.20}$$

Then for k large enough

$$
\begin{aligned}
\lambda(u) &= \lambda(r_k) - \int_{r_k}^{u} (1 - F(y))dy \geqq \lambda(r_k) - u(1 - F(r_k)) \\
&\geqq \lambda(r_k) - Dr_k(1 - F(r_k)) \\
&\geqq \lambda(r_k) - \eta\varepsilon D^{1-(1/\kappa)}(r_k^{1-(1/\kappa)}\rho_\kappa(r_k) + \lambda(r_k)) \qquad \text{(by (2.20))} \\
&= (1 - \eta\varepsilon D^{1-(1/\kappa)})\lambda(r_k) - \eta\varepsilon D^{1-(1/\kappa)} r_k^{1-(1/\kappa)}\rho_\kappa(r_k) \\
&\geqq (1 - (\eta\varepsilon + 2\eta)D^{1-(1/\kappa)})\lambda(r_k) \qquad \text{(by (2.17))}.
\end{aligned}
\tag{2.21}
$$

Similarly

$$
\begin{aligned}
\rho_\kappa(u) &= \rho_\kappa(r_k) + \frac{1}{\kappa}\int_{r_k}^{u} y^{(1/\kappa)-1}(1 - F(y))dy \\
&\leqq \rho_\kappa(r_k) + (u^{1/\kappa} - r_k^{1/\kappa})(1 - F(r_k)) \\
&\leqq \rho_\kappa(r_k) + (D^{1/\kappa} - 1)r_k^{1/\kappa}(1 - F(r_k)) \\
&\leqq (1 + \eta\varepsilon)\rho_\kappa(r_k) + \eta\varepsilon r_k^{(1/\kappa)-1}\lambda(r_k).
\end{aligned}
\tag{2.22}
$$

Consequently, for $u \in [r_k, Dr_k]$, we have

$$
\begin{aligned}
\frac{\lambda(u)/u}{\rho_\kappa(u)/u^{1/\kappa}} &= u^{(1/\kappa)-1}\frac{\lambda(u)}{\rho_\kappa(u)} \geqq \frac{u^{(1/\kappa)-1}(1 - (\eta\varepsilon + 2\eta)D^{1-(1/\kappa)})\lambda(r_k)}{(1 + \eta\varepsilon)\rho_\kappa(r_k) + \eta\varepsilon r_k^{(1/\kappa)-1}\lambda(r_k)} \\
&= \left(\frac{u}{r_k}\right)^{(1/\kappa)-1} \frac{(1 - (\eta\varepsilon + 2\eta)D^{1-(1/\kappa)})\lambda(r_k)/r_k}{(1 + \eta\varepsilon)\rho_\kappa(r_k)/r_k^{1/\kappa} + \eta\varepsilon\lambda(r_k)/r_k}.
\end{aligned}
\tag{2.23}
$$

Suppose $\gamma < \infty$ in (2.19). Then take $D = \gamma + 1$ and $u = s_k$. Divide out $\rho_\kappa(r_k)/r_k^{1/\kappa}$ on the right-hand side of (2.23) and use (2.17) and (2.19) to get

$$
\begin{aligned}
\liminf_{k\to\infty} \frac{\lambda(s_k)/s_k}{\rho_\kappa(s_k)/s_k^{1/\kappa}} &\geqq \gamma^{(1/\kappa)-1}\frac{(1 - (\eta\varepsilon + 2\eta)D^{1-(1/\kappa)})\varepsilon}{(1 + \eta\varepsilon) + \eta\varepsilon^2} \\
&\to \gamma^{(1/\kappa)-1}\varepsilon \qquad (\text{as } \eta \downarrow 0).
\end{aligned}
\tag{2.24}
$$

Since $\gamma \geqq 1$ and $1/\kappa > 1$, this contradicts (2.16).

Next suppose $\gamma = \infty$ in (2.19). Then take $u = Dr_k$, $D > 1$, in (2.23) to get, as in (2.24),

$$
\begin{aligned}
\liminf_{k\to\infty} \frac{\lambda(Dr_k)/(Dr_k)}{\rho_\kappa(Dr_k)/(Dr_k)^{1/\kappa}} &\geqq D^{(1/\kappa)-1}\frac{(1 - (\eta\varepsilon + 2\eta)D^{1-(1/\kappa)})\varepsilon}{(1 + \eta\varepsilon) + \eta\varepsilon^2} \\
&\to D^{(1/\kappa)-1}\varepsilon \qquad (\text{as } \eta \downarrow 0).
\end{aligned}
\tag{2.25}
$$

On the other hand we may take $u = Dr_k$ in (2.18) to get

$$\limsup_k \frac{\lambda(Dr_k)/(Dr_k)}{\rho_\kappa(Dr_k)/(Dr_k)^{1/\kappa}} \leq \varepsilon. \tag{2.26}$$

Since $D > 1$ and $1/\kappa > 1$, (2.25) and (2.26) are contradictory. Thus indeed (2.12) or (2.13) holds.

It is now easy to deduce a contradiction from (2.9). If (2.12) holds then (2.9) implies that

$$\infty > \int_{[1,\infty)} \frac{x^{1/\kappa} dF(x)}{\int_{[0,x)} y^{1/\kappa} dF(y)}, \tag{2.27}$$

because $\rho_\kappa(x) + x^{(1/\kappa)-1}\lambda(x) \sim \rho_\kappa(x) \sim \int_{[0,x)} y^{1/\kappa} dF(y)$ under (2.10) and (2.12) (integrate by parts in (2.6)). Now (2.27) implies that $\mathbf{E}(X^+)^{1/\kappa} < \infty$, by the Abel–Dini Theorem, but we assumed $\mathbf{E}(X^+)^{1/\kappa} = \infty$. Alternatively, if (2.13) holds then by (2.9)

$$\infty > c \int_{[1,\infty)} \frac{x dF(x)}{\int_x^\infty (1 - F(y))dy} \geq c \int_{[1,\infty)} \frac{x dF(x)}{\int_{(x,\infty)} y dF(y)}$$

for some $c > 0$, which is also impossible by Abel–Dini. This proves Lemma 2.1. □

Now we can prove (1.12). We define

$$C_n = \inf\left\{ x > 0 : \frac{\rho_\kappa(x)}{x^{1/\kappa}} + \frac{\lambda(x)}{x} \leq \frac{\delta}{n} \right\}. \tag{2.28}$$

Then $C_n \uparrow \infty$ as $n \to \infty$ and

$$\frac{n\rho_\kappa(C_n)}{C_n^{1/\kappa}} + \frac{n\lambda(C_n)}{C_n} = \delta, \tag{2.29}$$

while

$$\frac{\rho_\kappa(x)}{x^{1/\kappa}} + \frac{\lambda(x)}{x} \geq \frac{\delta}{n} \quad \text{for } x \leq C_n.$$

It follows from Lemma 2.1 that for n_0 large enough

$$\infty = \int_{(C_{n_0},\infty)} \frac{dF(x)}{\rho_\kappa(x)/x^{1/\kappa} + \lambda(x)/x}$$

$$\leq \sum_{n>n_0} \int_{(C_{n-1},C_n]} \left(\frac{n}{\delta}\right) dF(x)$$
$$= \delta^{-1} \sum_{n>n_0} n((1-F(C_{n-1})) - (1-F(C_n)))$$
$$= \delta^{-1} \sum_{j>n_0+1} (1-F(C_{j-1})) + \delta^{-1}(n_0+1)(1-F(C_{n_0})).$$

This shows that $\sum_{n\geqq 1}(1-F(C_n)) = \infty$.

Next we will show that

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n}(X_i^+ - \mathbf{E}X_i^+)}{n^\kappa} = \infty \quad \text{a.s.}. \tag{2.30}$$

For this it will be convenient to let

$$\hat{X}_i = X_i^+ - \mathbf{E}X_i^+ \quad \text{and} \quad \hat{S}_n = \sum_{i=1}^{n} \hat{X}_i. \tag{2.31}$$

Fix $a \in (0, 1/2)$ and choose

$$0 < \delta < \min\left(\frac{a}{2}, \frac{a^2}{32c_+\kappa}\right), \tag{2.32}$$

where c_+ is a constant, depending only on the distribution of X^+, which we now specify. Let $Y_1^x = (\hat{X}_1 \wedge x) \vee (-x)$ and define $\hat{A}(x) = \mathbf{E}(Y_1^x)$. Then, for $x \geqq \mathbf{E}X^+$,

$$\begin{aligned}
\hat{A}(x) &= \int_0^x \Big(\mathbf{P}(\hat{X}_1 > y) - \mathbf{P}(\hat{X}_1 \leqq -y)\Big)\, dy \\
&= \int_{\mathbf{E}X^+}^{x+\mathbf{E}X^+} \mathbf{P}(X^+ > y)dy - \int_{\mathbf{E}X^+-x}^{\mathbf{E}X^+} \mathbf{P}(X^+ \leqq y)dy \\
&= \int_0^{x+\mathbf{E}X^+} \mathbf{P}(X^+ > y)dy - \int_0^{\mathbf{E}X^+} \mathbf{P}(X^+ > y)dy - \int_0^{\mathbf{E}X^+} \mathbf{P}(X^+ \leqq y)dy \\
&= \int_0^{x+\mathbf{E}X^+} (1-F(y))dy - \mathbf{E}X^+ = -\int_{x+\mathbf{E}X^+}^{\infty} (1-F(y))dy.
\end{aligned} \tag{2.33}$$

Thus we have

$$-\hat{A}(x) = \lambda(x + \mathbf{E}X^+) \leqq \lambda(x), \quad \text{for } x \geqq \mathbf{E}X^+. \tag{2.34}$$

Also, for $x \geqq \mathbf{E}X^+$,

$$\begin{aligned} \frac{1}{2}\mathbf{E}(Y_1^x)^2 &= \int_0^x y[\mathbf{P}(\hat{X} > y) + \mathbf{P}(\hat{X} \leqq -y)]dy \\ &\leqq \int_0^x y\mathbf{P}(X > y)dy + \int_0^{\mathbf{E}X^+} (\mathbf{E}X^+ - y)\mathbf{P}(X^+ \leqq y)dy \\ &\leqq \int_0^x y(1 - F(y))dy + (\mathbf{E}X^+)^2/2 \\ &\leqq 2c_+ \int_0^x y(1 - F(y))dy = c_+U_+(x), \quad \text{say,} \end{aligned} \tag{2.35}$$

for some $c_+ > 0$, where c_+ is a constant depending only on the distribution of X^+. This is the value we take in (2.32).

Take $n > m > 1$, and write

$$\begin{aligned} &\mathbf{P}\Big\{\max_{m \leq j \leq n} \frac{\hat{S}_j}{C_j} > a\Big\} \\ &\geqq \sum_{j=m}^{n} \mathbf{P}\{\hat{S}_{j-1} > -aC_j\}\mathbf{P}\Big\{\frac{\hat{X}_j}{C_j} > 2a \geqq \max_{j+1 \leq k \leq n} \frac{\hat{X}_k}{C_k}\Big\}. \end{aligned} \tag{2.36}$$

Let $Z_k^j = (\hat{X}_k \wedge C_j) \vee (-C_j) = Y_k^{C_j}$. Now for $C_j > \mathbf{E}X^+$

$$\begin{aligned} &\mathbf{P}\{\hat{S}_{j-1} \leqq -aC_j\} \\ &\leqq \mathbf{P}\Big\{\sum_{k=1}^{j-1}(\hat{X}_k \wedge C_j) \vee (-C_j) - (j-1)\hat{A}(C_j) \leqq -aC_j - (j-1)\hat{A}(C_j)\Big\} \\ &\qquad + (j-1)\mathbf{P}\{\hat{X}_j < -C_j\} \\ &\leqq \mathbf{P}\Big\{\sum_{k=1}^{j-1}(Z_k^j - \mathbf{E}Z_k^j) \leqq -\frac{1}{2}aC_j\Big\}, \end{aligned} \tag{2.37}$$

because by (2.34)

$$-(j-1)\hat{A}(C_j) \leqq j\lambda(C_j) \leqq \delta C_j \ \ (\text{by } (2.29)) \quad \leqq \frac{1}{2}aC_j \ \ (\text{by } (2.32)),$$

while $\mathbf{P}\{\hat{X}_j < -C_j\} = \mathbf{P}\{X_1^+ < \mathbf{E}X^+ - C_j\} = 0$ when $C_j > \mathbf{E}X^+$. Chebychev's inequality applied to (2.37) gives

$$\mathbf{P}\{\hat{S}_{j-1} \leqq -aC_j\} \leqq \frac{4j\mathbf{E}(Z_1^j)^2}{a^2C_j^2}. \tag{2.38}$$

Notice that, since $1/\kappa < 2$,

$$U_+(x) = 2\int_0^x y(1-F(y))dy$$
$$\leqq 2x^{2-1/\kappa}\int_0^x y^{(1/\kappa)-1}(1-F(y))dy = 2\kappa x^{2-(1/\kappa)}\rho_\kappa(x).$$

Using this, (2.35), and (2.38) in (2.37) gives

$$\mathbf{P}\{\hat{S}_{j-1} > -aC_j\} \geqq 1 - \frac{(8c_+)2\kappa j C_j^{2-(1/\kappa)}\rho_\kappa(C_j)}{a^2C_j^2} = 1 - \frac{16c_+\kappa j\rho_\kappa(C_j)}{a^2C_j^{1/\kappa}}$$
$$\geqq 1 - \frac{16c_+\kappa\delta}{a^2} \qquad \text{(by (2.29))}$$
$$\geqq \frac{1}{2} \qquad \text{(by choice of } \delta \text{ in (2.32))}.$$

Returning to (2.36) we see that, for large enough m,

$$\mathbf{P}\Big\{\max_{m\leqq j\leqq n}\frac{\hat{S}_j}{C_j} > a\Big\} \geqq \frac{1}{2}\sum_{j=m}^{n}\mathbf{P}\Big\{\frac{\hat{X}_j}{C_j} > 2a \geqq \max_{j+1\leqq k\leqq n}\frac{\hat{X}_k}{C_k}\Big\}$$
$$= \frac{1}{2}\mathbf{P}\Big\{\max_{m\leqq j\leqq n}\frac{\hat{X}_j}{C_j} > 2a\Big\}.$$

Letting $n\to\infty$ then $m\to\infty$ shows that

$$\mathbf{P}\Big\{\frac{\hat{S}_j}{C_j} > a \text{ i.o.}\Big\} \geqq \frac{1}{2}\mathbf{P}\Big\{\frac{\hat{X}_j}{C_j} > 2a \text{ i.o.}\Big\}. \tag{2.39}$$

Since $a < 1/2$ and $\sum(1-F(C_n)) = \infty$ the right-hand side of (2.39) is 1/2, and consequently, by the Hewitt–Savage 0–1 law, the left-hand side of (2.39) is 1. Thus

$$\limsup_{n\to\infty}\frac{\hat{S}_n}{C_n} \geqq a \quad \text{a.s.}. \tag{2.40}$$

But by (2.29) and the fact that $\mathbf{E}(X^+)^{1/\kappa}=\infty$,

$$\delta^\kappa \frac{C_n}{n^\kappa} \geqq (\rho_\kappa(C_n))^\kappa \to \infty,$$

and so (2.40) gives

$$\limsup_{n\to\infty} \frac{\hat{S}_n}{n^\kappa} = \infty \text{ a.s..} \tag{2.41}$$

Our final step is to transfer this to S_n. By (2.41), for each $x>0$ there are w.p. 1 infinitely many integers n_k such that

$$\sum_{i=1}^{n_k} (X_i^+ - \mathbf{E}X_i^+) \geqq x n_k^\kappa. \tag{2.42}$$

Let n_k be the successive indices for which (2.42) occurs and let $\mathcal{G}$ be the σ-field generated by $\{X_i^+, i\geqq 1\}$. The event that (2.42) occurs i.o. and the values of the n_k are $\mathcal{G}$-measurable. Now let W_i have the conditional distribution of $-X_i+\mathbf{E}\{X_i \mid X_i<0\}$, given that $X_i<0$, and take the W_i independent. From Remark (vi) of [15], noting that the r.v.'s W_i are bounded below and have mean 0, we have that for some $c>0$,

$$\min_{M\geq 1} \mathbf{P}\Big\{\sum_{i=1}^{M} W_i \leqq 0\Big\} \geqq c > 0. \tag{2.43}$$

Moreover, for any event $G\in\mathcal{G}$, conditionally on an event of the form

$$G\cap\{n_k=N\}\cap\{X_i\geqq 0 \text{ when } i \text{ is one of the indices } i_1<i_2<\cdots<i_r\leqq N \text{ but } X_i<0 \text{ for } i\in[1,N]\setminus\{i_1,\dots,i_r\}\},$$

the random variables $X_i^- - \mathbf{E}\{X_i^- \mid X_i<0\}I[X_i<0]$, $1\leqq i\leqq N$, are conditionally independent. They take the value 0 for $i=\{i_1,\dots,i_r\}$, and have the distribution of W_i when $i\in[1,N]\setminus\{i_1,\dots,i_r\}$. It follows that for any k, a.e. on the event $\{n_k<\infty\}$,

$$\mathbf{P}\Big\{\sum_{i=1}^{n_k} [X_i^- - \mathbf{E}\{X_i^- \mid X_i<0\}I[X_i<0]] \leqq 0 \mid \mathcal{G}\Big\}$$
$$\geqq \min_{M\geq 1} \mathbf{P}\Big\{\sum_{i=1}^{M} W_i \leqq 0\Big\} \geqq c.$$

But then

$$\mathbf{P}\Big\{\sum_{i=1}^{n_k} [X_i^- - \mathbf{E}\{X_i^- \mid X_i<0\}I[X_i<0]] \leqq 0 \text{ for infinitely many } k$$

for which (2.42) holds $\mid \mathcal{G}\Big\} = 1$

a.e. on the set on which (2.42) occurs i.o.. Now for any n_k for which (2.42) occurs as well as $\sum_{i=1}^{n_k} [X_i^- - \mathbf{E}\{X_i^- \mid X_i < 0\} I[X_i < 0]] \leqq 0$, we have (recall $\mathbf{E}X = 0$)

$$\begin{aligned} S_{n_k} &= \sum_{i=1}^{n_k} (X_i^+ - \mathbf{E}X_i^+) - \sum_{i=1}^{n_k} (X_i^- - \mathbf{E}X_1^-) \\ &= \sum_{i=1}^{n_k} (X_i^+ - \mathbf{E}X_i^+) - \sum_{i=1}^{n_k} [X_i^- - \mathbf{E}\{X_i^- \mid X_i < 0\} I[X_i < 0]] \\ &\qquad - \mathbf{E}\{X^- \mid X < 0\} \sum_{i=1}^{n_k} I[X_i < 0] + n_k \mathbf{E}X_1^- \\ &\geqq x n_k^\kappa - \mathbf{E}\{X^- \mid X < 0\} \sum_{i=1}^{n_k} I[X_i < 0] + n_k \mathbf{E}X^-. \end{aligned}$$

Now by the law of the iterated logarithm,

$$\lim_{k\to\infty} \frac{1}{n_k^\kappa} \left[-\mathbf{E}\{X^- \mid X < 0\} \sum_{i=1}^{n_k} I[X_i < 0] + n_k \mathbf{E}X^- \right] = 0 \quad \text{a.s..}$$

Thus $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s., proving (1.12).

(1.13) is immediate of course from Theorem 1.

Now we give an example to demonstrate (1.14). This is simply based on the observation in [16], Proposition 3.2, that $S_n \xrightarrow{P} \infty$ implies $S_n/n^\kappa \xrightarrow{P} \infty$ for $0 \leqq \kappa < 1$. The latter of course implies $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s., so we need only find a distribution function F for which $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, $\mathbf{E}(X^+)^{1/\kappa} < \infty = \mathbf{E}(X^-)^{1/\kappa}$ for some $1/2 < \kappa < 1$, but $A(x)/(xF(-x)) \to \infty$ as $x \to \infty$. The last condition implies $S_n \xrightarrow{P} \infty$ by [13], Theorem 2.1. The following example qualifies. Let $\delta > 0$ and

$$1 - F(x) = \frac{1}{x^{1+\delta}}, \quad x \geqq a_1, \quad F(-x) = \frac{1}{x(\log x)^2}, \quad x > a_2,$$

and

$$1 - F(x) = c_1, \quad 0 < x < a_1, \quad F(-x) = c_2, \quad 0 < x \leqq a_2.$$

Here a_1, a_2, c_1, c_2 are positive constants. Note that F has a jump of size $1 - c_1 - c_2$ at 0. We have

$$\mathbf{E}X^+ = \int_0^{a_1} c_1\,dx + \int_{a_1}^{\infty} \frac{dx}{x^{1+\delta}} = a_1 c_1 + \frac{1}{\delta a_1^\delta}$$

and

$$\mathbf{E}X^- = \int_0^{a_2} c_2 dx + \int_{a_2}^{\infty} \frac{dx}{x(\log x)^2} = a_2 c_2 + \frac{1}{\log a_2}.$$

We can choose the constants so that

$$\mathbf{E}X = (a_1 c_1 + 1/(\delta a_1^\delta)) - (a_2 c_2 + 1/\log a_2) = 0;$$

take, for example

$$a_1 = 10, \quad a_2 = e^k, \quad c_1 = \frac{1}{2}, \quad c_2 = \frac{5 + 1/(\delta 10^\delta) - 1/k}{e^k},$$

where k is so large that $c_2 < \frac{1}{2}$. Then $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, and for large x

$$A(x) = \int_0^x (1 - F(y) - F(-y))dy = -\int_x^\infty (1 - F(y) - F(-y))dy$$
$$= \int_x^\infty F(-y)dy - \int_x^\infty (1 - F(y))dy = \frac{1}{\log x} - \frac{1}{\delta x^\delta} \sim \frac{1}{\log x}.$$

Thus

$$\frac{A(x)}{xF(-x)} \sim \log x \to \infty$$

so $S_n \xrightarrow{P} \infty$ as required, yet $\mathbf{E}(X^+)^{1/\kappa} < \infty$ for all $\kappa > 1/(1+\delta)$ and $\mathbf{E}(X^-)^{1/\kappa} = \infty$ for all $0 < \kappa < 1$. This proves (1.14).

(1.15) is just the contrapositive of the following lemma.

LEMMA 2.2. *Suppose* $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, $\mathbf{E}(X^+)^{1/\kappa} < \infty$, *and*

$$\int_0^\infty \left(\int_x^\infty F(-y)dy \right)^{\kappa/(1-\kappa)} dx < \infty. \tag{2.44}$$

Then

$$\limsup_{n\to\infty} \frac{S_n}{n^\kappa} \leq 0 \quad a.s.. \tag{2.45}$$

PROOF OF LEMMA 2.2. If $\mathbf{E}(X^-)^{1/\kappa} < \infty$ then $S_n/n^\kappa \to 0$ a.s. by (2.2), so (2.45) holds. We therefore assume $\mathbf{E}(X^-)^{1/\kappa} = \infty$ and consequently

$F(-x) > 0$ for all $x > 0$ for the remainder of this proof. We have, since $\mathbf{E}X_i^+ = \mathbf{E}X_i^-$,

$$\frac{S_n}{n^\kappa} = \frac{\sum_{i=1}^n (X_i^+ - \mathbf{E}X_i^+)}{n^\kappa} + \frac{\sum_{i=1}^n (\mathbf{E}X_i^- - X_i^-)}{n^\kappa}. \tag{2.46}$$

The first term on the right-hand side of (2.46) is $o(1)$ a.s. by the Marcinkiewicz–Zygmund law. Thus, letting

$$\overline{X}_i = \mathbf{E}X_i^- - X_i^- \quad \text{and} \quad \overline{S}_n = \sum_{i=1}^n \overline{X}_i,$$

it will suffice to show that $\limsup_{n\to\infty} \overline{S}_n/n^\kappa \leqq 0$ a.s.. For $\delta > 0$, $x > 1$, define

$$C(x) = \inf\left\{y : \int_y^\infty F(-z)dz \leqq \frac{\delta}{x^{1-\kappa}}\right\}, \tag{2.47}$$

so that for large x

$$x^{1-\kappa} \int_{C(x)}^\infty F(-z)dz = \delta. \tag{2.48}$$

An argument just like that in (2.33) shows that for $x \geqq \mathbf{E}X^-$

$$\overline{A}(x) := \mathbf{E}((\overline{X}_i \wedge x) \vee (-x)) = \int_{x+\mathbf{E}X^-}^\infty F(-y)dy \leqq \int_x^\infty F(-y)dy,$$

so

$$x^{1-\kappa}\overline{A}(C(x)) \leqq \delta. \tag{2.49}$$

Note also that the $\overline{X}_i = \mathbf{E}X_i^- - X_i^-$ are bounded above, so by Remark (vi) of [15] we have for some $\Delta > 0$

$$\min_{n\geqq 1} \mathbf{P}\{\overline{S}_n \geqq 0\} \geqq \Delta > 0. \tag{2.50}$$

It follows that

$$\mathbf{P}\{\max_{1\leqq j\leqq n} \overline{S}_j > x\} = \sum_{j=1}^n \mathbf{P}\{\max_{1\leqq k\leqq j-1} \overline{S}_k \leqq x < \overline{S}_j\}$$

$$\text{(2.51)} \qquad \leqq \frac{1}{\Delta}\sum_{j=1}^{n}\mathbf{P}\{\max_{1\leqq k\leqq j-1}\overline{S}_k\leqq x<\overline{S}_j,\ \sum_{i=j+1}^{n}\overline{X}_i\geqq 0\}$$
$$\leqq \frac{1}{\Delta}\mathbf{P}\{\overline{S}_n>x\}.$$

Now set

$$\overline{Z}_i^k=(\overline{X}_i\wedge C(2^k))\vee(-C(2^k)),$$

and use (2.51) and (2.49) to write

$$\mathbf{P}\{\max_{1\leqq n\leqq 2^k}\overline{S}_n>2\delta 2^{\kappa k}\}\leqq\frac{1}{\Delta}\mathbf{P}\{\overline{S}_{2^k}>2\delta 2^{\kappa k}\}$$
$$\leqq\frac{1}{\Delta}\mathbf{P}\Big\{\sum_{i=1}^{2^k}(\overline{X}_i\wedge C(2^k))\vee(-C(2^k))-2^k\overline{A}(C(2^k))>\delta 2^{\kappa k}\Big\}$$
$$\text{(2.52)} \qquad +\frac{2^k}{\Delta}\mathbf{P}\{\overline{X}_1>C(2^k)\}$$
$$=\frac{1}{\Delta}\mathbf{P}\Big\{\sum_{i=1}^{2^k}(\overline{Z}_i^k-\mathbf{E}\overline{Z}_i^k)>\delta 2^{\kappa k}\Big\}$$

for k large enough; note that $\mathbf{P}\{\overline{X}_1>C(2^k)\}=0$ for k so large that $C(2^k)>\mathbf{E}X^-$. We have (cf. (2.35))

$$\mathbf{E}(\overline{Z}_i^k)^2\leqq 4c_-\int_0^{C(2^k)} yF(-y)dy=2c_-U_-(C(2^k)),\quad\text{say,}$$

where c_- is some constant depending only on the distribution of X^-. From (2.52) and Chebychev's inequality we get

$$\text{(2.53)} \qquad \mathbf{P}\{\max_{1\leqq n\leqq 2^k}\overline{S}_n>2\delta 2^{\kappa k}\}\leqq\frac{2c_-2^kU_-(C(2^k))}{\Delta\delta^2 2^{2\kappa k}}=\frac{cU_-(C(2^k))}{\delta^2 2^{(2\kappa-1)k}}$$

for some constant $c>0$. Next,

$$\sum_{k\geq 1}\frac{U_-(C(2^k))}{2^{(2\kappa-1)k}}=\sum_{k\geq 1}\frac{1}{2^{(2\kappa-1)k}}\sum_{j=1}^{k}2\int_{C(2^{j-1})}^{C(2^j)}xF(-x)dx$$
$$+U_-(C(1))\sum_{k\geq 1}\frac{1}{2^{(2\kappa-1)k}}$$

$$= 2\sum_{j\geqq 1}\int_{C(2^{j-1})}^{C(2^j)} xF(-x)dx \sum_{k\geqq j}\frac{1}{2^{(2\kappa-1)k}} + \frac{U_-(C(1))}{2^{2\kappa-1}-1} \tag{2.54}$$

$$= \frac{2^{2\kappa}}{2^{2\kappa-1}-1}\sum_{j\geqq 1}\int_{C(2^{j-1})}^{C(2^j)} \frac{xF(-x)dx}{2^{(2\kappa-1)j}} + \frac{U_-(C(1))}{2^{2\kappa-1}-1}.$$

Now by (2.47), $x \leqq C(2^j)$ implies $\int_x^\infty F(-y)dy \geqq \delta/2^{j(1-\kappa)}$, so (2.53) and (2.54) give, for some c_1, c_2,

$$\begin{aligned} &\delta^{1/(1-\kappa)}\sum_{k\geqq 1}\mathbf{P}\{\max_{1\leqq n\leqq 2^k}\overline{S}_n > 2\delta 2^{\kappa k}\} \\ &\leqq c_1\int_{C(1)}^{\infty}\left(\int_x^\infty F(-y)dy\right)^{(2\kappa-1)/(1-\kappa)} xF(-x)dx + c_2. \end{aligned} \tag{2.55}$$

Integration by parts shows that the last integral converges if (and only if) (2.44) holds. Consequently by the Borel–Cantelli lemma

$$\limsup_{k\to\infty}\frac{\max_{1\leqq n\leqq 2^k}\overline{S}_n}{2^{\kappa k}} \leqq 2\delta \quad \text{a.s.}. \tag{2.56}$$

Thus for large k and $2^{k-1} < j \leqq 2^k$,

$$\overline{S}_j \leqq [\max_{2^{k-1}\leqq n\leqq 2^k}\overline{S}_n] \vee 0 \leqq 3\delta 2^{\kappa k} \leqq 3\delta 2^\kappa j^\kappa.$$

Given a large n choose $k = k(n)$ so that $2^{k-1} \leqq n < 2^k$. Then

$$\frac{\overline{S}_n}{n^\kappa} \leqq \frac{2^\kappa[\max_{1\leqq j\leqq 2^k}\overline{S}_j]\vee 0}{2^{\kappa k}}.$$

This shows that $\limsup_{n\to\infty}\overline{S}_n/n^\kappa \leqq 3\delta 2^\kappa$. Let $\delta\downarrow 0$ to get (2.45).

PROOF OF THEOREM 3. Since $\mathbf{E}(X^+)^{1/\kappa} < \infty$, we have by the Marcinkiewicz–Zygmund law (see (2.2)) that

$$\frac{\sum_{i=1}^n [X_i^+ - \mathbf{E}\{X_i \mid X > 0\}I[X_i > 0]]}{n^\kappa} \to 0 \quad \text{a.s.},$$

so we may replace X_i by the constant $\mathbf{E}\{X \mid X>0\}$ when $X_i>0$, without influencing the value of $\limsup S_n/n^\kappa$. We may therefore assume that X^+ can take only one value >0. We denote by $\widetilde{X}_i$ the modified random variable which is obtained by replacing X_i by $\mathbf{E}\{X \mid X>0\}$ when $X_i>0$. Its characteristic function is Φ. We further write $\widetilde{S}_n=\sum_1^n \widetilde{X}_i$.

Now let $\sigma_1<\sigma_2<\cdots$ be the successive strict upward ladder indices of the random walk $\widetilde{S}_n$ and take $\sigma_0=0$. Then the $\widetilde{S}_{\sigma_{k+1}}-\widetilde{S}_{\sigma_k}$ are i.i.d. and take values in $(0,\mathbf{E}\{X \mid X>0\}]$, so that

$$\frac{1}{n}\widetilde{S}_{\sigma_n}=\frac{1}{n}\sum_{k=1}^{n}(\widetilde{S}_{\sigma_k}-\widetilde{S}_{\sigma_{k-1}})\to\mathbf{E}\{\widetilde{S}_{\sigma_1}\}\in(0,\mathbf{E}\{X \mid X>0\}] \quad \text{a.s.}. \tag{2.57}$$

Since

$$\limsup_{n\to\infty}\widetilde{S}_n/n^\kappa=\infty \quad \text{a.s.} \tag{2.58}$$

if and only if

$$\limsup_{n\to\infty}\frac{\widetilde{S}_{\sigma_n}}{\sigma_n^\kappa}=\infty \quad \text{a.s.}, \tag{2.59}$$

we see from (2.57) that (2.58) is equivalent to

$$\liminf_{n\to\infty}\frac{\sigma_n}{n^{1/\kappa}}=0 \quad \text{a.s.}.$$

By [25], Theorem 1, this is in turn equivalent to

$$\int_1^\infty \exp\left\{-\lambda\frac{(m(x))^{1/(1-\kappa)}}{x}\right\}\frac{dx}{x}=\infty \tag{2.60}$$

for all $\lambda>0$, where for $x>0$,

$$m(x)=\mathbf{E}(\sigma_1\wedge x)=\sum_{k\le x}k\mathbf{P}\{\sigma_1=k\}+x\mathbf{P}\{\sigma_1>x\}.$$

We now estimate $\mathbf{E}\{\sigma_1\wedge x\}$. For brevity write $K=\mathbf{E}\{X \mid X>0\}$ and $C_1=\mathbf{P}\{\widetilde{X}=\mathbf{E}\{X \mid X>0\}\}=\mathbf{P}\{X>0\}$. Then $\widetilde{X}_i^+=0$ or $=K$. Therefore,

$$\begin{aligned}\mathbf{P}\{\sigma_1=n\}&=\mathbf{P}\{\widetilde{S}_{n-1}\in(-K,0],\widetilde{S}_i\leqq 0, 0\leqq i\leqq n-1\}\mathbf{P}\{\widetilde{X}=K\}\\&=C_1\mathbf{P}\{\widetilde{S}_{n-1}\in(-K,0],\widetilde{S}_i\leqq 0,0\leqq i\leqq n-1\}.\end{aligned} \tag{2.61}$$

Next we use a simple argument based on the fact that all cyclical permutations of $\tilde{X}_1, \ldots, \tilde{X}_{n-1}$ are equally likely (see [23]; proof of Proposition 32.5). If ν is any index $\leqq n-1$ at which $\max\limits_{0 \leqq i \leqq n-1} \tilde{S}_i$ is achieved, and $\tilde{S}_{n-1} = \sum_{i=1}^{n-1} \tilde{X}_i \leqq 0$, then at least the cyclical permutation

$$\tilde{X}_{\nu+1}, \ldots, \tilde{X}_{n-1}, \tilde{X}_1, \ldots, \tilde{X}_\nu, \tag{2.62}$$

has all the partial sums $\leqq 0$, as one easily checks. Therefore, by (2.61)

$$\mathbf{P}\{\sigma_1 = n\} \geqq \frac{C_1}{n} \mathbf{P}\{\tilde{S}_{n-1} \in (-K, 0]\}. \tag{2.63}$$

Now first assume that $\{\tilde{S}_n\}$ is aperiodic, in the sense of [23], Definition 2.2. Since $\mathbf{E}\tilde{X} = 0$, $\{\tilde{S}_n\}$ is interval recurrent. From this it follows that for any fixed number $L > 0$ and set $A \subset [-L, L]$ and any open interval $I = (a - \eta, a + \eta)$, say, there exists a $1 \leqq j = j(L, I) < \infty$ and a constant $C_2 = C_2(L, I) > 0$ such that uniformly in n,

$$\begin{aligned}
&\sum_{i=1}^{j} \mathbf{P}\{\tilde{S}_{n+i} \in I \mid \tilde{S}_n \in A\} \\
&\geqq \min_{x \in A} \sum_{i=1}^{j} \mathbf{P}\{\tilde{S}_i + x \in I\} \\
&\geqq \min_{\substack{p \in \mathbb{Z} \\ |p| \leqq 2L/\eta}} \sum_{i=1}^{j} \mathbf{P}\left\{\left|\tilde{S}_i + p\frac{\eta}{2} - a\right| < \frac{\eta}{2}\right\} \\
&\geqq C_2.
\end{aligned}$$

It follows that for any fixed $\delta > 0$ there exists some j and $C_3 > 0$ such that

$$\begin{aligned}
\sum_{i=1}^{j} \mathbf{P}\{\sigma_1 = n + i\} &\geqq \frac{C_1}{n} \sum_{i=1}^{j} \mathbf{P}\{\tilde{S}_{n-1+i} \in (-K, 0]\} \\
&\geqq \frac{C_3}{n} \int_0^\delta \mathbf{P}\{|\tilde{S}_{n-1}| < s\} ds.
\end{aligned} \tag{2.64}$$

If $\{\tilde{S}_n\}$ is periodic, then it can take all values $k\lambda$, $k \in \mathbb{Z}$, for some $\lambda \neq 0$. Then the preceding argument still goes through if we restrict x to multiples of λ and if I contains a multiple of λ. Therefore, (2.64) is valid also in the periodic case.

For an estimate in the opposite direction we introduce the set

$$\Gamma = \Gamma_n = \Big\{x = (x_1, \dots, x_n) : \sum_{i=1}^{j} x_i \leqq 0,\ 1 \leqq j \leqq n-1, \text{ and } \sum_{i=1}^{n} x_i \in (0, K]\Big\},$$

and for any sequence of length n, $x = (x_1, \dots, x_n)$, we define its cyclical permutations $\tau^k x = (x_{k+1}, \dots, x_n, x_1, \dots, x_k)$, $0 \leqq k \leqq n-1$. By (2.61)

$$\mathbf{P}\{\sigma_1 = n\} = \mathbf{P}\{(\tilde{X}_1, \dots, \tilde{X}_n) \in \Gamma\}. \tag{2.65}$$

Next we note that for any $x \in \Gamma$, none of the permutations $\tau^k x$ with $1 \leqq k \leqq n-1$ lies in Γ, because

$$\sum_{i=k+1}^{n} x_i = \sum_{i=1}^{n} x_i - \sum_{i=1}^{k} x_i > 0.$$

Moreover for $x \in \Gamma$ and $k < l$, we must have $\tau^k x \neq \tau^l x$. Indeed $\tau^k x = \tau^l x$ would imply that the periodic extension of x with period n would also have period $l - k$ and then also period $p := \text{g. c. d.}(n, l-k)$. But this would force $\sum_{i=1}^{p} x_i > 0$, because $\sum_{i=1}^{n} x_i > 0$, and $\sum_{i=1}^{p} x_i > 0$ is impossible for $x \in \Gamma$ and $p < n$. Therefore, given $\tau^k x$ for some $x \in \Gamma$, one can find k and x uniquely, and the sets $\tau^k \Gamma := \{\tau^k x : x \in \Gamma\}$, $0 \leqq k \leqq n-1$, are disjoint. Finally, if we take into account that

$$\sum_{i=1}^{n} (\tau^k x)_i = \sum_{i=1}^{n} x_i \in (0, K]$$

for any $x \in \Gamma$, we obtain

$$\begin{aligned}\mathbf{P}\{\sigma_1 = n\} = \mathbf{P}\{(\tilde{X}_1, \dots, \tilde{X}_n) \in \Gamma\} &= \frac{1}{n}\mathbf{P}\Big\{(\tilde{X}_1, \dots, \tilde{X}_n) \in \bigcup_{k=0}^{n-1} \tau^k \Gamma\Big\} \\ &\leqq \frac{1}{n}\mathbf{P}\{\tilde{S}_n \in (0, K]\}.\end{aligned}$$

Analogously to (2.64) we then also have for any fixed $\delta > 0$ that there exist constants $C_4 < \infty$ and $j < \infty$ so that

$$\mathbf{P}\{\sigma_1 = n\} \leqq \frac{C_4}{n} \int_0^{\delta} \sum_{i=1}^{j} \mathbf{P}\{|\widetilde{S}_{n+i}| < s\} ds. \tag{2.66}$$

Now

$$m(x) = \mathbf{E}(\sigma_1 \wedge x) \uparrow \mathbf{E}\sigma_1 = \infty, \quad (x \to \infty)$$

because $\mathbf{E}X = 0$ (see [9], Theorem 9.2). It follows from this and (2.64) that for some constant $C_5 > 0$

$$m(x) = \sum_{n \geqq 1} (n \wedge x) \mathbf{P}\{\sigma_1 = n\}$$

$$\geqq C_5 \sum_{i=1}^{j} \sum_{n \geqq 1} (n \wedge x) \mathbf{P}\{\sigma_1 = n + 1 + i\}$$

$$\geqq \sum_{n \geqq 1} \frac{C_5 C_3}{n} (n \wedge x) \int_0^\delta \mathbf{P}\{|\widetilde{S}_n| < s\} ds.$$

Similarly, by means of (2.66),

$$m(x) \leqq \sum_{n \geqq 1} \frac{C_6}{n} (n \wedge x) \int_0^\delta \mathbf{P}\{|\widetilde{S}_n| < s\} ds.$$

We may therefore replace $m(x)$ in (2.60) by

$$\sum_{n \geqq 1} \int_0^\delta \left(1 \wedge \frac{x}{n}\right) \mathbf{P}\{|\widetilde{S}_n| < s\} ds, \tag{2.67}$$

and if we define $(1 \wedge x/0) = 1$, then we may even start the sum at $n = 0$.

The proof of (1.16) now only requires a few manipulations from analysis. One easily checks that $(1 \wedge x/n)$ lies between two constant multiples of

$$\int_0^1 e^{-ny/x} dy.$$

Moreover

$$\int_0^\delta \mathbf{P}\{|\widetilde{S}_n| < s\} ds = \frac{1}{\pi} \int_{-\infty}^{\infty} \Phi^n(\theta) \frac{1 - \cos \delta\theta}{\theta^2} d\theta \tag{2.68}$$

(see Chung and Fuchs [4]). Therefore the sum in (2.67) (starting with $n = 0$) lies between two constant multiples of

$$\begin{aligned} &\sum_{n=0}^{\infty} \int_0^1 e^{-ny/x} dy \int_{-\infty}^{\infty} \Phi^n(\theta) \frac{1 - \cos \delta\theta}{\theta^2} d\theta \\ &= \int_0^1 dy \int_{-\infty}^{\infty} \frac{1 - \cos \delta\theta}{\theta^2} \cdot \frac{1}{1 - e^{-y/x}\Phi(\theta)} d\theta. \end{aligned} \tag{2.69}$$

Note that the interchange of the summation and integration here is justified by the following estimates

$$\sum_{n=0}^{\infty}\int_0^1 e^{-ny/x}dy\int_{-\infty}^{\infty}|\Phi(\theta)|^n\frac{1-\cos\delta\theta}{\theta^2}d\theta$$

$$\leqq\sum_{n=0}^{\infty}\int_0^1 e^{-2\lfloor n/2\rfloor y/x}dy\int_{-\infty}^{\infty}|\Phi(\theta)|^{2\lfloor n/2\rfloor}\frac{1-\cos\delta\theta}{\theta^2}d\theta$$

$$=2\pi\sum_{m=0}^{\infty}\int_0^1 e^{-2my/x}dy\int_0^{\delta}\mathbf{P}\{|T_m|<s\}ds,$$

where $T_m=\sum_1^m\widetilde{X}_i-\sum_1^m\widetilde{X}'_i$, with all $\widetilde{X}_i,\widetilde{X}'_j$ i.i.d. (because $\widetilde{X}_i-\widetilde{X}'_i$ has characteristic function $|\Phi(\theta)|^2$). But $\mathbf{E}X=0$ and X not degenerate at 0 implies that $\mathbf{P}\{X>0\}>0$ and $\mathbf{P}\{X<0\}>0$, so that $\widetilde{X}$ is not degenerate. We then have for some constant $C_7<\infty$ that

$$\int_0^{\delta}\mathbf{P}\{|T_m|<s\}ds\leqq\delta\mathbf{P}\{|T_m|<\delta\}\leqq\frac{C_7}{\sqrt{m+1}},$$

by a general concentration function inequality ([7], Theorem 3.1). Thus

$$\sum_{m=0}^{\infty}\int_0^1 e^{-2my/x}dy\int_0^{\delta}\mathbf{P}\{|T_m|<s\}ds=O\Big(\sum_{m=0}^{\infty}\Big(1\wedge\frac{x}{m}\Big)\frac{1}{\sqrt{m+1}}\Big)<\infty.$$

Thus the interchange of summation and integration is permissible by Fubini's theorem.

Replacing $m(x)$ in (2.60) by the right-hand side of (2.69) with $\delta=1$ and changing the variable gives (1.16). □

PROOF OF THEOREM 4. (a) Let $\kappa>1$ or $\frac{1}{2}<\kappa<1$, $\mathbf{E}|X|<\infty$ and $\mathbf{E}X=0$, and $\limsup_{n\to\infty}|S_n|/n^{\kappa}>0$ a.s.. If $\limsup_{n\to\infty}|S_n|/n^{\kappa}<\infty$ a.s., then $\mathbf{E}|X|^{1/\kappa}<\infty$ so $|S_n|/n^{\kappa}\to 0$ a.s. by (2.2), which is a contradiction.

(b) We consider the various cases, excluding the two mentioned.

The case $\kappa>1$ is covered by Corollary 1 of [2].

For $\kappa=1$, we know that $\limsup_{n\to\infty}S_n/n$ can only take the values $+\infty$ or $-\infty$ when $\mathbf{E}|X|=\infty$ by Corollary 3 of [11], proving (1.19) in this case. If $\mathbf{E}|X|<\infty$ and $\mathbf{E}X=0$ then neither condition in (1.19) can occur for $\kappa=1$, so the result is true vacuously.

Next let $0 \leqq \kappa \leqq \frac{1}{2}$. If $\limsup_{n\to\infty} S_n/n^\kappa > 0$ a.s., then $\limsup_{n\to\infty} S_n = \infty$ a.s., so $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. by (1.9). (Note that (1.19) is trivial if $F(0-) = 0$ or $F(0) = 1$.)

Next let $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X \neq 0$. If $\limsup_{n\to\infty} S_n/n^\kappa > 0$ a.s., then $\mathbf{E}X > 0$ (by 1.10)), so $S_n/n^\kappa \to \infty$ a.s. by the strong law of large numbers, and (1.19) is true. When $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| = \infty$ and $\limsup_{n\to\infty} S_n/n < \infty$ a.s., then we have $S_n/n \to -\infty$ a.s. by Corollary 3 of [11]. This is not possible when $\limsup_{n\to\infty} S_n/n^\kappa > 0$ a.s., so the latter implies $\limsup_{n\to\infty} S_n/n = \infty$ a.s. and consequently $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s.. Again (1.19) is true.

When $\kappa = 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X \neq 0$, (1.19) is obviously not true by the strong law of large numbers.

Finally, let $\frac{1}{2} < \kappa < 1$. We will use a result of Klass [18], [19] to give a random walk with $\mathbf{E}|X| < \infty$ and $\mathbf{E}X = 0$ and

$$0 < \limsup_{n\to\infty} \frac{S_n}{n^\kappa} < \infty \quad \text{a.s.,} \tag{2.70}$$

showing that (1.19) is not true in this case either. To do this, define as in Klass ([19], Equation (2.1), p. 152) a positive function $K(x)$ (uniquely) satisfying, for $x > 0$,

$$K^2(x) = x\mathbf{E}X^2 I(|X| \leqq K(x)) + xK(x)\mathbf{E}|X| I(|X| > K(x)).$$

Clearly $K(x) \to \infty$ as $x \to \infty$, and integration by parts shows that

$$K^2(x) = 2x \int_0^{K(x)} y\mathbf{P}(|X| > y)dy + xK(x) \int_{K(x)}^{\infty} \mathbf{P}(|X| > y)dy. \tag{2.71}$$

Then by Theorem 2.5 of Klass [19],

$$1 \leqq \limsup_{n\to\infty} \frac{S_n}{l_2 n K(n/l_2 n)} \leqq 1.5 \quad \text{a.s.} \tag{2.72}$$

provided

$$\sum_{n \geqq e} \mathbf{P}\{X > l_2 n K(n/l_2 n)\} < \infty. \tag{2.73}$$

Here $l_2 n = \log\log n$ for $n > e$. We can satisfy (2.73) by taking X to be bounded above, so to prove (2.70) we need merely find a random walk, whose increments are bounded above, with $\mathbf{E}|X| < \infty$, $\mathbf{E}X = 0$, and, for some $c > 0$,

$$l_2 n K(n/l_2 n) \sim c n^\kappa \quad (n \to \infty). \tag{2.74}$$

To do this, let the distribution F of X put all its mass in $(0,\infty)$ on one point and satisfy

$$F(-x) \sim cx^{-1/\kappa}(l_2x)^{1-1/\kappa}, \qquad x\to\infty, \tag{2.75}$$

and $\mathbf{E}X=0$. (Note that $\mathbf{E}|X|<\infty$ when (2.75) holds.) Since X is bounded above we have for some x_1 and large x

$$\begin{aligned}\int_0^x y\mathbf{P}(|X|>y)dy &\sim (2-1/\kappa)\int_{x_1}^x \frac{cydy}{y^{1/\kappa}(l_2y)^{(1-\kappa)/\kappa}} + \text{const} \\ &\sim \frac{x^{2-1/\kappa}c}{(l_2x)^{(1-\kappa)/\kappa}}, \qquad \text{as } x\to\infty.\end{aligned}$$

Also, as $x\to\infty$,

$$\int_x^\infty \mathbf{P}(|X|>y)dy \sim (2-1/\kappa)\int_x^\infty \frac{cdy}{y^{1/\kappa}(l_2y)^{(1-\kappa)/\kappa}} \sim \frac{(2-1/\kappa)cx^{1-(1/\kappa)}}{(1/\kappa-1)(l_2x)^{(1-\kappa)/\kappa}}.$$

Thus $K(x/l_2x)$ satisfies, by (2.71),

$$\begin{aligned}\frac{K^2(x/l_2x)}{(x/l_2x)} &= 2\int_0^{K(x/l_2x)} y\mathbf{P}(|X|>y)dy + K(x/l_2x)\int_{K(x/l_2x)}^\infty \mathbf{P}(|X|>y)dy \\ &\sim \frac{(K(x/l_2x))^{2-1/\kappa}c}{(1-\kappa)(l_2(K(x/l_2x)))^{(1-\kappa)/\kappa}},\end{aligned}$$

or

$$(K(x/l_2x))^{1/\kappa} \sim \frac{cx}{(1-\kappa)(l_2x)(l_2(K(x/l_2x)))^{(1-\kappa)/\kappa}}. \tag{2.76}$$

Clearly then, $\log(K(x/l_2x))$ is bounded above and below by multiples of $\log x$, so $l_2(K(x/l_2x))\sim l_2x$. Thus

$$(K(x/l_2x))^{1/\kappa} \sim \frac{cx}{(1-\kappa)(l_2x)^{1/\kappa}},$$

proving (2.74). Hence (2.70) holds by (2.72). □

Remark. It is easy to see that (1.15) holds for the example just given, so this example also demonstrates that (1.15) is not sufficient for $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s..

PROOF OF THEOREM 5. (a) Keep $\kappa \geqq 0$ and $a > 0$. It follows trivially from the definitions that $\limsup_{n\to\infty} |S_n|/n^\kappa > a$ a.s. implies $T_\kappa(a) < \infty$ a.s., so Theorem 1 shows that we have $T_\kappa(a) < \infty$ a.s. for all $a > 0$ under the specified conditions.

(b) Keep $\kappa = 1$ and $0 < |\mathbf{E}X| \leqq \mathbf{E}|X| < \infty$. Suppose $\mathbf{E}X > 0$. By the recurrence of the random walk $\sum_{i=1}^n (X_i - \mathbf{E}X)$, we have

$$\begin{aligned} 1 &= \mathbf{P}\{S_n - n\mathbf{E}X > 0 \text{ i.o.}\} = \mathbf{P}\{S_n > n\mathbf{E}X \text{ i.o.}\} \leqq \mathbf{P}\{(|S_n|/n) > |\mathbf{E}X| \text{ i.o.}\} \\ &\leqq \mathbf{P}\{\max_{n\geqq 1}(|S_n|/n) > |\mathbf{E}X|\} = \mathbf{P}\{T_1(|\mathbf{E}X|) < \infty\}. \end{aligned}$$

Using a similar argument when $\mathbf{E}X < 0$, we see that $T_1(|\mathbf{E}X|) < \infty$ a.s. when $\mathbf{E}X \neq 0$, so $T_1(a) < \infty$ a.s. for all $a \leqq |\mathbf{E}X|$.

Next take $a > |\mathbf{E}X|$ and $\delta > 0$. By the strong law of large numbers we have $|S_n|/n \to |\mathbf{E}X|$ a.s. so we can choose $m_0(a,\delta)$ so large that

$$\mathbf{P}\{\max_{m<j\leqq n}(|S_j|/j) > a\} \leqq \mathbf{P}\{(|S_j|/j) > a \text{ for some } j > m\} \leqq \delta/2 \tag{2.77}$$

whenever $n \geqq m \geqq m_0$. We can then choose $x_0 = x_0(m_0, \delta)$ so large that

$$\mathbf{P}\{\max_{1\leqq j\leqq m_0}(|S_j|/j) > x\} \leqq \delta/2 \tag{2.78}$$

whenever $x \geqq x_0$. Thus for $x \geqq x_0 \wedge a$ and $n > m_0$ we have

$$\begin{aligned} &\mathbf{P}\{\max_{1\leqq j\leqq n}(|S_j|/j) > x\} \\ &\leqq \mathbf{P}\{\max_{1\leqq j\leqq m_0}(|S_j|/j) > x\} + \mathbf{P}\{\max_{m_0<j\leqq n}(|S_j|/j) > a\} \leqq \delta. \end{aligned} \tag{2.79}$$

Letting $n \to \infty$ in this shows that $\mathbf{P}\{\sup_{j\geqq 1}(|S_j|/j) > x\} \leqq \delta$ for $x \geqq x_0$, which proves $\mathbf{P}\{T_1(x) = \infty\} \to 1$ as $x \to \infty$.

(c) Now let $\mathbf{E}|X|^{1/\kappa} < \infty$ and either $\kappa > 1$ or $\frac{1}{2} < \kappa \leqq 1$ (so $\mathbf{E}|X| < \infty$) and $\mathbf{E}X = 0$. Then by the Marcinkiewicz–Zygmund law (see (2.2)) we have $|S_n|/n^\kappa \to 0$ a.s.. The same working as in (2.77)–(2.79), with the divisor j replaced by j^κ, and a positive but otherwise arbitrary, shows that $\mathbf{P}\{T_\kappa(x) = \infty\} \to 1$ as $x \to \infty$. □

PROOF OF THEOREM 6. (a) Keep $0 \leqq \kappa \leqq 1$ and suppose $T^*_\kappa(a) < \infty$ a.s. for some fixed $a \geqq 0$. Take $T^*_0 = 0 = S_0$, $T^*_1 = T^*_\kappa(a)$, and for $k = 2, 3, \ldots$, define

$$T^*_k = \min\left\{n > T^*_{k-1} : S_n > S_{T^*_{k-1}} + a(n - T^*_{k-1})^\kappa\right\}.$$

Let $\Delta_k = T^*_k - T^*_{k-1}$, $k \geqq 1$. Then the Δ_k are i.i.d., each with the same distribution as T^*_1, and so $\Delta_k < \infty$ a.s.. Hence $T^*_k < \infty$ a.s., $k = 1, 2, \ldots$, and

$T_k^* \to \infty$ a.s. as $k \to \infty$. Now for all $k \geqq 1$

$$\begin{aligned} S_{T_k^*} &= \sum_{j=1}^{k} (S_{T_j^*} - S_{T_{j-1}^*}) > a \sum_{j=1}^{k} (T_j^* - T_{j-1}^*)^\kappa \\ &= a \sum_{j=1}^{k} \Delta_j^k > a \Big(\sum_{j=1}^{k} \Delta_j \Big)^\kappa \qquad \text{(when } 0 \leqq \kappa \leqq 1) \\ &= a (T_k^*)^\kappa. \end{aligned}$$

Thus, for all k, $S_{T_k^*} > a(T_k^*)^\kappa$ a.s., proving that $\limsup_{n\to\infty} S_n/n^\kappa \geqq a$ a.s..

Now take $\kappa > 1$. Let X be such that $X \geqq x_0$ a.s., where $x_0 > 0$ is a constant, but $\mathbf{E}(X^{1/\kappa}) < \infty$. Then when $a < x_0$

$$1 = \mathbf{P}\{X_1 > a\} \leqq \mathbf{P}\{\max_{j \geqq 1}(S_j/j^\kappa) > a\},$$

so that $T_\kappa^*(a) < \infty$ a.s.. But $S_n/n^\kappa \to 0$ a.s. by the Marcinkiewicz–Zygmund law (2.2).

(b) Keep $0 \leqq \kappa \leqq 1$, $a > 0$, and let $T_\kappa^*(a) < \infty$ a.s.. Then $\limsup_{n\to\infty} S_n/n^\kappa \geqq a > 0$ a.s. by part (a), so $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. by Theorem 4(b), except possibly in the cases $\kappa = 1$ and $\mathbf{E}|X| < \infty$, $\mathbf{E}X \neq 0$, or $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X = 0$. This proves (1.24). In the exceptional cases, we may have

$$\limsup_{n\to\infty} \frac{S_n}{n^\kappa} = a \in (0, \infty) \quad \text{a.s.,} \tag{2.80}$$

as was shown in Theorem 4 (b), and if (2.80) occurs, then for all $x < a$, $\mathbf{P}\{S_n/n^\kappa > x \text{ i.o.}\} = 1$, and hence $T_\kappa^*(x) < \infty$ a.s.. A similar proof to that of Theorem 5 shows that, also, $\mathbf{P}\{T_\kappa^*(x) = \infty\} \to 1$ as $x \to \infty$.

(c) Part (a) of the present theorem shows that, when $0 \leqq \kappa \leqq 1$, $T_\kappa^*(a) < \infty$ a.s. for some $a > 0$ if and only if $\limsup_{n\to\infty} S_n/n^\kappa > b$ a.s. for some $b > 0$. In turn, with the exception of the cases $\kappa = 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X \neq 0$, or possibly $\frac{1}{2} < \kappa < 1$, $\mathbf{E}|X| < \infty$ and $\mathbf{E}X = 0$, this occurs if and only if $\limsup_{n\to\infty} S_n/n^\kappa = \infty$ a.s. (by Theorem 4 (b)). Then $T_\kappa^*(a) < \infty$ a.s. for all $a > 0$ apart from those exceptional cases. We can read off the corresponding analytic equivalences from Theorem 2 (b)–(e).

(d) Suppose $T_1^*(a) < \infty$ a.s. for some $a > 0$. Then by part (a) of the present theorem, we have $\limsup_{n\to\infty} S_n/n \geqq a$ a.s.. If $\mathbf{E}|X| = \infty$ then by Kesten ([11], Corollary 3, p. 1195) we have $\limsup_{n\to\infty} S_n/n = \infty$ a.s. and $J_+ = \infty$ (by

(1.7)), and hence $T_1^*(a) < \infty$ a.s. for all $a > 0$. If $\mathbf{E}|X| < \infty$, then by the strong law of large numbers, $\mathbf{E}X \geq a$, hence $\mathbf{E}X > 0$. Now

$$\mathbf{P}\left\{\frac{S_n}{n} > \mathbf{E}X \text{ i.o.}\right\} = \mathbf{P}\left\{\sum_{i=1}^{n}(X_i - \mathbf{E}X) > 0 \text{ i.o.}\right\} = 1$$

so $T_1^*(a) < \infty$ a.s. even for $a = \mathbf{E}X$, hence for $0 \leq a \leq \mathbf{E}X$.

Conversely, if $J_+ = \infty$ then $\limsup_{n\to\infty} S_n/n = \infty$ a.s. by (1.7), and hence $T_1^*(a) < \infty$ a.s. for all $a > 0$. If $0 < \mathbf{E}X \leq \mathbf{E}|X| < \infty$ then $T_1^*(a) < \infty$ a.s. for $a = \mathbf{E}X$, hence for $0 \leq a \leq \mathbf{E}X$, as just shown above. □

ACKNOWLEDGEMENTS. Much of this work was done while the second author was visiting the Mathematics Department of Cornell University. He is grateful for its hospitality. The research of the first author and travel of the second author was supported in part by the NSF through a grant to Cornell University.

REFERENCES

[1] CHOW, Y. S. and TEICHER, H., *Probability theory. Independence, interchangeability, martingales*, 2nd ed., Springer Texts in Statistics, Springer-Verlag, New York – Berlin, 1988. *MR* **89e**:60001

[2] CHOW, Y. S and ZHANG, C-H., A note on Feller's strong law of large numbers, *Ann. Probab.* **14** (1986), 1088–1094. *MR* **87j**:60047

[3] 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

[4] CHUNG, K. L. and FUCHS, W. H. J., On the distribution of values of sums of random variables, *Mem. Amer. Math. Soc.* (1951), no. 6, 12 pp. *MR* **12**, 722e

[5] ERICKSON, K. B., On the strong law of large numbers when the mean is undefined, *Trans. Amer. Math. Soc.* **185** (1973), 371–381. *MR* **49** #1579

[6] ERICKSON, K. B., Recurrence sets of normed random walk in R^d, *Ann. Probab.* **4** (1976), 802–828. *MR* **54** #14108

[7] ESSEEN, C. G., On the concentration function of a sum of independent random variables, *Z. Wahrsch. Verw. Gebiete* **9** (1968), 290–308. *MR* **37** #6974

[8] GRIFFIN, P. S. and MCCONNELL, T. R., L^p-boundedness of the overshoot in multidimensional renewal theory, *Ann. Probab.* **23** (1995), 2022–2056. *MR* **97d**:60119

[9] GUT, A., *Stopped random walks. Limit theorems and applications*, Applied Probability. A Series of the Applied Probability Trust, 5, Springer-Verlag, New York – Berlin, 1988. *MR* **88m**:60085

[10] HAHN, M. G. and KLASS, M. J., Approximation of partial sums of arbitrary i.i.d. random variables and the precision of the usual exponential upper bound, *Ann. Probab.* **25** (1997), 1451–1470.

[11] KESTEN, H., The limit points of a normalized random walk, *Ann. Math. Statist.* **41** (1970), 1173–1205. *MR* **42** #1222

[12] KESTEN, H., A universal form of the Chung-type law of the iterated logarithm, *Ann. Probab.* **25** (1997), 1588–1620.

[13] KESTEN, H. and MALLER, R. A., Infinite limits and infinite limit points of random walks and trimmed sums, *Ann. Probab.* **22** (1994), 1473–1513. *MR* **95m**:60055

[14] KESTEN, H. and MALLER, R. A., Two renewal theorems for general random walks tending to infinity, *Probab. Theory Related Fields* **106** (1996), 1–38. *MR* **97i**:60113

[15] KESTEN, H. and MALLER, R. A., Divergence of a random walk through deterministic and random subsequences, *J. Theoret. Probab.* **10** (1997), 395–427.
[16] KESTEN, H. and MALLER, R. A., Random walks crossing high level curved boundaries, *J. Theoret. Probab.* (1998).
[17] KLASS, M. J., Properties of optimal extended-valued stopping rules for S_n/n, *Ann. Probab.* **1** (1973), 719–757. *MR* **54** #14086
[18] KLASS, M. J., Toward a universal law of the iterated logarithm. I, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **36** (1976), 165–178. *MR* **54** #3822
[19] KLASS, M. J., Toward a universal law of the iterated logarithm. II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **39** (1977), 151–165. *MR* **56** #6808
[20] KLASS, M. J., The finite mean LIL bounds are sharp, *Ann. Probab.* **12** (1984), 907–911. *MR* **85j**:60048
[21] PRUITT, W. E., General one-sided laws of the iterated logarithm, *Ann. Probab.* **9** (1981), 1–48. *MR* **82k**:60066
[22] SIEGMUND, D., *Sequential analysis. Tests and confidence intervals*, Springer Series in Statistics, Springer-Verlag, New York – Berlin, 1985. *MR* **87h**:62145
[23] SPITZER, F., *Principles of random walks*, 2nd ed., Graduate Texts in Mathematics, Vol. 34, Springer-Verlag, New York, 1976. *MR* **52** #9383
[24] STONE, C., The growth of a recurrent random walk, *Ann. Math. Statist.* **37** (1966), 1040–1041. *MR* **33** #5004
[25] ZHANG, C-H., The lower limit of a normalized random walk, *Ann. Probab.* **14** (1986), 560–581. *MR* **87h**:60131

(Received May 14, 1997)

DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NY 14853-7901
U.S.A.
kesten@math.cornell.edu

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF WESTERN AUSTRALIA
NEDLANDS 6907
WESTERN AUSTRALIA
maller@maths.uwa.edu.au

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 253–259

ERDŐS–RÉNYI–SHEPP LAWS FOR DEPENDENT RANDOM VARIABLES

R. KIESEL and U. STADTMÜLLER

To the memory of Alfréd Rényi

Abstract

We prove an Erdős–Rényi–Shepp law for the partial sums of a uniform strong mixing stationary sequence.

1. Introduction and main results

While there is a large amount of literature on versions of the Erdős–Rényi–Shepp law for sequences of independent, identically distributed (i.i.d.) random variables, see e.g. [11, 3, 4, 5, 6, 7, 12, 13], not much is known for dependent random variables (see [9] for a first result in this direction). Using a recent large deviation result by Bryc [1] we proceed to a more general setting.

Let $\{X_n\}$ be a stationary sequence. We define $\mathcal{F}_n^m := \sigma(X_k : n \leqq k \leqq m)$ the canonical σ-algebra generated by $X_n, \ldots, X_m$ and the ϕ-mixing coefficient

$$\phi_n := \sup_{k \geqq 1} \sup \left\{ |\mathbf{P}(B|A) - \mathbf{P}(B)| : A \in \mathcal{F}_1^k, \mathbf{P}(A) \neq 0; B \in \mathcal{F}_{k+n}^\infty \right\}.$$

We say that a sequence $\{X_n\}$ is ϕ-mixing if $\phi_n \to 0$ for $n \to \infty$.

We shall need the following hypergeometric rate of convergence

$$e^{Kn}\phi_n \to 0 \quad (n \to \infty) \quad \text{for each } K \geqq 0. \tag{1.1}$$

We have the following large deviation theorem by Bryc [1].

THEOREM B. *Let $\{X_n\}$ be a stationary ϕ-mixing sequence of random variables such that $|X_1| \leqq C < \infty$ and* (1.1) *holds. Define $Z_n = (X_1 + \ldots + X_n)/n$, $n \geqq 1$. Then the limit*

$$\lim_{n \to \infty} n^{-1} \log \mathbf{E}\{\exp(n\lambda Z_n)\} = L(\lambda)$$

exists for each $\lambda \in \mathbb{R}$ and the function $I : \mathbb{R} \to [0, \infty]$ defined by

$$I(x) := \sup_\lambda \{x\lambda - L(\lambda)\}$$

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60G50.
Key words and phrases. Erdős–Rényi laws, mixing random variables.

is a convex, continuous function. For $\{Z_n\}$ the large deviation principle holds true with the good rate function I, that is,

$$\limsup_{n\to\infty} n^{-1}\log \mathbf{P}\,(Z_n\in F) \leqq -\inf_{x\in F} I(x) \tag{1.2}$$

for each closed set $F\subseteqq \mathbb{R}$, and

$$\liminf_{n\to\infty} n^{-1}\log \mathbf{P}\,(Z_n\in G) \geqq -\inf_{x\in G} I(x) \tag{1.3}$$

for each open set $G\subseteqq \mathbb{R}$.

LEMMA. *Let $\{X_n\}$ be a stationary ϕ-mixing sequence of random variables, such that $|X_1|\leqq C<\infty$, $\mathbf{E}(X_1)=0$ but $\mathbf{E}(X_1^2)>0$ and (1.1) holds. Then we have with $x_0:=\sup\{x\geqq 0;\ I(x)=0\}$, $x_1:=\sup\{x>0;\ I(x)<\infty\}$ and $A:=\sup_{0<x<x_1}\{I(x)\}$ (where $\sup_{\emptyset}\{I(.)\}:=-\infty$) that*

(i) *$L(.)$ is convex and hence continuous on $\mathbb{R}$ and $L(\lambda)\leqq C\lambda$, $\lambda\geqq 0$;*

(ii) *$0\leqq x_0\leqq x_1\leqq C$;*

(iii) *If $x_0<x_1$, $I:[x_0,x_1)\to[0,A)$ is continuous and strictly increasing and hence $I^{\leftarrow}:[0,A)\to[x_0,x_1)$ exists.*

CONVENTION. To obtain our main result in a closed form we define

$$I^{\leftarrow}(x):=x_1 \quad \text{if } x\geqq \max\{0,A\}.$$

Using (1.2) resp. (1.3) for $F=[a,\infty)$ resp. $G=(a,\infty)$ with any $a\in(x_0,x_1)$ we obtain by the strict monotonicity of I that for any sufficiently small $\varepsilon>0$ and n sufficiently large

$$\mathbf{P}(Z_n\geqq a)\leqq \exp\{-nI(a-\varepsilon)\} \tag{1.4}$$

and

$$\mathbf{P}(Z_n> a)\geqq \exp\{-nI(a+\varepsilon)\}. \tag{1.5}$$

We consider in the sequel the following random variables

$$V_n:=\max_{0\leq k\leq n-b_n}\frac{S_{k+b_n}-S_k}{b_n} \tag{1.6}$$

with $S_n:=X_1+\dots+X_n$ and $b_n:=[c\log n]$ for $c>0$.

We can now state our main result, in which $A:=\sup_{0<x<x_1}\{I(x)\}$ as above.

THEOREM. *Let $\{X_n\}$ be a stationary, ϕ-mixing sequence of random variables, such that $|X_1|\leqq C<\infty$, $\mathbf{E}(X_1)=0$ but $\mathbf{E}(X_1^2)>0$ and (1.1) holds. Then we have for any $c>0$*

$$\lim_{n\to\infty} V_n = I^{\leftarrow}(1/c)\ \textit{a.s.}.$$

2. Proofs

PROOF OF THE LEMMA. It is obvious that $\psi_n(\lambda) := n^{-1} \log\{\mathbf{E}(e^{n\lambda Z_n})\} \leqq \lambda C$ for nonnegative arguments and exists for all λ since $-C \leqq Z_n \leqq C$. As a limit of convex functions (cumulant generating functions) $L(.)$ is convex and hence continuous. Next we shall show that $I(x) = \infty$ for all $x > C$. This follows from the fact that (use $\mathbf{E}(X_1) = 0$)

$$I(x) = \sup_{\lambda \geqq 0}\{\lambda x - L(\lambda)\} \geqq \sup_{\lambda \geqq 0}\{\lambda x - \lambda C\} = \infty \quad \text{for } x > C.$$

Hence $x_1 \leqq C$. By definition $I(.)$ is strictly positive for $x > x_0$. By Lemma 2.2.5 in [8] it is a good rate function and is hence convex and lower semicontinuous. If $x_0 < x_1$ then it is continuous in (x_0, x_1) and we show that $I(.)$ is strictly increasing on $[x_0, x_1)$. By general arguments $I(x)$ is nondecreasing for $x \geqq 0$ (see e.g. [8], pp. 28 or [10], p. 4). We show now that $I(.)$ is increasing, i.e., we have $I(x_2) < I(x_3)$ if $x_0 \leqq x_2 < x_3 < x_1$. By the last step we can restrict ourselves to prove the case $x_2 > x_0$. There exists a sequence (λ_n) such that $I(x_2) = \lim_{n\to\infty} (\lambda_n x_2 - L(\lambda_n))$ and $\lambda_n \geqq \delta$ with some $\delta \in (0,1)$, since otherwise $I(x_2) = 0$ in contradiction to the last step. For $0 < \varepsilon < (x_3 - x_2)\delta$ we have for n large enough

$$\begin{aligned} I(x_2) &\leqq \lambda_n x_2 - L(\lambda_n) + \varepsilon \\ &< \lambda_n\left(x_3 - \frac{\varepsilon}{\delta}\right) - L(\lambda_n) + \varepsilon \\ &\leqq \lambda_n x_3 - L(\lambda_n) \leqq I(x_3), \end{aligned}$$

giving the Lemma. □

PROOF OF THE THEOREM. Let us begin with the cases $A > -\infty$, i.e. $A > 0$ and $c > 1/A$.

(a) Our first claim is

$$\limsup_{n\to\infty} V_n \leqq I^{\leftarrow}(1/c) \ \text{a.s..} \tag{2.1}$$

Choose $\varepsilon > 0$ and set $x := I^{\leftarrow}(1/c) + 2\varepsilon < x_1$. Define for $n \geqq 1$ the events

$$A_n := \{\omega : V_n \geqq x\}.$$

We have to show that

$$\mathbf{P}(A_n \ \text{i.o.}) = 0. \tag{2.2}$$

Hence we estimate $\mathbf{P}(A_n)$. Using successively the sub-additivity of the probability measure, the stationarity of the sequence $\{X_n\}$ and (1.4) we obtain for n large enough

$$\mathbf{P}(A_n) = \mathbf{P}\left(\bigcup_{k=0}^{n-b_n}\left\{\omega : \frac{S_{k+b_n} - S_k}{b_n} \geqq x\right\}\right)$$

$$\leqq n\mathbf{P}\left(\frac{S_{b_n}}{b_n}\geqq x\right)$$
$$\leqq n\exp\{-b_nI(x-\varepsilon)\}=n\exp\{-b_nI(I^{\leftarrow}(1/c)+\varepsilon)\}$$
$$\leqq n\exp\{-(1+\delta)\log n\}$$

for some $\delta>0$ by the Lemma.

Now choose T the smallest integer such that $T\delta>1$; then

$$\sum_{n=1}^{\infty}\mathbf{P}(A_{n^T})<\infty. \tag{2.3}$$

Hence $\mathbf{P}(A_{n^T}\text{ i.o.})=0$. Setting for $n\geqq 1$

$$A_n^*:=\left\{\omega:\max_{0\leq k\leq n-b_n}\frac{S_{k+b_n-1}-S_k}{b_n-1}\geqq x\right\}$$

similarly we see that

$$\sum_{n=1}^{\infty}\mathbf{P}(A_{n^T}^*)<\infty. \tag{2.4}$$

Now for n large enough we have $b_{(n+1)^T}-b_{n^T}\leqq 1$ and using (2.3) and (2.4) this implies that (2.2) holds true (see [2], p. 100 for this type of argument).

(b) Our second claim is

$$\liminf_{n\to\infty}V_n\geqq I^{\leftarrow}(1/c)\text{ a.s.}. \tag{2.5}$$

Choose $\varepsilon>0$ so that $x:=I^{\leftarrow}(1/c)-2\varepsilon>x_0$. Define for $n\geqq 1$ the events

$$B_n:=\{\omega:V_n\leqq x\}.$$

We have to show that $\mathbf{P}(B_n\text{ i.o.})=0$. By the Borel–Cantelli lemma it suffices to show

$$\sum_{n=1}^{\infty}\mathbf{P}(B_n)<\infty. \tag{2.6}$$

Let (d_n) be any sequence of positive integers with $d_n\to\infty$ $(n\to\infty)$. Then we have

$$\mathbf{P}(B_n)=\mathbf{P}\left(\bigcap_{k=0}^{n-b_n}\left\{\omega:\frac{S_{k+b_n}-S_k}{b_n}\leqq x\right\}\right)$$
$$\leqq\mathbf{P}\left(\bigcap_{i=1}^{\left[\frac{n-b_n}{b_n+d_n}\right]-1}\left\{\omega:\frac{S_{i(b_n+d_n)}-S_{i(b_n+d_n)-b_n}}{b_n}\leqq x\right\}\right).$$

For $i=1,\dots,\left[\frac{n-b_n}{b_n+d_n}\right]-1$ we define the events

$$E_i := \left\{\omega : \frac{S_{i(b_n+d_n)} - S_{i(b_n+d_n)-b_n}}{b_n} \leqq x\right\}.$$

Using standard techniques for sequences of ϕ-mixing, stationary random variables we get for any positive integer N

$$\begin{aligned}
\mathbf{P}\left(\bigcap_{i=1}^{N} E_i\right) &= \mathbf{P}\left(\bigcap_{i=1}^{N-1} E_i E_N\right) \\
&\leqq (\phi_{d_n} + \mathbf{P}(E_N))\,\mathbf{P}\left(\bigcap_{i=1}^{N-1} E_i\right) \\
&= (\phi_{d_n} + \mathbf{P}(E_1))\,\mathbf{P}\left(\bigcap_{i=1}^{N-1} E_i\right) \\
&\leqq (\phi_{d_n} + \mathbf{P}(E_1))^{N-1}\,\mathbf{P}(E_1).
\end{aligned}$$

Using our large deviation estimate (1.5) and stationarity we get for n large enough

$$\begin{aligned}
\mathbf{P}(E_1) = \mathbf{P}\left(\frac{S_{b_n}}{b_n} \leqq x\right) &= 1 - \mathbf{P}\left(\frac{S_{b_n}}{b_n} > x\right) \\
&\leqq 1 - \exp\{-b_n I(x+\varepsilon)\} = 1 - \exp\{-b_n I(I^{\leftarrow}(1/c) - \varepsilon)\} \\
&\leqq 1 - \exp\{-b_n((1-\delta)/c)\} = 1 - \exp\{-(1-\delta)\log n\}
\end{aligned}$$

for some $\delta > 0$, by the Lemma. Combining the above estimates we get, with $d_n := [\log n]$,

$$\begin{aligned}
\mathbf{P}(B_n) &\leqq (1 - (\exp\{-(1-\delta)\log n\} - \phi_{d_n}))^{\left[\frac{n-b_n}{(c+1)\log n}\right]-2} \\
&= \exp\left\{\left(\left[\frac{n-b_n}{(c+1)\log n}\right] - 2\right)\log(1 - (\exp\{-(1-\delta)\log n\} - \phi_{d_n}))\right\} \\
&\leqq \exp\left\{-\left(\left[\frac{n-b_n}{(c+1)\log n}\right] - 2\right)\left(n^{-(1-\delta)} - \phi_{d_n}\right)\right\} \\
&\leqq \exp\left\{-\frac{1}{2}\left(\left[\frac{n-b_n}{(c+1)\log n}\right] - 2\right)n^{-(1-\delta)}\right\} \leqq \exp\left\{-\bar{c}n^{\delta}/\log n\right\},
\end{aligned}$$

where $\bar{c}$ is a positive constant. Observe that we used the inequality $\log(1-x) \leqq -x$ (for $x<1$) and then the hypergeometric mixing rate in the next to last inequality.

By the above estimate we get (2.6) and hence claim (b) is true.

Combining claims (a) and (b) the proof for the main case is complete.

In the remaining cases ($A=-\infty$, i.e., $x_0=x_1$ or $A>0$ but $0<c\leqq 1/A$) we have for any $x>x_1=I^{\leftarrow}(1/c)$ that $I(x)=\infty$ and hence we get the following large deviation inequality. For any $M>0$ there exists some $n_0\in\mathbb{N}$ such that for $n\geqq n_0$

$$\mathbf{P}(Z_n\geqq x)\leqq\exp\{-nM\}$$

holds. Similar arguments as above lead to the upper bound

$$\limsup_{n\to\infty} V_n\leqq x_1.$$

If $x_1=0$ we are done, since replacing X by $-X$ leads to $\lim\limits_{n\to\infty} V_n=0$ a.s.. If $0<x_0=x_1$ then for any $x<x_1$ we have $I(x)=0$ and hence for any $\varepsilon>0$ we obtain for n large enough

$$\mathbf{P}(Z_n\geqq x)\geqq\exp\{-n\varepsilon\}$$

and therefrom we can deduce the lower bound (2.5) as above. Finally if $x_0<x_1$ and $x<x_1$ then we have $cI(x)<cA\leqq 1$ since $c\leqq 1/A$ now. Again our reasoning above goes through, giving the desired result. □

REFERENCES

[1] Bryc, W., On large deviations for uniformly strong mixing sequences, *Stochastic Process. Appl.* **41** (1992), 191–202. *MR* **93h**:60037

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

[3] Csörgő, M. and Shao, Q-M., A self-normalized Erdős–Rényi type strong law of large numbers, *Stochastic Process. Appl.* **50** (1994), 187–196. *MR* **95d**: 60058

[4] 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

[5] de Acosta, A. and Kuelbs, J., Limit theorems for moving averages of independent random vectors, *Z. Wahrsch. Verw. Gebiete* **64** (1983), 67–123. *MR* **84i**:60011

[6] 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

[7] Deheuvels, P. and Steinebach, J., Exact convergence rates in strong approximation laws for large increments of partial sums, *Probab. Theory Related Fields* **76** (1987), 369–393. *MR* **88m**:60081

[8] Dembo, A. and Zeitouni, O., *Large deviations techniques and applications*, Jones and Bartlett Publishers, Boston, MA, 1993. *MR* **95a**:60034

[9] Deo, C.M., A note on Erdős–Rényi law of large numbers, *Canad. Math. Bull.* **19** (1976), 497–500. *MR* **56** #3924

[10] Deuschel, J-D. and Stroock, D. W., *Large deviations*, Academic Press, Boston, 1989. *MR* **90h**:60026

[11] Erdős, P. and Rényi, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111. *MR* **42** #6907

[12] KIESEL, R. and STADTMÜLLER, U., Erdős–Rényi–Shepp laws and weighted sums of independent identically distributed random variables, *J. Theoret. Probab.* **9** (1996), 961–982. *MR* **97m**:60035

[13] SHEPP, L.A., A limit law concerning moving averages, *Ann. Math. Statist.* **35** (1964), 424–428. *MR* **29** #4091

(Received December 12, 1996)

DEPARTMENT OF STATISTICS
BIRKBECK COLLEGE
UNIVERSITY OF LONDON
MALET STREET
LONDON WC1E 7HX
ENGLAND

r.kiesel@statistics.bbk.ac.uk

and

ABTEILUNG MATHEMATIK III
UNIVERSITÄT ULM
D–89069 ULM
GERMANY

stamue@mathematik.uni-ulm.de

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 261–271

POINTWISE BOUNDED APPROXIMATION

J. KOMLÓS and G. TUSNÁDY

To the memory of A. Rényi

Abstract

We show that a sequence of partial sums of i.i.d. random variables can be approximated by a sequence of normally distributed random variables in such a way that the difference is finite almost surely.

1. Introduction

V. M. Zolotarev posed the following peculiar question. Let S_n be a sequence of partial sums of i.i.d. random variables. Is it possible to approximate S_n by a sequence T_n of normally distributed random variables in such a way that $\sup_n |S_n - T_n|$ is finite almost surely ?

Observe that here, unlike in standard embedding questions, it is not assumed that T_n are partial sums, or, for that matter, anything about the joint distributions of the random variables T_n (they do not even have to be joint normal).

This was a question that grew out from his work on embeddings using higher order terms in the Cornish–Fisher expansion $(S_n - an)/(\sigma\sqrt{n}) = \sum_{i=0}^{k} n^{-i/2} p_k(N_n) + \varepsilon_n n^{-k/2}$, where a and σ^2 are the mean and variance of the terms in S_n, N_n are standard normal, and p_k are polynomials. (He has the following result – see in [11]: If the terms in S_n have $r > 4$ moments and satisfy the Cramér condition, then in the above mentioned Cornish–Fisher expansion with $k = \lceil r - 4 \rceil - 1$ one has $\varepsilon_n \to 0$ a.s.. Our Lemma 2 below is a particular instance of his theorem.)

In this paper we give an affirmative answer to the above question. In fact, we will construct a pointwise bounded approximation for a quite general class of random sequences S_n, T_n.

1991 *Mathematics Subject Classification.* Primary 60F15, 60G50; Secondary 05D15.
Key words and phrases. Coupling, partial sums.

Research supported by the Hungarian National Foundation for Scientific Research Grant #1905.

0081–6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

Let us first explain why we called the question peculiar. The central limit theorem easily implies that the partial sums S_n can be approximated by normal variables T_n in such a way that the difference $S_n - T_n$ is stochastically bounded uniformly in n (that is, for every $\varepsilon > 0$ there is a K such that $\mathbf{P}(|S_n - T_n| > K) < \varepsilon$ for all n). In this approximation, only the individual distributions of S_n and T_n appear, and it is not clear at first sight whether the joint distributions of S_n matter or not in Zolotarev's question. If the joint distributions of the S_n are given then simple-minded applications of the central limit theorem are doomed, for the above mentioned uniform stochastic boundedness alone is not sufficient to guarantee pointwise boundedness. An embedding of the pair (S_n, T_n) is a joint distribution of the variables (S_n, T_n) with the prescribed marginals. This joint distribution determines the conditional distribution of T_n given S_n. One is tempted to use the conditional distribution of T_n given S_n independently for different n, but it will not work, for independent errors add up. The crucial point in our construction is that we use the same randomization (that is, a kind of mixture) for different n. It turns out that this embedding works regardless how we specify the joint distributions of the sequence S_n.

Usually the joint distributions of both S_n and T_n are fitted (see in Bretagnolle and Massart [2], Csörgő and Hall [4], Csörgő and Révész [5], Komlós, Major and Tusnády [7], Major [8], Tusnády [10]). In this situation a bounded error embedding is certainly impossible since the error terms tend to infinity pointwise according to the theorem of Bártfai [1].

In the following theorem S_n and T_n are general sequences of random variables.

THEOREM 1. *Let $\varepsilon(x)$ be a monotone decreasing positive function with $\lim_{x\to\infty} \varepsilon(x) = 0$, and Δ_n a sequence of positive numbers with $\lim_{n\to\infty} \Delta_n = \infty$. If the (marginal) distributions of the random variables S_n, T_n satisfy the following condition:*

$$\mathbf{P}(a < S_n < b) \leqq (1 + \varepsilon(b-a))\mathbf{P}(a < T_n < b) \quad \text{for all } a, b \text{ with } 1 \leqq b - a \leqq \Delta_n, \tag{1}$$

then, given arbitrary joint distributions for the S_n, one can construct (on some probability space) sequences S_n, T_n with the prescribed marginal and joint distributions such that

$$\sup_n |S_n - T_n| < \infty \quad \text{almost surely.} \tag{2}$$

REMARK. Note that the condition (1) is equivalent to the existence of an $\varepsilon(x)$ as above and a non-negative sequence δ_n with $\lim_{n\to\infty} \delta_n = 0$ such that

$$\mathbf{P}(a < S_n < b) \leqq (1 + \varepsilon_n(b-a))\mathbf{P}(a < T_n < b) \quad \text{for all } a, b \text{ with } b - a \geqq 1, \tag{3}$$

where $\varepsilon_n(x)$ is the truncated sequence $\varepsilon_n(x) = \max\{\varepsilon(x), \delta_n\}$. (Indeed, choose $\delta_n = \varepsilon(\Delta_n/2)$, and partition the interval (a,b) into intervals of lengths between $\Delta_n/2$ and Δ_n.)

THEOREM 2. *Let* $S_n = \sum_{k=1}^{n} X_k$, *where* X_k *are i.i.d. continuous random variables with*

$$\mathbf{E}X_1 = 0, \quad \mathbf{E}X_1^2 = 1, \quad \mathbf{E}|X_1|^r < \infty \tag{4}$$

for some $r \geqq 4$. *If they satisfy the Cramér condition*

$$\limsup_{|t|\to\infty} |\mathbf{E}e^{itX_1}| < 1 \tag{5}$$

then one can construct (on some probability space) two sequences S_n', T_n *such that the joint distributions of* S_n' *are the same as those of* S_n, T_n *are normal with* $\mathbf{E}T_n = 0$, $\mathbf{E}T_n^2 = n$, *and*

$$\sup_n |S_n' - T_n| < \infty \quad \textit{almost surely.}$$

The strategy of our proof is the following. The natural candidate, the quantile transform, does not work for the construction of a pair (S_n', T_n) which satisfies Theorem 2. But we show in Lemma 2 with the help of the Cornish–Fisher expansion that this condition works if S_n is approximated by $T_n + \kappa\left(\frac{T_n^2}{n} - 1\right)$ with an appropriate constant κ. Theorem 1 enables us to replace this approximating sequence by the sequence T_n. The main idea in the proof of Theorem 1 is to construct, for a fixed sequence S_n, random variables T_n and U for which $|S_n - T_n| \leqq U$, and the variable U is *independent* of the sequence S_n. This independence enables us to ensure that the sets where $|S_n - T_n|$ is large are contained in a set which does not depend on the index n.

2. Proof of Theorem 1

2.1. A matching lemma

LEMMA 1. *Given random variables* X, Y, Z *such that* $X \leqq Y$, *and for all* a, b,

$$\mathbf{P}(a < X \leqq Y < b) \leqq \mathbf{P}(a < Z < b). \tag{6}$$

Then there are random variables X', Y', Z' *(on some probability space) such that the joint distribution of* X', Y' *is the same as that of* X, Y, *the distribution of* Z' *is the same as that of* Z, *and* $X' \leqq Z' \leqq Y'$.

Note that the above sufficient condition is also necessary.

PROOF. Strassen's [9] marriage lemma gives the following condition for the existence of such an embedding. For an arbitrary two-dimensional open set $V \subset \{(x,y): x \leqq y\}$,

$$\mathbf{P}((X,Y) \in V) \leqq \mathbf{P}(Z \in W), \tag{7}$$

where $W = W(V) = \{w : a \leqq w \leqq b \text{ for some } (a,b) \in V\}$.

W is open, and thus it is a countable union of disjoint open intervals. If W is one single open interval, say $W = (a,b)$, then V is a subset of $\{(x,y): a < x \leqq y < b\}$, and thus (6) implies (7). □

2.2. The construction

The main idea is to generate a positive random variable U controlling the size of $|S_n - T_n|$. In fact, we will construct T_n as a *mixture* of $T_n(U)$, such that the difference $|S_n - T_n(U)|$ is bounded by U for large n. As we will see it is to our advantage to choose U to be independent of the whole process S_n.

Given S_n (on a large enough probability space), we first truncate S_n at a (very high) level M_n by choosing a number x_n such that $\mathbf{P}(x_n - 1/2 < T_n \leqq x_n) > 0$ and $\mathbf{P}(x_n \leqq T_n < x_n + 1/2) > 0$, and then defining

$$S'_n = \begin{cases} S_n & \text{if } |S_n| \leqq M_n \\ x_n & \text{otherwise.} \end{cases}$$

We assume that the alternative condition (3) is satisfied, replace the function $\varepsilon(x)$ with a new function $\eta(x)$, and set $\eta_n(x) = \max\{\eta(x), \delta_n\}$. M_n and $\eta(x)$ will be chosen to satisfy the following conditions.

(A) $\sum_n \mathbf{P}(|S_n| > M_n) < \infty$.

(B) $\mathbf{P}(|S_n| > M_n) \leqq \delta_n \mathbf{P}(a < T_n < a+1)$ whenever a is such that $a < x_n < a+1$.

(C) $\eta(4x) \geqq \mathbf{P}(|T_n| \geqq x)$ for all $x > M_n$.

(D) $\eta(x) \geqq 2\varepsilon(x)$ and $\eta(4M_n) \geqq \delta_n$.

Now we choose a random variable $U \geqq u_0$, independent of the whole sequence S_n, with

$$\mathbf{P}(U \geqq u) \geqq \eta(2u) \quad \text{for } u \geqq u_0,$$

where u_0 is such that $u_0 \geqq 1$ and $\eta(2u_0) \leqq 1$.

We will prove (under the conditions of Theorem 1) the following inequality:

$$\mathbf{P}(a < S'_n - U \leqq S'_n + U < b) \leqq \mathbf{P}(a < T_n < b) \quad \text{for all } a, b. \tag{8}$$

Thus, condition (6) is met for

$$X = S'_n - U, \quad Y = S'_n + U, \quad Z = T_n.$$

Hence, by Lemma 1, we can put T_n in between $S'_n - U$ and $S'_n + U$. That is, $|S'_n - T_n| \leqq U$, and (2) is proven, since $S'_n = S_n$ for large n almost surely.

Lemma 1 only gives the three-dimensional distributions $F_n(s,u,t)$ of (S'_n, U, T_n) separately for each n. In applying it in our situation, we use further independent randomizations to generate the sequence T_n, using the conditional distributions $F_n(t|s,u)$ obtained from $F_n(s,u,t)$, by assuming that T_n is conditionally independent of the sequence $[S_m : m \neq n]$ under fixed S'_n, U. □

2.3. Proof of (8)

If $b-a<1$ then the left-hand side of (8) is 0, so we may assume $b-a \geqq 1$. For the same reason, we may also assume that the interval (a,b) intersects the interval $(-M_n, M_n)$. We start with the obvious inequality

$$\mathbf{P}(a < S'_n - U \leqq S'_n + U < b) \leqq \mathbf{P}(a < S'_n < b)\,\mathbf{P}(U < m),$$

where $m = \min\{(b-a)/2,\ M_n - a,\ M_n + b\}$. We distinguish four cases: (i) $b-a \leqq 4M_n$, (ii) $b-a > 4M_n$ and $a \geqq -M_n$, (iii) $b-a > 4M_n$ and $b \leqq M_n$, (iv) $a < -M_n$ and $b > M_n$. In case (i), we have

$$\mathbf{P}(U < m) \leqq \mathbf{P}(U < (b-a)/2) \leqq 1 - \eta(b-a) = 1 - \eta_n(b-a).$$

In case (ii),

$$\begin{aligned}\mathbf{P}(U < m) &\leqq \mathbf{P}(U < M_n - a) \leqq \mathbf{P}(U < 2M_n)\\ &\leqq 1 - \eta(4M_n) = 1 - \eta_n(4M_n) \leqq 1 - \eta_n(b-a).\end{aligned}$$

Case (iii) is similar.

Conditions (B) and (D), together with (3) imply

$$\begin{aligned}\mathbf{P}(a < S'_n < b) &\leqq \mathbf{P}(a < S_n < b) + \mathbf{P}(|S_n| > M_n, \text{ and } a < x_n < b)\\ &\leqq \mathbf{P}(a < S_n < b) + \delta_n \mathbf{P}(a < T_n < b)\\ &\leqq (1 + \varepsilon_n(b-a) + \delta_n)\mathbf{P}(a < T_n < b)\\ &\leqq (1 + \eta_n(b-a))\mathbf{P}(a < T_n < b).\end{aligned}$$

That is, (3) holds with S'_n replacing S_n and η_n replacing ε_n. Using it, we get, in the first three of the above four cases,

$$\mathbf{P}(a < S'_n - U \leqq S'_n + U < b) \leqq \mathbf{P}(a < S'_n < b)(1 - \eta_n(b-a)) \leqq \mathbf{P}(a < T_n < b)$$

as required. It remains to consider case (iv). Let $c = \min\{|a|, b\}$. Then, by (C),

$$\mathbf{P}(U < m) \leqq \mathbf{P}(U < 2c) \leqq 1 - \eta(4c) \leqq \mathbf{P}(|T_n| < c) \leqq \mathbf{P}(a < T_n < b).$$

2.4. Conditions A, B, C, D

(A) and (B) can be ensured by simply choosing M_n large enough. To satisfy (C), we may apply the following fact with $f_n(x) = \mathbf{P}(|T_n| \geqq x)$, the sequence a_n equal to M_n already chosen according to (A) and (B), $g(x) = \eta(4x)$, and b_n being the new choice for M_n.

FACT. Let $f_n(x)$ be monotone decreasing functions tending to 0 as $x \to \infty$, and a_n be any sequence. Then there is a sequence b_n and a function $g(x)$ such that $b_n \geqq a_n$, $\lim g(x) = 0$ as $x \to \infty$, and $g(x) \geqq f_n(x)$ for all $x \geqq b_n$.

Finally, to satisfy (D) it is enough to further increase η.

3. Proof of Theorem 2

3.1. Cornish–Fisher expansion

The Edgeworth expansion is an approximation of the distribution function of partial sums. When working with embeddings, one translates these approximations to random variables along the lines of the expansions of Cornish and Fisher [3].

LEMMA 2. *Let S_n be as in Theorem 2. Assume also that the probability space is sufficiently rich, and let T_n be defined by the quantile transformation:*

$$F_n(S_n) = G_n(T_n), \tag{9}$$

where $F_n(x)$ and $G_n(x) = \Phi(x/\sqrt{n})$ are the distribution functions of S_n and T_n. Then

$$S_n - \left(T_n + \kappa\left(\frac{T_n^2}{n} - 1\right)\right) \to 0 \quad \textit{almost surely,}$$

where $\kappa = (1/6)\mathbf{E}X_1^3$.

PROOF. This lemma is a particular case of the result of Zolotarev mentioned in the introduction, but since his results were reported without proofs, we give a proof of Lemma 2 here for the sake of completeness.

We will use the Edgeworth expansion (Theorem 3 in XVI.4 in Feller [6]): Conditions (4), (5) with integer $r \geqq 3$ imply that, as $n \to \infty$,

$$F_n(x\sqrt{n}) - \Phi(x) - \varphi(x)\sum_{k=3}^{r} n^{1-k/2}R_k(x) = o(n^{1-r/2}) \tag{10}$$

uniformly in x. Here R_k is a polynomial depending only on the moments $\mathbf{E}X_1^j$, $j = 3, 4, ..., k$, but not on n and r (or otherwise on the distribution of X_1); in particular $R_3(x) = \kappa(1 - x^2)$.

We apply the following technical lemma.

LEMMA 3. *There exist positive constants c, d such that the following estimate holds for any x and y such that $|y| \leqq c/(1+|x|)$:*

$$\Phi(x+y-z) \leqq \Phi(x) + y\varphi(x) \leqq \Phi(x+y+z),$$

where $z = d(1+|x|)y^2$.

PROOF. We show that the choice $c = 1/4$, $d = 2$ is appropriate. Let us apply the Taylor formula

$$\Phi(x+y+u) - [\Phi(x) + \varphi(x)y] = \varphi(x)u + \varphi'(\xi)(y+u)^2/2,$$

where ξ is between x and $x+y+u$. If u is such that here in the right side the modulus of the first term is larger than that of the second one, then the sign of the right-hand side is dictated by the sign of u. Thus it is enough to show that this is the case with $|u| = z$. On applying $\varphi'(\xi) = -\xi\varphi(\xi)$, and $z \leqq cd|y| = |y|/2$, it is enough to show that

$$d(1+|x|)y^2 \geqq |\xi|\, e^{(x^2-\xi^2)/2}\, (3|y|/2)^2/2.$$

Here y^2 cancels out, the term $|\xi|$ is less than $1+|x|$, and the exponent $(x^2-\xi^2)/2$ is easily seen to be less than $3/8$. □

Let us denote the left-hand side of (10) by C, and let $t<1$. Then, for some $\varepsilon > 0$,

$$\left|\frac{C}{\varphi(x)}\right| = o\left(n^{-1/2}/\varphi(x)\right) = o(n^{-\varepsilon}) = o\left(\frac{1}{1+|x|}\right)$$

uniformly for $|x| \leqq t\sqrt{\log n}$ as $n \to \infty$. Similarly, for

$$B = B(x,n) = \sum_{k=3}^{r} n^{1-k/2} R_k(x)$$

we have

$$|B| = O\left((\log n)^{3/2}/\sqrt{n}\right) = o\left(\frac{1}{1+|x|}\right)$$

uniformly for $|x| \leqq t\sqrt{\log n}$ as $n \to \infty$.

The law of iterated logarithm implies that

$$|S_n|/\sqrt{n} \leqq t\sqrt{\log n}$$

almost surely for any positive t and large enough n. Hence, for such S_n, Lemma 3 applies with $x = S_n/\sqrt{n}$ and $y = B \pm C/\varphi(x)$ provided n is large enough.

In our case S_n is continuous, thus the quantile transform (9) defines a unique function $H_n(x)$ such that $F_n(x\sqrt{n}) = \Phi(H_n(x))$. Lemma 3 and formula (10) imply that

$$|H_n(x) - (x+B)| \leqq C/\varphi(x) + 8(1+|x|)B^2,$$

which goes to zero almost surely provided $|S_n| \leqq t\sqrt{n\log n}$.

Thus we proved already that $T_n - \sqrt{n}D_n(S_n/\sqrt{n})$ goes to zero almost surely, where $D_n(x) = x + B$. Hence

$$T_n - \left(S_n + \kappa\left(\frac{S_n^2}{n} - 1\right)\right) \to 0 \quad \text{almost surely}$$

whence

$$S_n - \left(T_n - \kappa\left(\frac{T_n^2}{n} - 1\right)\right) \to 0 \quad \text{almost surely.} \qquad \square$$

3.2. A simple inequality

LEMMA 4. *For every positive $\beta < 1$ there is an $\alpha > 1$ and a threshold σ_0 such that, for all $\sigma \geqq \sigma_0$,*

$$\mathbf{P}(a < \sigma Z + Z^2 < b) \leqq (1+\sigma^{-\beta})\mathbf{P}(a < \sigma Z < b) \text{ for all } |a| \leqq \sigma^\alpha,\ 1 \leqq b - a \leqq 2,$$

where Z is standard normal.

COROLLARY. *Let Z be standard normal. For every positive $\beta < 1$ there is an $\alpha > 1$ and a threshold σ_0 such that, for all $\sigma \geqq \sigma_0$,*

$$\mathbf{P}(a < \sigma Z + Z^2 < b) \leqq (1+\sigma^{-\beta})\,\mathbf{P}(a < \sigma Z < b)$$

for all intervals (a, b) of length at least 1 intersecting the interval $(-\sigma^\alpha, \sigma^\alpha)$.

PROOF. Set $X = \sigma Z$, $Y = \sigma Z + Z^2$, and let us denote the corresponding densities by f and g, respectively. Then

$$f(x) = \frac{1}{\sigma}\varphi\left(\frac{x}{\sigma}\right) \quad \text{and} \quad g(x) = \frac{\varphi(z_1) + \varphi(z_2)}{\sqrt{\sigma^2 + 4x}},$$

where

$$z_1 = \frac{\sqrt{\sigma^2+4x} - \sigma}{2}, \quad z_2 = \frac{-\sqrt{\sigma^2+4x} - \sigma}{2}.$$

It is enough to show that

$$\frac{g(x)}{f(x)} \leqq 1 + \frac{1}{\sigma^\beta}$$

holds true if $|x| \leqq \sigma^\alpha$, $\sigma > \sigma_0$, and α, σ_0 are chosen appropriately. The term $\varphi(z_2)$ here is negligible because

$$z_2 \leqq -\frac{3}{4}\sigma \leqq -\frac{1}{4}\sigma \leqq \frac{x}{\sigma}.$$

For the term $\varphi(z_1)$, if $x<0$ then $|z_1|>|x|$, thus in this case it is enough to prove that

$$\frac{\sigma}{\sqrt{\sigma^2+4x}} \leqq 1+\frac{1}{\sigma^\beta},$$

which is an elementary fact. If $x>0$ then $\sigma \leqq \sqrt{\sigma^2+4x}$, thus it is enough to show that

$$\frac{\varphi(z_1)}{\varphi\left(\dfrac{x}{\sigma}\right)} \leqq 1+\frac{1}{\sigma^\beta},$$

which follows from the inequality

$$0 \leqq \left(1+\frac{u}{2}\right) - \sqrt{1+u} \leqq \frac{u^2}{8},$$

valid for all $u \geqq 0$ (in our case $u = 4x^2/\sigma^2$). □

3.3. Proof of Theorem 2

Lemma 2 already defines a sequence T_n (through the quantile transformation (9)). With this T_n define $T_n' = (T_n + \kappa T_n^2/n)$, and

$$T_n'' = \begin{cases} T_n' & \text{if } |T_n'| \leqq n^{\alpha/2} \\ 0 & \text{otherwise,} \end{cases}$$

where α was defined in Lemma 4. Then, by Lemma 2, $S_n - T_n''$ is bounded almost surely. Now we still have to define another normal sequence $\tilde{T}_n$ approximating T_n''. The Corollary to Lemma 4, applied with $Z = T_n/\sqrt{n}$ and $\sigma = \sqrt{n}$, implies that Condition (3) holds for T_n'' and T_n, with arbitrary $\varepsilon(x)$ and $\delta_n = 2n^{-2\alpha}$. The application of Theorem 1 concludes the proof. □

4. Concluding remarks

It is very likely that three moments are enough, and that Cramér's condition is not needed in Theorem 2. We needed it only in Lemma 2, where probably it is sufficient to assume that X_1 is a non-lattice variable. As a matter of fact, we do not need the full strength of Lemma 2, only the boundedness of $S_n - (T_n + \kappa T_n^2/n)$ is used. For this, probably the existence of some moments is enough.

The main differences betweeen previous embeddings of partial sums and the one developed here are the following:

the joint distributions of normal approximation are not fitted here, and

the Cornish–Fisher expansion is used instead of a simple one-term normal approximation.

One may ask what would be the result of a joined strategy: to use the diadic scheme with a Cornish–Fisher like expansion for conditional distributions (if there are any).

Our proof guaranteed the finiteness of $\sup_n |S_n - T_n|$ but not the finiteness of its expectation. This method probably gives $\mathbf{P}(\sup_n |S_n - T_n| > x) = O(1/x)$, but nothing better. We believe that for any embedding, $\mathbf{E} \sup_n |S_n - T_n| = \infty$.

Concerning Theorem 1, it would be interesting to have a simple characterization for all sequences of distributions F_n, G_n such that for any sequence S_n with marginals F_n there is a sequence T_n with marginals G_n such that (2) holds.

Referee's remark. It can be seen by means of a small trick that – as the authors guessed – the Cramér condition formulated in (5) can be dropped from the conditions of Theorem 2. Indeed, let us consider a random variable Z, $\mathbf{E}Z = 0, \mathbf{E}Z^2 = 1$, with finite moments whose characteristic function is concentrated in a finite interval (such a random variable exists), and which is independent of the sequence S_n. If we replace the random variable S_n by $\bar{S}_n = Z + S_{n-1}$, then the distribution of $\bar{S}_n$ satisfies the Edgeworth expansion (10). Hence the proof of Theorem 2 yields that Theorem 2 holds without the assumption (5) if S_n is replaced by $\bar{S}_n$. But then it also holds for the original sequence S_n.

The condition about the existence of four moments in Theorem 2 can be weakened, but it remains an open question whether the existence of three moments suffices, as the authors guess. This is an intriguing question, because a positive answer to it would mean that the conditions which are needed for the stochastic boundedness of the single random variables $|S_n - T_n|$ are also sufficient for the stochastic boundedness of the expression $\sup_n |S_n - T_n|$ with an appropriate construction.

REFERENCES

[1] Bártfai, P., Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungsfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation, *Studia Sci. Math. Hungar.* **1** (1966), 161–168. *MR* **35** #6218

[2] Bretagnolle, J. and Massart, P., Hungarian constructions from the nonasymptotic viewpoint, *Ann. Probab.* **17** (1989), 239–256. *MR* **91b**:60026

[3] Cornish, E. A. and Fisher, R. A., Moments and cumulants in the specification of distributions, *Rev. Inst. Int. Statist. La Haye* **5** (1937), 307–320. *Jb. Fortschritte Math.* **63**, 1081

[4] Csörgő, S. and Hall, P., The Komlós–Major–Tusnády approximations and their applications, *Austral. J. Statist.* **26** (1984), 189–218. *MR* **86f**:60039

[5] Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Akadémiai Kiadó, Budapest, 1981. Probability and mathematical statistics, Academic Press, New York, 1981. *MR* **84d**:60050

[6] Feller, W., *An introduction to probability theory and its applications*, Vol. II, Second Edition, John Wiley & Sons, Inc., New York – London – Sydney, 1971, p. 541. *MR* **42** #5292

[7] Komlós, J., Major, P. and Tusnády, G., An approximation of partial sums of independent RV's, and the sample DF, I-II, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **32** (1975), 111–131, **34** (1976), 33–58. *MR* **51** #11605b, *MR* **53** #6697

[8] Major, P., On the invariance principle for sums of independent identically distributed random variables, *J. Multivariate Anal.* **8** (1978), 487–517. *MR* **80f**:60034

[9] Strassen, V., The existence of probability measures with given marginals, *Ann. Math. Statist.* **36** (1965), 423–439. *MR* **31** #1693

[10] Tusnády, G., Strong invariance principles, *Recent developments in statistics* (Proc. European Meeting Statisticians, Grenoble, 1976), Ed. by J. Barra, North-Holland, Amsterdam, 1977, 289–300. *MR* **56** #16750

[11] Zolotarev, V. M., Asymptotic expansion with probability one of sums of independent random variables, *Teor. Veroyatnost. i Primenen.* **36** (1991), 770–772. *MR* **93c**:60067 and *Theory of Probability and its Applications* **36** (1992), 800–802.

(Received June 30, 1997)

DEPARTMENT OF MATHEMATICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NJ 08903
U.S.A.

komlos@math.rutgers.edu

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

tusnady@math-inst.hu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 273–304

ALMOST SURE FUNCTIONAL LIMIT THEOREMS PART I. THE GENERAL CASE

P. MAJOR

Dedicated to the memory of A. Rényi

Abstract

In this paper we formulate and prove the almost sure functional limit theorem in fairly general cases. This limit theorem is a result which states that if a stochastic process $X(t,\omega)$, $t \geqq 0$, is given on a probability space with some nice properties, then an appropriate probability measure λ_T can be defined on the interval $[1,T]$ for all $T>1$ in such a way that for almost all ω the distributions of the appropriate normalizations of the trajectories $X_t(\cdot,\omega)=X(t\cdot,\omega)$, considered as random variables $\xi_T(t)$, $t\in[1,T]$, on the probability spaces $([1,T],\mathcal{A},\lambda_T)$ with values in a function space have a weak limit independent of ω as $T\to\infty$. We shall consider self-similar processes which appear in different limit theorems. The almost sure functional limit theorem will be formulated and proved for them and their appropriate discretization under weak conditions. We also formulate and prove a coupling argument which makes it possible to prove the almost sure functional limit theorem for certain processes which converge to a self-similar process. In the second part of this work we shall prove and generalize — with the help of the results of the first part — some known almost sure functional limit theorems for independent random variables.

1. Introduction

The following "almost sure central limit theorem" is a popular subject in recent research. Let $X_1(\omega), X_2(\omega), \dots$ be a sequence of iid. random variables, $EX_1=0$, $EX_1^2=1$, $S_n(\omega)=\sum_{k=1}^{n} X_k(\omega)$ on a probability space $(\Omega,\mathcal{A},P)$. (In the sequel we denote by $(\Omega,\mathcal{A},P)$ the probability space where the random variables we are considering exist.) Then

$$\lim_{n\to\infty}\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}I\left(\frac{S_k(\omega)}{\sqrt{k}}<u\right)=\Phi(u)\quad\text{for almost all } \omega\in\Omega \tag{1.1}$$

and all numbers u, where $I(\mathbf{A})$ denotes the indicator function of a set $\mathbf{A}$, and $\Phi(u)$ is the standard normal distribution function. This result was discovered by Brosamler [2] and Schatte [7]. It states that appropriately normalized partial sums of iid. random variables satisfy not only the

1991 *Mathematics Subject Classification.* Primary 60F17; Secondary 28D05.

Key words and phrases. Almost sure invariance principle, Ornstein–Uhlenbeck process, ergodic theorem, self-similar processes, compactness.

0081-6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

central limit theorem, but for a typical $\omega \in \Omega$ the weighted averages of the functions $g_k(u,\omega) = I\left(S_k(\omega) < u\sqrt{k}\right)$ with appropriate weights converge to the normal law. Later this result was formulated in a more general form which states that not only the weighted averages of the functions $I\left(S_k(\omega) < u\sqrt{k}\right)$ converge to the normal distribution function for a typical ω, but a similar result also holds for sequences of random broken lines or polygons $G_n(u) = G_n(u,\omega)$, $n = 1, 2, \dots$, defined in an appropriate way on the interval $[0, 1]$ by means of the partial sums $S_1(\omega), \dots, S_n(\omega)$.

Define a random measure $\mu_n = \mu_n(\omega)$ for all n by attaching an appropriate weight $a_k = a_{k,n}$ to the functions $G_k(u,\omega)$ for all $1 \leqq k \leqq n$. Then these measures converge weakly to the Wiener measure for almost all ω. Such a result is called an almost sure functional limit theorem. Later we formulate this notion in a more detailed form.

The almost sure central (and also the functional) limit theorem shows some similarity to the ergod theorem which states — in physical terminology — that the space and time averages of ergodic sequences agree. In the case of the almost sure central limit theorem an analogous result holds for the normalized partial sums $\frac{S_k(\omega)}{\sqrt{k}}$, $k = 1, 2, \dots$. Now the time average is replaced by a weighted time average, where the k-th term gets weight $a_k = a_{k,n} = \frac{1}{\log(n+1)} \log \frac{k+1}{k} \sim \frac{1}{k \log n}$, $1 \leqq k \leqq n$, in the n-th block instead of the weight $\frac{1}{n}$ given to the first n terms in the ergod theorem. On the other hand, $\frac{S_n(\omega)}{\sqrt{n}}$ is asymptotically normally distributed, with expectation zero and variance one. Hence the right-hand side in formula (1.1) equals $\lim_{n\to\infty} EI\left(\frac{S_n(\omega)}{\sqrt{n}} < u\right)$, and this expression resembles to a space average. This similarity of the almost sure central limit theorem to the ergod theorem may be put even stronger by an appropriate time scaling to be explained later.

The relation between the ergod theorem and almost sure central (and functional) limit theorem is deeper than the above mentioned formal analogy. It was pointed out, — by our knowledge it was discovered by Brosamler [2], Fisher [5] and Lacey and Philipp [6] — that these theorems can be deduced from the ergod theorem applied to the Ornstein–Uhlenbeck process.

In the present paper we discuss how the almost sure central and functional limit theorem can be generalized and proved by means of the ergod theorem in a natural way. The proof has two main ingredients. The first one is to show that a result analogous to the almost sure functional limit theorem holds for the Wiener process. This can be deduced from the ergod theorem for the Ornstein–Uhlenbeck process. This is an ergodic process which can be obtained from the Wiener process by means of a well-known transformation. The second ingredient is to show that, since the random polygons or broken lines constructed from the partial sums of independent random variables in a natural way behave similarly to the Wiener process, the almost sure central limit theorem for the Wiener process also implies this result for the random polygons (or broken lines) made from normalized partial sums of independent random variables.

First we show that the method of proving the almost sure functional limit theorem for the Wiener process by means of the ergod theorem for the Ornstein–Uhlenbeck process can be generalized for a large class of other processes, for the so-called self-similar processes. The stationarity property of the Ornstein–Uhlenbeck process is equivalent to the self-similarity property of the Wiener process, a property which holds for all self-similar processes. Actually, self-similar processes are those processes which appear as the limit in different limit theorems. Similarly to the construction of the Ornstein–Uhlenbeck process generalized Ornstein–Uhlenbeck processes can be constructed as the transforms of self-similar processes. These generalized Ornstein-Uhlenbeck processes are stationary processes, and the application of the ergod theorem for them enables us to prove the almost sure functional limit theorem for general self-similar processes. Then with the help of some further work we can also prove the almost sure functional limit theorem for their appropriate discretized versions.

In the next step we want to find a good coupling argument which enables us to prove the almost sure invariance principle not only for (self-similar) limit processes but also for processes in the domain of their attraction. To carry out such a program a coupling argument has to be introduced which is adapted to the present problem. We shall do it by introducing a notion we call the Property A.

In Part II of this work we shall prove the almost sure functional limit theorem for independent random variables whose partial sums converge to the normal or to a stable law. In the proofs we shall exploit that the Wiener process and the stable process are self-similar, hence the results of the present paper can be applied for them. Then we can prove, by applying the coupling argument of the present paper, the almost sure invariance principle for independent random variables which satisfy certain (weak) conditions.

There are other processes which are natural candidates for almost sure functional limit theorem type results, e.g. random processes in the domain of attraction of a self-similar process subordinated to a Gaussian process (see Dobrushin [3]). But such problems will not be discussed here.

Several results of the present paper can be traced down in earlier works. Our main goal is to explain the main ideas behind these results and to present a unified treatment of various problems in this subject. The first part of this work considers general results where no independence type condition is assumed. In the second part different arguments — the techniques worked out for the study of independent random variables — are applied, and we deal there with almost sure functional limit theorems for independent random variables. This paper consists of three sections. In Section 2 we formulate the main results, and Section 3 contains the proofs.

2. The main results of the paper

To formulate our results first we recall the definition of self-similar processes with self-similarity parameter α and define with their help a new process which we call a generalized Ornstein–Uhlenbeck process.

DEFINITION of self-similar processes. We call a stochastic process $X(u,\omega)$, $u \geqq 0$, $X(0,\omega) \equiv 0$, self-similar with self-similarity parameter α, $\alpha > 0$, if

$$X(u,\omega) \stackrel{\Delta}{=} \frac{X(Tu,\omega)}{T^{1/\alpha}}, \quad 0 \leqq u < \infty, \tag{2.1}$$

for all $T > 0$, where $\stackrel{\Delta}{=}$ means that the processes at the two sides of the equation have the same distribution. (Here we consider the distribution of the whole process $X(u,\omega)$, $u \geqq 0$, and not only its one-dimensional distributions.)

The Wiener process is self-similar with self-similarity parameter $\alpha = 2$. Similarly, for all stable laws G with parameter α, $0 < \alpha < 2$, $\alpha \neq 1$, a so-called stable process $X(u,\omega)$ can be constructed which has independent and stationary increments, $X(0,\omega) \equiv 0$, which is self-similar with self-similarity parameter α, and the distribution function of $X(1,\omega)$ is G. The case $\alpha = 1$ is exceptional. In this case (except the special case when $X(1,\omega)$ has symmetric distribution) only a modified version of formula (2.1) holds, where a norming factor const. $\log T$ must be added with an appropriate non-zero constant to one side in formula (2.1). Another example for self-similar processes was given by Dobrushin in paper [3], who could construct new kind of self-similar processes subordinated to a Gaussian process. He constructed them by working with non-linear functionals of Gaussian processes.

Now we introduce the following notion:

DEFINITION of generalized Ornstein-Uhlenbeck processes. Let $X(u,\omega)$, $u \geqq 0$, be a self-similar process with self-similarity parameter $\alpha > 0$. We call the process $Z(t,\omega)$, $-\infty < t < \infty$, defined by formula

$$Z(t,\omega) = \frac{X(e^t,\omega)}{e^{t/\alpha}}, \quad -\infty < t < \infty, \tag{2.2}$$

the generalized Ornstein–Uhlenbeck process corresponding to the process $X(u,\omega)$.

Let us remark that the generalized Ornstein–Uhlenbeck process corresponding to the Wiener process is the usual Ornstein–Uhlenbeck process.

A Wiener process $W(t,\omega)$, $t \ge 0$, has continuous trajectories, the trajectories of a stable process $X(t,\omega)$ are so-called càdlàg (continue à droite, limite à gauche) functions, i.e. all trajectories $X(\cdot,\omega)$ are continuous from the right, and have a left-hand side limit in all points $t > 0$. Hence the Wiener process $W(t,\omega)$ and any of its scaled version $A_T W(Tt,\omega)$, $0 \le t \le 1$, where $T > 0$ and $A_T > 0$ are arbitrary constants, can be considered as random variables taking values in the space $C([0,1])$ of continuous functions on the interval $[0,1]$. The processes $X(t,\omega)$, $A_T X(tT,\omega)$, $0 \le t \le 1$, where $X(t,\omega)$, $0 \le t < \infty$, is a stable process, can be considered as random variables on the space $D([0,1])$ of càdlàg functions on the interval $[0,1]$.

We shall work not only in the space $C([0,1])$ but also in the space $D([0,1])$. To work in the space $D([0,1])$ one has to handle some unpleasant technical problems. But since we also want to investigate stable processes in Part II of this work, we also have to work in this space. We shall apply the book of P. Billingsley [1] as the main reference for this subject.

We consider both spaces $C([0,1])$ and $D([0,1])$ with the usual topology, and the Borel σ-algebra generated by this topology. Both spaces can be endowed with a metric which induces this topology, and with which these spaces are separable, complete metric spaces. A detailed discussion and proof of these results and definitions can be found in the book of P. Billingsley [1]. Since we shall need the exact form of these metrics we recall these results. In the $C([0,1])$ space the supremum metric $\rho(x,y) = \sup\limits_{0 \le t \le 1} |x(t) - x(s)|$ is considered. In the space $D([0,1])$ the following metric $d_0(\cdot,\cdot)$ satisfies these properties: For a pair of functions $x, y \in D([0,1])$ $d_0(x,y) \le \varepsilon$, if there exists such a homeomorphism $\lambda(t)\colon [0,1] \to [0,1]$ of the interval $[0,1]$ into itself for which $\lambda(0) = 0$, $\sup\limits_{t \ne s} \log \left| \frac{\lambda(t) - \lambda(s)}{t - s} \right| \le \varepsilon$, and $|x(t) - y(\lambda(t))| \le \varepsilon$ for all $t \in [0,1]$. (See for instance Theorems 14.1 and 14.2 in Billingsley's book [1].) In the sequel we shall apply these metrics in the spaces $C([0,1])$ and $D([0,1])$, and denote them by $\rho(\cdot,\cdot)$.

Let us also recall that given some probability measures μ_T on a metric space $\mathbf{K}$ indexed by $T \in [1,\infty)$ or $T = \{A_1, A_2, \dots\}$, $\lim\limits_{n \to \infty} A_N = \infty$, the measures μ_T converge weakly to a measure μ on $\mathbf{K}$ as $T \to \infty$ if $\lim\limits_{T \to \infty} \int\limits_{\mathbf{K}} \mathcal{F}(x) \mu_T(dx) = \int\limits_{\mathbf{K}} \mathcal{F}(x) \mu(dx)$ for all continuous and bounded functionals $\mathcal{F}$ on the space $\mathbf{K}$. The next result states the almost sure functional limit theorem for a self-similar process which satisfies some additional conditions. The proof is based on the ergod theorem applied for the generalized Ornstein–Uhlenbeck process corresponding to this self-similar process.

THEOREM 1. *Let $X(u,\omega)$ be a self-similar process with continuous or càdlàg trajectories, and $Z(t,\omega)$ the generalized Ornstein–Uhlenbeck process corresponding to it. The process $Z(t,\omega)$, $-\infty < t < \infty$, is stationary. Let us assume that the process $Z(t,\omega)$ is not only stationary, but also ergodic.*

Then for all measurable and bounded functionals $\mathcal{F}$ on the space $C([0,1])$ or $D([0,1])$ (depending on whether the trajectories of $X(\cdot,\omega)$ are continuous or only càdlàg functions)

$$\lim_{T\to\infty}\frac{1}{\log T}\int_1^T \frac{1}{t}\mathcal{F}(X_t(u,\omega))\,dt = E\mathcal{F}(X_1(u,\omega)) \quad \textit{for almost all } \omega, \tag{2.3}$$

where

$$X_t(u,\omega) = \frac{X(ut,\omega)}{t^{1/\alpha}}, \quad 0 \leqq u \leqq 1,\ t > 0. \tag{2.4}$$

Let us define for all $\omega\in\Omega$ and $T\geqq 1$ the (random) probability measure $\mu_T(\omega)$ in the space $C([0,1])$ or $D([0,1])$ which is concentrated on the trajectories $X_t(\omega)$, $1\leqq t\leqq T$, and takes the value $X_t(\omega)$, $1\leqq t\leqq T$, with probability $\frac{1}{\log T}\frac{dt}{t}$. More formally, for a measurable set $\mathbf{A}\subset C([0,1])$ or $\mathbf{A}\subset D([0,1])$ put $\mu_T(\omega)(\mathbf{A}) = \lambda_T\{t\colon X_t(\omega)\in\mathbf{A}\}$, where λ_T is a measure on $[1,T]$ defined by the formula $\lambda_T(\mathbf{C}) = \frac{1}{\log T}\int_{\mathbf{C}}\frac{dt}{t}$ for all measurable sets $\mathbf{C}\subset[1,T]$.

The following version of Formula (2.3) also holds: For almost all $\omega\in\Omega$ the probability measures $\mu_T(\omega)$ converge weakly to the distribution of the process $X_1(u,\omega)$ defined in (2.4) with $t=1$, which we denote by μ_0 in the sequel. In other words, there is a set of probability one such that if ω is in this set then relation (2.3) holds for this ω and all *bounded and continuous functionals $\mathcal{F}$.*

If $X(u,\omega)$ is a Wiener or stable process, then the generalized Ornstein–Uhlenbeck process corresponding to it is not only stationary, but also ergodic. Hence the results of Theorem 1 are applicable in this case.

We want to prove a discretized version of the above result, where the measures $\mu_T(\omega)$ concentrated in the set of trajectories $X_t(\omega)$, $1\leqq t\leqq T$, are replaced by some measures $\mu_N(\omega)$ which are concentrated on a set of trajectories $X_{a(j,N)}(\omega)$ with appropriate weights, and the numbers $a(j,N)$ constitute a finite set. Then we want to make a further discretization, where the trajectories $X_{a(j,N)}$ are replaced by their discretized version. To prove these results in the case when the trajectories of the process $X(\cdot,\omega)$ are càdlàg functions we impose the following additional condition:

$$P\left(\lim_{t\to 1-0} X(t,\omega) = X(1,\omega)\right) = 1. \tag{2.5}$$

First we formulate a result which serves as the basis of the discretization results formulated later.

THEOREM 2. *Let $X(u,\omega)$, $X_t(u,\omega)$, $\mu_T(\omega)$ and μ_0 be the same as in Theorem 1. Let us assume that the conditions of Theorem 1 are satisfied, and also the additional condition* (2.5) *holds in the case when the process $X(\cdot,\omega)$ has càdlàg trajectories. Let us define, similarly to the trajectories $X_t(\cdot,\omega)$ defined in* (2.4), *the following transformed functions $x_t=x_t(\cdot)$ of a function $x\in C([0,1])$ or $x\in D([0,1])$ by the formula*

$$x_t(u)=x_{t,\alpha}(u)=t^{-1/\alpha}x(ut), \quad 0\leqq u\leqq 1,\ 0<t\leqq 1, \tag{2.4'}$$

where α is the self-similarity parameter of the underlying self-similar process $X(\cdot,\omega)$. Then for almost all $\omega\in\Omega$

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\mu_T(\omega)\left(\sup_{1-\varepsilon\leqq s,t\leqq 1}\rho(x_s,x_t)>\delta\right)=0 \quad \textit{for all } \delta>0, \tag{2.6}$$

where $\rho(\cdot,\cdot)$ is the metric whose definition was recalled before Theorem 1, and with which $C([0,1])$ or $D([0,1])$ are separable, complete metric spaces. (Let us recall that the (random) measure $\mu_T(\omega)$ is concentrated on the trajectories $X_u(\cdot,\omega)$, $1\leqq u\leqq T$, of the process $X(\cdot,\omega)$ defined by formula (2.4).)

Condition (2.5) had to be imposed to control the behaviour of the trajectories of the processes $X_t(u,\omega)$ in the end point $u=1$. This is not a strict restriction. For instance the next simple Lemma 1 gives a sufficient condition for its validity. It implies in particular, that the stable processes with self-similarity parameter α, $0<\alpha<2$, $\alpha\neq 1$, satisfy relation (2.5).

LEMMA 1. *Let $X(\cdot,\omega)$ be a self-similar process with self-similarity parameter $\alpha>0$ which is also a process with stationary increments, and whose trajectories are càdlàg functions. Then it satisfies relation* (2.5).

Now we formulate the result about "possible discretization" of the measures μ_T in the result of Theorem 1. Before this we make some comments which can explain the content of this result.

For all $T>1$ let us consider the probability space $([1,T],\mathcal{A},\lambda_T)$, where $\mathcal{A}$ is the Borel σ-algebra, and λ_T is the measure defined in the formulation of Lemma 1. Fix an $\omega\in\Omega$, and let us consider the random variable $\xi(t)$, $1\leqq t\leqq T$, as $\xi(t)=X_t(\cdot,\omega)$, defined in formula (2.4), in the probability space $([1,T],\mathcal{A},\lambda_T)$. This is a random variable which takes its value in the space $C([0,1])$ or $D([0,1])$, and it has distribution $\mu_T(\omega)$. Let us consider the above construction with some $T=B_N$, together with a dense splitting $1=B_{N,1}<B_{N,2}<\cdots<B_{N,k_N}=B_N$ of the interval $[1,B_N]$. Let us define the random variable $\hat{\xi}(t)$ such that $\hat{\xi}(t)=\xi(B_{k,N})=X_{B_{k,N}}(\cdot,\omega)$ if $t\in[B_{k,N},B_{k+1,N}]$. This random variable is close to the previously defined random variable $\xi(t)$, hence it is natural to expect that if $\hat{\mu}_{B_n}(\omega)$ denotes its distribution, then the measures $\hat{\mu}_{B_N}(\omega)$ have the same weak limit as the measures $\mu_{B_N}(\omega)$ as $N\to\infty$. The first statement of Theorem 3 is a result

of this type. Then we prove that an appropriate small modification of the functions $\xi(B_{k,N})=X_{B_k,N}(\cdot,\omega)$ does not change the limit behaviour of the measures $\hat\mu_{B_N}(\omega)$. The second statement of Theorem 3 is such a result.

THEOREM 3. *Let us assume that the conditions of Theorem 1 and Theorem 2 are satisfied. For all $N=0,1,\dots$ let us consider a finite increasing sequence of real numbers $1=B_{1,N}<B_{2,N}<\dots<B_{k_N,N}$, and for the sake of simpler notation let us denote $B_{k_N,N}$ by B_N. Let us assume that these sequences satisfy the following properties:*

$$\lim_{N\to\infty} B_N=\infty,\qquad \lim_{N\to\infty}\frac{\log B_{j,N}}{\log B_N}=0 \quad \textit{for all fixed } j, \tag{2.7}$$

and

$$\lim_{j\to\infty}\ \sup_{(k,N):\ j\leqq k<N}\frac{B_{k+1,N}}{B_{k,N}}=1.$$

Moreover, assume the following strengthened form of the relation $\lim\limits_{N\to\infty}B_N=\infty$:

$$\lim_{j\to\infty}\ \inf_{N:\ N\geqq j} B_{j,N}=\infty. \tag{2.8}$$

For all $\omega\in\Omega$ define the (random) measures $\hat\mu_N(\omega)$, $N=1,2,\dots$, with the help of the sequences $1=B_{1,N}<B_{2,N}<\dots<B_{k_N,N}$ in the following way:

The measure $\hat\mu_N(\omega)$, $N=1,2,\dots$, is concentrated on the trajectories $X_{B_{j,N}}(\cdot,\omega)$, $1\leqq j<k_N$, where $X_t(\cdot,\omega)$ is defined in (2.4), and

$$\hat\mu_N(\omega)(X_{B_{j,N}}(\cdot,\omega))=\frac{1}{\log B_N}\int_{B_{j,N}}^{B_{j+1,N}}\frac{1}{u}\,du=\frac{1}{\log B_N}\log\frac{B_{j+1,N}}{B_{j,N}},\qquad 1\leqq j<k_N. \tag{2.9}$$

Then for almost all ω the measures $\hat\mu_N(\omega)$ converge weakly to μ_0 defined in Theorem 1.

For all $\omega\in\Omega$ let us also define the following random broken lines $\bar X_{B_{j,N}}(\cdot,\omega)$ which are "discretizations" of the trajectories $X_{B_{j,N}}(\cdot,\omega)$.

$$\bar X_{B_{j,N}}(s,\omega)=X_{B_{j,N}}\left(\frac{B_{l-1,N}}{B_{j,N}},\omega\right)\quad \textit{if}\quad \frac{B_{l-1,N}}{B_{j,N}}\leqq s<\frac{B_{l,N}}{B_{j,N}},$$
$$1\leqq l\leqq j,\ 1\leqq j<k_N,\quad \textit{and}\ \bar X_{B_{j,N}}(1,\omega)=X_{B_{j,N}}(1,\omega),$$

where $B_{0,N}=0$. (The definition $B_{0,N}=0$ is needed to define $\bar X_{B_{j,N}}(s,\omega)$ also for $0\leqq sB_{j,N}<B_{1,N}$.)

Define the measures $\bar\mu_N(\omega)$ (with the help of the already defined measures $\hat\mu_N(\omega)$) as
(2.9′)
$$\bar\mu_N(\omega)(\bar X_{B_{j,N}}(\cdot,\omega))=\hat\mu_N(\omega)(X_{B_{j,N}}(\cdot,\omega))=\frac{1}{\log B_N}\log\frac{B_{j+1,N}}{B_{j,N}},\quad 1\leqq j<k_N.$$

Then for almost all $\omega\in\Omega$ the probability measures $\bar\mu_N(\omega)$ converge weakly to the probability measure μ_0 defined in Theorem 1 as $N\to\infty$.

We have defined $\bar X_{B_{j,N}}(\cdot,\omega)$ as a broken line with discontinuities and not as a polygon where the values of $X_{B_{j,N}}$ in the points $\frac{B_{l,N}}{B_{j,N}}$ are connected by linear segments. The reason for working with broken lines is that we want to prove results which are valid also in the case when the processes $X_t(\cdot,\omega)$ take their values in $D([0,1])$ but not necessarily in the space $C([0,1])$. In the general case the results we want to prove are valid only when broken lines are considered. In the case of processes with continuous trajectories we also could have defined them as random polygons. Moreover, it follows from some results of the general theory (see e.g. Section 18 in Billingsley's book [1]) that if the distribution of the processes consisting of the above defined random broken lines converge to a measure in the $C([0,1])$ space, then the distributions of the naturally defined random polygon version of these processes have the same limit in the $C([0,1])$ space.

Let $\xi_n(\omega)$, $n=1,2,\dots$, be a sequence of random variables, and let us define the partial sums $S_n(\omega)=\sum\limits_{k=1}^{n}\xi_k(\omega)$, $n=1,2,\dots$, $S_0(\omega)\equiv0$. Let us also consider two appropriate monotone increasing numerical sequences A_n and B_n, $n=0,1,\dots$, of positive numbers such that

$$\text{(2.10)}\qquad B_0=0,\quad \lim_{n\to\infty}A_n=\infty,\quad \lim_{n\to\infty}B_n=\infty,\quad\text{and}\quad \lim_{n\to\infty}\frac{B_{n+1}}{B_n}=1.$$

For all $k=1,2,\dots$ let us consider the partition $0=s_{0,k}\leqq s_{1,k}\leqq\dots\leqq s_{k,k}$ of the interval $[0,1]$, defined by the formula $s_{j,k}=\frac{B_j}{B_k}$, $0\leqq j\leqq k$. Let us also define with the help of the quantities $\xi_n(\omega)$, A_n and B_n, $n=1,2,\dots$ the following random broken lines $S_k(s,\omega)$, $0\leqq s\leqq 1$, $k=1,2,\dots$,
(2.11)
$$S_k(s,\omega)=\frac{S_{j-1}(\omega)}{A_k}\quad\text{if } s_{j-1,k}\leqq s<s_{j,k},\ 1\leqq j\leqq k,\quad S_k(1,\omega)=\frac{S_k(\omega)}{A_k}.$$

Now we introduce the following definition:

Definition of the almost sure functional limit theorem. Let $\xi_n(\omega)$, $n=1,2,\dots$, be a sequence of random variables, and let two monotone increasing sequences of non-negative real numbers A_n and B_n, $n=1,2,\dots$, be given which satisfy formula (2.10). Let us consider the random broken lines $S_k(s,\omega)$, $0 \leqq s \leqq 1$, defined with the help of their partial sums $S_k(\omega)$, $k=1,2,\dots$, by formula (2.11). For all $\omega \in \Omega$ and $N=1,2,\dots$, define the random measure $\mu_N(\omega)$ in the following way: The measure $\mu_N(\omega)$ is concentrated on the random broken lines $S_k(\cdot,\omega)$, $1 \leqq k < N$, and

$$\mu_N(\omega)(S_k(\cdot,\omega)) = \frac{1}{\log \frac{B_N}{B_1}} \log \frac{B_{k+1}}{B_k}, \quad 1 \leqq k < N. \tag{2.12}$$

We say that the sequence of random variables $\xi_n(\omega)$, $n=1,2,\dots$, satisfies the almost sure functional limit theorem with weight functions A_n and B_n, $n=1,2,\dots$, and limit measure μ_0 on the space $D([0,1])$ if for almost all $\omega \in \Omega$ the probability measures $\mu_N(\omega)$ converge weakly to the measure μ_0 as $N \to \infty$. In the special case when the limit measure μ_0 is the Wiener measure we say that these random variables satisfy the almost sure functional central limit theorem.

If the limit measure μ_0 is concentrated in the space $C([0,1])$, then the broken lines $S_k(\cdot,\omega)$ can be replaced by a natural modification which is a random polygon. Then we can consider a version of the measures $\mu_N(\omega)$ which are defined in the same way as the original ones, only the random processes $S_k(\cdot,\omega)$ are replaced by their random polygon version. Then the convergence of the original measures $\mu_N(\omega)$ to μ_0 in the space $D([0,1])$ implies the convergence of their modified version in the $C([0,1])$ space with the same limit. Let us also remark that although we allowed fairly large freedom in the choice of the sequence A_n in the definition of the almost sure functional limit theorem, nevertheless we shall always choose it in a very special way. Namely, in all almost sure functional limit theorems we shall prove the limit measure is the distribution of a self-similar process with a self-similarity parameter $\alpha > 0$ restricted to the interval $[0,1]$, and A_n is chosen as $A_n = B_n^{1/\alpha}$.

Let us remark that if the random variables $\xi_k(\omega)$ satisfy the almost sure functional central limit theorem with weight functions $A_n = \sqrt{n}$ and $B_n = n$, — and in Part II we shall prove that under the conditions imposed for the validity of formula (1.1) this is the case, — then they also satisfy relation (1.1). To see this, fix a real number u and define the functional $\mathcal{F} = \mathcal{F}_t$ in the space $C([0,1])$ by the formula $\mathcal{F}(x) = 1$ if $x(1) < u$, and $\mathcal{F}(x) = 0$ if $x(1) \geqq u$, where $x \in C([0,1])$, i.e. it is a continuous function on the interval $[0,1]$. This functional $\mathcal{F}$ is continuous with probability one with respect to the Wiener measure μ_0. Hence $\int \mathcal{F}(x)\, d\mu_n(\omega)(x) \to \int \mathcal{F}(x)\, d\mu_0(x)$ for almost all ω. This relation is equivalent to formula (1.1). Indeed, the right-hand side of this relation equals the right-hand side of formula (1.1), while the left-hand side is a slight modification of the left-hand side of (1.1). The difference between these formulas is that the weights $\frac{1}{k}$ in (1.1) are replaced by $\log \frac{k+1}{k}$ in the other formula, and summation goes from 1 to $n-1$ instead of summation from 1 to n. Since $\log \frac{k+1}{k} = \frac{1}{k} + O\left(\frac{1}{k^2}\right)$ these two relations are equivalent.

We formulate the following statement because of its importance in later applications in form of a Corollary.

COROLLARY. *Let $X(\cdot,\omega)$ be a self-similar process with self-similarity parameter $\alpha > 0$ such that its trajectories are in the $C([0,1])$ or $D([0,1])$ space, it satisfies relation (2.5), and the generalized Ornstein–Uhlenbeck process corresponding to it is ergodic. Let t_n, $n = 0,1,\dots$, $t_0 = 0$, be an increasing sequence of real numbers such that $\lim_{n\to\infty} t_n = \infty$, $\lim_{n\to\infty} \frac{t_{n+1}}{t_n} = 1$. Put $\eta_n(\omega) = X(t_n,\omega) - X(t_{n-1},\omega)$, $B_n = t_n$, $A_n = B_n^{1/\alpha}$, $n = 1,2,\dots$. Then the sequence $\eta_n(\omega)$, $n = 1,2,\dots$, satisfies the almost sure functional limit theorem with weight functions A_n and B_n and limit measure μ_0 which is the distribution of the process $X(u,\omega)$, restricted to $0 \leqq u \leqq 1$.*

To prove this Corollary define the process $X'(u,\omega) = A_1^{-1} X(B_1 u,\omega)$ and observe that it has the same distribution as the process $X(u,\omega)$. Define the real numbers $B_{k,N} = \frac{t_k}{t_1}$, $1 \leq k \leq N$, consider the random broken lines $\bar{X}'_{B_{j,N}}(\cdot,\omega)$, $1 \leqq j \leqq N$, and the random measure $\bar{\mu}_N(\omega)$ defined in the formulation of Theorem 3 with this process $X'(\cdot,\omega)$ and these numbers $B_{k,N}$, (with the choice $k_N = N$), and apply Theorem 3, — whose conditions are satisfied, — for these random measures $\bar{\mu}_N(\omega)$.

On the other hand, define the random broken lines $S_k(s,\omega)$ by formula (2.11) with $B_N = t_N$, $A_N = B_n^{1/\alpha}$ and the partial sums $S_k(\omega) = \sum_{l=1}^{k} (X(t_l,\omega) - X(t_{l-1},\omega))$, and let us also define the measure $\mu_N(\omega)$ by formula (2.12) with these random broken lines. Then a comparison shows that the above defined broken lines $\bar{X}'_{B_{j,N}}(\cdot,\omega)$ and $S_j(\cdot,\omega)$ and also their distributions, the random measures $\bar{\mu}_N(\omega)$ and $\mu_N(\omega)$ agree. Hence the second statement of Theorem 3 implies the almost sure functional limit theorem in this case.

If a sequence of random variables $\xi_n(\omega)$, $n = 1,2,\dots$, is close to this sequence $\eta_n(\omega)$, then it is natural to except that this new sequence satisfies the same almost sure functional limit theorem. We want to give a good coupling argument that enables us to prove this for a large class of processes $\xi_n(\omega)$. For this aim we define a Property A. We prove that if Property A holds for a pair of sequences of random variables $(\xi_n(\omega), \eta_n(\omega))$, $n = 1,2,\dots$, and the sequence $\eta_n(\omega)$, $n = 1,2,\dots$, satisfies the almost sure functional limit theorem, then the sequence $\xi_n(\omega)$, $n = 1,2,\dots$ also satisfies the almost sure functional limit theorem with the same norming constants and limit law.

DEFINITION of Property A. Let $\eta_n(\omega)$, $n = 1,2,\dots$, be a sequence of random variables which satisfies the almost sure functional limit theorem with a limit measure μ_0 in the space $C([0,1])$ or $D([0,1])$ and some weight functions A_n and B_n satisfying relation (2.10). Let us also assume that the limit measure μ_0 is the distribution of the restriction of a self-similar process $X(u,\omega)$ with self-similarity parameter $\alpha>0$ to the interval $0\leqq u\leqq 1$, and the weight functions A_n and B_n are such that $A_n=B_n^{1/\alpha}$.

Define the indices $N(n)$ as $N(n)=\inf\{k\colon B_k\geqq 2^n\}$, $n=0,1,\dots$. The pairs of sequences of random variables $(\xi_n(\omega),\eta_n(\omega))$, $n=1,2,\dots$, satisfy Property A if for all $\varepsilon>0$ and $\delta>0$ there exists a sequence of random variables $\bar\xi_n(\omega)=\bar\xi_n(\varepsilon,\delta,\omega)$, $n=1,2,\dots$, whose (joint) distribution agrees with the (joint) distribution of the sequence $\xi_n(\omega)$, $n=1,2,\dots$, and the partial sums $\bar S_n(\omega)=\sum\limits_{k=1}^{n}\bar\xi_k(\omega)$ and $T_n(\omega)=\sum\limits_{k=1}^{n}\eta_k(\omega)$ satisfy the following relation:

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{k=1}^{N(n)}\log\frac{B_{k+1}}{B_k}I\left(\left\{\frac{\sup\limits_{0\leqq j\leqq k}|\bar S_j(\omega)-T_j(\omega)|}{A_k}>\varepsilon\right\}\right)<\delta \tag{2.13}$$

for almost all $\omega\in\Omega$, where $I(A)$ denotes the indicator function of the set A.

REMARK. Let us remark that the joint distribution of the random variables $\xi_n(\omega)$, $n=1,2,\dots$, determines whether it satisfies the almost sure invariance principle. It is not important how and on which probability space these random variables are constructed. This can be seen for instance by applying the following "canonical representation" of the sequence $\xi_n(\omega)$, $n=1,2,\dots$, on the probability space $(\Omega,\mathcal{A},P)$. Define the space $(R^\infty,\mathcal{B}^\infty,\bar\mu)$, where $R^\infty=\{(x_1,x_2,\dots)\colon x_j\in R,\ j=1,2,\dots\}$, $\mathcal{B}^\infty$ is the Borel σ-algebra on R^∞, $\bar\mu(\mathbf{B})=P((\xi_1,\xi_2,\dots)\in\mathbf{B})$ for $\mathbf{B}\in\mathcal{B}^\infty$, and define the random variables $\bar\xi_n(x_1,x_2,\dots)=x_n$, $n=1,2,\dots$, on this space. Then the random variables $\bar\xi_n$ on the space $(R^\infty,\mathcal{B}^\infty,\bar\mu)$ have the same joint distribution as the random variables $\xi_n(\omega)$, and these two sequences satisfy the almost sure invariance principle simultaneously.

THEOREM 4. *Let $\eta_n(\omega)$, $n=1,2,\dots$, be a sequence of random variables which satisfies the almost sure functional limit theorem, and let a pair of sequences of random variables $(\xi_n(\omega),\eta_n(\omega))$, $n=1,2,\dots$, satisfy Property A. Then the sequence of random variables $\xi_n(\omega)$, $n=1,2,\dots$, also satisfies the almost sure functional limit theorem with the same weight functions A_n and B_n and limit measure μ_0 as the sequence of random variables $\eta_n(\omega)$.*

We shall prove in Part II of this work that Theorem 4 is applicable in several interesting cases. We shall prove with the help of a Basic Lemma formulated there that when partial sums of independent random variables are considered, then an appropriate construction satisfies the conditions of Theorem 4 under general conditions. In such a way it will turn out that the necessary and sufficient conditions of limit theorems for normalized partial sums of independent random variables are also sufficient conditions for the almost sure functional limit theorem.

We shall prove still another result which states that a small perturbation of the weight functions B_n does not affect the validity of the almost sure functional limit theorem. The reason to prove such a result is the following. We have certain freedom in the choice of the weight functions B_n, and there are cases when no "most natural choice" of the weight functions exists. We want to show that different natural choices yield equivalent results. Let us remark that a modification of the weight functions B_n also implies a modification of the random broken lines $S_n(t,\omega)$ appearing in the definition of the almost sure functional limit theorem.

THEOREM 5. *Let a sequence of random variables* $\xi_n(\omega)$, $n=1,2,\dots$, *satisfy the almost sure functional limit theorem with some limit measure* μ_0 *and weight functions* B_n, $A_n=B_n^{1/\alpha}$ *with some* $\alpha>0$, $n=0,1,\dots$, *which satisfies relation* (2.11). *Let* $\bar B_n$, $n=0,1,\dots$, $\bar B_0=1$, *be another monotone increasing sequence such that* $\lim\limits_{n\to\infty}\frac{\bar B_n}{B_n}=1$. *Put* $\bar A_n=\bar B_n^{1/\alpha}$. *Then the sequence of random variables* $\xi_n(\omega)$ *also satisfies the almost sure functional limit theorem with the limit measure* μ_0 *and weight functions* $\bar B_n$ *and* $\bar A_n$.

3. Proof of the results

PROOF OF THEOREM 1. We can write

$$Z(t+T,\omega)=\frac{X(e^{t+T},\omega)}{e^{(t+T)/\alpha}}\stackrel{\Delta}{=}\frac{X(e^{t},\omega)}{e^{(t+T)/\alpha}e^{-T/\alpha}}=\frac{X(e^{t},\omega)}{e^{t/\alpha}}=Z(t,\omega)$$

for all $-\infty<T<\infty$. Hence the process $Z(t,\omega)$, $-\infty<t<\infty$, is stationary. If it is not only stationary, but also ergodic, then the ergod theorem can be applied for the process $Z(\cdot,\omega)$ and all bounded and measurable functionals $\mathcal{G}$ on the space $(R^{(-\infty,\infty)},\mathcal{B},\mu)$, where $R^{(-\infty,\infty)}$ is the space of functions on the interval $(-\infty,\infty)$, $\mathcal{B}_0$ is the σ-algebra induced by the usual Borel (product) topology on $R^{(-\infty,\infty)}$, μ is the distribution of the process $Z(\cdot,\omega)$ on the space $(R^{(-\infty,\infty)},\mathcal{B}_0)$, and $\mathcal{B}$ is the closure of the σ-algebra $\mathcal{B}_0$ with respect to the measure μ. This means that $\mathbf{B}\in\mathcal{B}$ if and only if there exists some $\mathbf{B}_0\in\mathcal{B}_0$ such that $\mu(\mathbf{B}_0\Delta\mathbf{B})=0$ for the symmetric difference $\mathbf{B}_0\Delta\mathbf{B}$, or more precisely there is a $\mathcal{B}_0$-measurable set $\mathbf{C}$ such that $\mu(\mathbf{C})=0$ and $\mathbf{B}_0\Delta\mathbf{B}\subset\mathbf{C}$. Furthermore, we introduce the shift operators $\mathbf{T}_s$ defined by the formula $\mathbf{T}_s(z(\cdot))=z(s+\cdot)$ for all $z(\cdot)\in R^{(-\infty,\infty)}$ and put $Z_s(v,\omega)=$

$Z(s+v,\omega)$, $-\infty<v<\infty$. Then the ergod theorem implies that

$$\lim_{T\to\infty}\frac{1}{\log T}\int_0^{\log T}\mathcal{G}(\mathbf{T}_s(Z(u,\omega))\,ds=\lim_{T\to\infty}\frac{1}{\log T}\int_0^{\log T}\mathcal{G}(Z_s(u,\omega))\,ds$$
$$(3.1)\qquad =E\mathcal{G}(Z(u,\omega))\quad\text{for almost all } \omega\in\Omega.$$

Given a bounded measurable functional $\mathcal{F}$ on the space $C([0,1])$ or $D([0,1])$ let us extend it to the space of all measurable functions on the space $R^{[0,1]}$ of all functions on the interval $[0,1]$ by defining $\mathcal{F}(x)=0$ if the function $x=x(\cdot)$ is not in the space $C([0,1])$ or $D([0,1])$. Then we define the functional $\mathcal{G}=\mathcal{G}(\mathcal{F})$ on the space $R^{(-\infty,\infty)}$ by the formula $\mathcal{G}(z)=\mathcal{F}(x_z)$ with $x_z(u)=u^{1/\alpha}z(\log u)$, $0<u\leqq 1$, $z(0)=0$. We can write

$$\frac{1}{\log T}\int_1^T\frac{1}{t}\mathcal{F}(X_t(\cdot,\omega))\,dt=\frac{1}{\log T}\int_0^{\log T}\mathcal{F}(X_{e^s}(\cdot,\omega))\,ds$$
$$=\frac{1}{\log T}\int_0^{\log T}\mathcal{G}(Z_s(\cdot,\omega))\,ds,$$

since $\mathcal{G}(Z_s(\cdot,\omega))=\mathcal{F}(X_{e^s}(\cdot,\omega))$. Indeed,

$$x_{Z_s(\cdot,\omega)}(u)=u^{1/\alpha}Z_s(\log u,\omega)=u^{1/\alpha}Z(s+\log u,\omega)=u^{1/\alpha}\frac{X(e^{s+\log u},\omega)}{e^{(s+\log u)/\alpha}}$$
$$=\frac{X(ue^s,\omega)}{e^{s/\alpha}},\qquad\text{for all } 0\leqq u\leqq 1,$$

hence $x_{Z_s(\cdot,\omega)}=X_{e^s}(\cdot,\omega)$, where $X_s(\cdot,\omega)$ was defined in (2.4). This relation (with the choice $s=0$) implies in particular that

$$E\mathcal{G}(Z(\cdot,\omega))=E\mathcal{G}(Z_0(\cdot,\omega))=E\mathcal{F}(X_1(\cdot,\omega)).$$

These identities together with relation (3.1) and the definition of the measures $\mu_T(\omega)$ introduced in the formulation of Theorem 1 imply that

$$
\begin{aligned}
&\lim_{T\to\infty}\int \mathcal{F}(x)\,d\mu_T(\omega)(x)\\
&=\lim_{T\to\infty}\frac{1}{\log T}\int_1^T\frac{1}{t}\mathcal{F}(X_t(\cdot,\omega))\,dt\\
&=\lim_{T\to\infty}\frac{1}{\log T}\int_0^{\log T}\mathcal{G}(Z_s(\cdot,\omega))\,ds=E\mathcal{G}(Z(\cdot,\omega))\\
&=E\mathcal{F}(X_1(\cdot,\omega))=\int\mathcal{F}(x)\,d\mu_0(x)\quad\text{for almost all }\omega\in\Omega.
\end{aligned}
\tag{3.2}
$$

To prove Theorem 1 we have to show that relation (3.2) holds simultaneously for all bounded and continuous functionals $\mathcal{F}$ for almost all $\omega\in\Omega$, and the exceptional set of $\omega\in\Omega$ of measure zero should not depend on the functional $\mathcal{F}$. We prove this with the help of the following

Lemma A. *Under the conditions of Theorem 1 the closure of the set of (random) measures $\mu_T(\omega)$, $T\geqq 1$, are compact in the topology defining weak convergence of probability measures in the space $C([0,1])$ or $D([0,1])$ (depending on where the distribution of the process $X(\cdot,\omega)$ is defined) for almost all $\omega\in\Omega$.*

Proof of Lemma A. We apply the result that a set of probability measures μ_T on a separable complete metric space (endowed with the topology inducing weak convergence) is compact if and only if for all $\varepsilon>0$ there is a compact set $\mathbf{K}=\mathbf{K}(\varepsilon)$ on the metric space such that $\mu_T(\mathbf{K})\geqq 1-\varepsilon$ for all measures μ_T. Both spaces $C([0,1])$ and $D([0,1])$ can be endowed with a metric which turns them to a separable complete metric space. (See e.g. Theorems 6.1 and 6.2, 14.1 in Billingsley's book [1].) Because of these results the following statement has to be proved. For almost all $\omega\in\Omega$ and all $\varepsilon>0$ there exists a compact set $\mathbf{K}=\mathbf{K}(\varepsilon,\omega)$ in the space $C([0,1])$ or $D([0,1])$ such that $\mu_T(\omega)(\mathbf{K})\geqq 1-\varepsilon$ for all $T\geqq 1$. In the proof we shall apply formula (3.2) which is valid for all bounded and measurable functionals $\mathcal{F}$ and some classical results which describe the compact sets in $C([0,1])$ and $D([0,1])$. These results can be found for instance in the book of Billingsley [1]. (Theorem 8.2 gives a description of compact sets in $C([0,1])$ and Theorem 14.4 a description of compact sets in $D([0,1])$.)

Let us first consider the case when the distributions of the processes $X_T(\cdot,\omega)$ defined in formula (2.4) are in the $C([0,1])$ space. We shall prove that for almost all $\omega\in\Omega$ and all $\varepsilon>0$ and $\eta>0$ there exist some numbers

$K = K(\varepsilon, \omega)$ and $\delta = \delta(\varepsilon, \eta, \omega) > 0$ such that

$$\mu_T(\omega)\left(x \in C([0,1]) \colon \sup_{0 \leqq u \leqq 1} |x(u)| \geqq K\right) \leqq \varepsilon, \quad \text{and}$$

(3.3)

$$\mu_T(\omega)\left(x \in C([0,1]) \colon |w_x(\delta)| \geqq \eta\right) \leqq \varepsilon,$$

for all $T \geqq 1$, where $w_x(\delta) = \sup_{|t-s| \leqq \delta} |x(t) - x(s)|$ for a function $x \in C([0,1])$. First we show that relation (3.3) implies that for almost all $\omega \in \Omega$ and all $T \geqq 1$ and $\varepsilon > 0$ there exists a compact set $\mathbf{K}(\varepsilon) = \mathbf{K}(\varepsilon, \omega) \subset C([0,1])$ for which $\mu_T(\omega)(\mathbf{K}(\varepsilon)) \geqq 1 - \varepsilon$. Indeed, let us fix some $\varepsilon > 0$, and consider the sets

$$\mathbf{J}_0 = \left(x \in C([0,1]) \colon \sup_{0 \leqq u \leqq 1} |x(u)| > K\right)$$

and

$$\mathbf{J}_n = \left(x \in C([0,1]) \colon |w_x(\delta_n)| > 2^{-n}\varepsilon\right), \quad n = 1, 2, \dots$$

with such constants $K = K(\varepsilon, \omega)$ and $\delta_n = \delta_n(\varepsilon, \omega)$ for which $\mu_T(\omega)(\mathbf{J}_n) \leqq \varepsilon 2^{-n-1}$, $n = 0, 1, \dots$, $T \geqq 1$. Such sets $\mathbf{J}_n$ really exist because of relation (3.3). (The numbers K and δ_n in the definition of the sets $\mathbf{J}_n$ and thus the sets $\mathbf{J}_n$ may depend on ω.) Define the set $\mathbf{K}(\varepsilon) = \bigcap_{n=0}^{\infty} \bar{\mathbf{J}}_n$, where $\bar{\mathbf{J}}$ is the complement of the set $\mathbf{J}$. Then $\mathbf{K}(\varepsilon)$ is a compact set in $C([0,1])$, and for almost all ω and $T \geqq 1$ $\mu_T(\omega)(\mathbf{K}(\varepsilon)) \geqq 1 - \varepsilon$. Applying this result for all $\varepsilon_n = 2^{-n}$, $n = 1, 2, \dots$, we get a set of $\bar{\Omega}$ of probability one, such that for all $\omega \in \bar{\Omega}$, $T \geqq 1$ and $\varepsilon > 0$ there exists a compact set $\mathbf{K}(\varepsilon) = \mathbf{K}(\varepsilon, \omega)$ such that $\mu_T(\omega)(\mathbf{K}(\varepsilon)) \geqq 1 - \varepsilon$. In such a way we reduced the proof of Lemma A in the case of continuous trajectories $X(\cdot, \omega)$ to the proof of relation (3.3).

To prove formula (3.3) we shall apply relation (3.2) with appropriate functionals $\mathcal{F}_1$ and $\mathcal{F}_2$ on the space $C([0,1])$. Put

$$\mathcal{F}_1(x) = \mathcal{F}_{1,K}(x) = I\left(\sup_{0 \leqq u \leqq 1} |x(u)| \geqq K\right)$$

and

$$\mathcal{F}_2(x) = \mathcal{F}_{2,\delta,\eta}(x) = I\left(\sup_{s,t \in [0,1] \colon |t-s| \leqq \delta} |x(s) - x(t)| \geqq \eta\right)$$

with appropriate constants $K > 0$, $\eta > 0$ and $\delta > 0$. For fixed $\varepsilon > 0$ and $\eta > 0$ the constants $K = K(\varepsilon) > 0$ and $\delta = \delta(\varepsilon, \eta) > 0$ can be chosen in such a way that $E\mathcal{F}_1(X_1(\cdot,\omega))<\varepsilon^2$ and $E\mathcal{F}_2(X_1(\cdot,\omega))<\varepsilon^2$. Then, because of formula (3.2) for almost all $\omega\in\Omega$ there exists such a threshold $T_0=T_0(\omega)$ for which $\int \mathcal{F}_i(x)\,d\mu_T(\omega)(x)\leqq\varepsilon$ for all $T\geqq T_0(\omega)$ and $i=1,2$. Since $\mathcal{F}_i(x)=0$ or $\mathcal{F}_i(x)=1$, $i=1,2$, this relation implies that $\mu_T(\omega)(x\colon \mathcal{F}_i(x)\neq 0)\leqq\varepsilon$, for $T\geqq T_0(\omega)$, $i=1,2$. This means that relation (3.3) holds for $T\geqq T_0(\omega)$. Furthermore, since $X_{aT}(u,\omega)=a^{-1/\alpha}X_T(au,\omega)$ for all $0<a\leqq 1$,

$$\mu_t(\omega)\left(x\colon \sup_{0\leqq u\leqq 1}|x(u)|\geqq K\right)\leqq\mu_{T_0(\omega)}(\omega)\left(x\colon \sup_{0\leqq u\leqq 1}|x(u)|\geqq KT_0(\omega)^{-1/\alpha}\right),$$

and

$$\begin{aligned}&\mu_t(\omega)\,(x\in C([0,1])\colon |w_x(\delta)|\geqq\eta)\\ &\leqq\mu_{T_0(\omega)}(\omega)\left(x\in C([0,1])\colon |w_x(\delta T_0(\omega))|\geqq\eta T_0(\omega)^{-1/\alpha}\right)\end{aligned}$$

if $1\leqq t\leqq T_0(\omega)$. These probabilities can be taken small by choosing a sufficiently large $K>0$ and sufficiently small $\delta>0$ which depend only on $T_0(\omega)$. Hence relation (3.3) holds not only for $T\geqq T_0(\omega)$ but also for all $T\geqq 1$ with a possible modification of the constants $\delta(\varepsilon,\eta,\omega)$ and $K(\omega)$ in it.

The proof in the case when the processes $X_T(\cdot,\omega)$ defined in (2.4) take their values in the space $D([0,1])$ is similar, hence we only indicate the necessary modifications. Because of the description of compact sets in the space $D([0,1])$ (found for instance in Theorem (14.4) in Billingsley's book [1]) we can reduce the proof of Lemma A in this case, by a natural modification of the argument presented after the formulation of formula (3.3), to the following modified version of relation (3.3): For all $\varepsilon>0$ and $\eta>0$ there exist some $K>0$ and $\delta>0$ such that

$$\begin{aligned}&\mu_T(\omega)\left(x\in D([0,1])\colon \sup_{0\leqq u\leqq 1}|x(u)|\geqq K\right)\leqq\varepsilon,\\ &\mu_T(\omega)\left(x\in D([0,1])\colon |w''_x(\delta)|\geqq\eta\right)\leqq\varepsilon,\\ &\mu_T(\omega)(x\in D([0,1])\colon w_x[0,\delta)\geqq\eta)\leqq\varepsilon\\ &\mu_T(\omega)(x\in D([0,1])\colon w_x[1-\delta,1)\geqq\eta)\leqq\varepsilon\end{aligned}\tag{3.3'}$$

for all $T\geqq 1$, where

$$w''_x(\delta)=\sup_{0\leqq t_1\leqq t\leqq t_2,|t_2-t_1|\leqq\delta}\min\{|x(t)-x(t_1)|,|x(t_2)-x(t)|\},$$

and $w_x[a,b)=\sup_{a\leqq s,t<b}|x(t)-x(s)|$ for all numbers $0\leqq a<b\leqq 1$.

The proof of formula (3.3′) is similar to that of formula (3.3). Let us introduce the functionals

$$\mathcal{F}_1(x)=I\left(\sup_{0\leqq t\leqq 1}|x(t)|\geqq K\right),\quad \mathcal{F}_2(x)=I\left(w''_x(\delta)\geqq\eta\right),$$
$$\mathcal{F}_3(x)=I\,(w_x[0,\delta)\geqq\eta)\quad\text{and } \mathcal{F}_4(x)=I\,(w_x[1-\delta,1)\geqq\eta)$$

on the space $D([0,1])$, where the constants $K=K(\varepsilon)$ and $\delta=\delta(\varepsilon,\eta)$ will be appropriately chosen. Let us observe that with their appropriate choice we can achieve that $E\mathcal{F}_i(X(\cdot,\omega))\leqq\varepsilon^2$ for $i=1,2,3,4$. To see this it is enough to observe that for all $x\in D([0,1])$ $\sup\limits_{0\leqq t\leqq 1}|x(t)|<\infty$, $\lim\limits_{\delta\to 0}w''_x(\delta)=0$ (see e.g. formulas (14.8) and (14.46) in Billingsley's book [1]), $\lim\limits_{\delta\to 0}w_x[0,\delta)=0$ and $\lim\limits_{\delta\to 0}w_x[1-\delta,1)=0$. These functionals $\mathcal{F}_i$ take values 0 and 1, and formula (3.3′) can be proved similarly to (3.3) with the help of relation (3.2). In such a way Lemma A is proved.

Now we turn back to the proof of Theorem 1. We prove with the help of Lemma A, formula (3.2) and a compactness argument that for almost all $\omega\in\Omega$ the sequence of measures $\mu_T(\omega)$ converges weakly to μ_0 as $T\to\infty$. First we show that for all $\varepsilon_0>0$ and $\varepsilon>0$ there exists a set $\Omega_0=\Omega_0(\varepsilon_0,\varepsilon)\subset\Omega$ and a compact set $\mathbf{K}=\mathbf{K}(\varepsilon_0,\varepsilon)$ in $C([0,1])$ or $D([0,1])$ such that $P(\Omega_0)\geqq 1-\varepsilon_0$ and $\mu_T(\omega)(\mathbf{K})\geqq 1-\varepsilon$ for all $\omega\in\Omega_0$ and $T\geqq 1$. This can be deduced from formulas (3.3) in the space $C([0,1])$ and from formula (3.3′) in the space $D([0,1])$ by an argument similar to the proof of the compactness of the measures $\mu_T(\omega)$ by means of these relations. Thus for instance in the space $C([0,1])$ we define the sets $\mathbf{J}_n$, $n=1,2,\dots$, and $\mathbf{K}=\mathbf{K}(\varepsilon)$ similarly to the definition given after formula (3.3) with the only difference that in this case the numbers K and δ_n appearing in the definition of the sets $\mathbf{J}_n$ are chosen independently of ω in such a way that $P(\{\omega\colon \mu_T(\omega)(\mathbf{J}_n)\leqq\varepsilon 2^{-n-1}$ for all $T\geqq 1\})\geqq 1-\varepsilon_0 2^{-n-1}$. The argument in the case of the $D([0,1])$ space with the help of relation (3.3′) is similar.

For a large number $L>0$ let $\mathbf{F}(L)$ denote the class of continuous and bounded functionals $\mathcal{F}$ on the space $C([0,1])$ or $D([0,1])$ such that $|\mathcal{F}(x)|\leqq L$ for all $x\in C([0,1])$ or $x\in D([0,1])$. Fix an $\varepsilon_0>0$ and $\varepsilon>0$, and choose a set $\Omega_0\subset\Omega$ and a compact set $\mathbf{K}=\mathbf{K}(\varepsilon_0,\varepsilon,L)$ in such a way that $P(\Omega_0)\geqq 1-\varepsilon_0$ and $\mu_T(\omega)(\mathbf{K})\geqq 1-\dfrac{\varepsilon}{L}$ for all $\omega\in\Omega_0$ and $T\geqq 1$. Fix two small numbers $\eta>0$ and $\delta>0$, and let the set $\mathbf{F}(L,\varepsilon_0,\varepsilon,\eta,\delta)\subset\mathbf{F}(L)$ consist of those functionals $\mathcal{F}\in\mathbf{F}(L)$ for which $\sup\limits_{x,y\in\mathbf{K},\,\rho(x,y)\leqq\delta}|\mathcal{F}(x)-\mathcal{F}(y)|\leqq\eta$. For all $\delta>0$ fix a finite δ-net in the compact set $\mathbf{K}$ corresponding to it, i.e. a finite set $\mathbf{J}_\delta=\{x_1,\dots,x_r\}\subset\mathbf{K}$ such that for all $x\in\mathbf{K}$ $\min\limits_{1\leqq s\leqq r}\rho(x,x_s)\leqq\delta$. Such a δ-net really exists because of the compactness of the set $\mathbf{K}$.

Consider the above fixed numbers $\varepsilon_0>0$, $\varepsilon>0$ and $L>0$, together with the sets Ω_0 and $\mathbf{K}$ corresponding to them. First we show that there exists an $\Omega'_0\subset\Omega_0$ such that $P(\Omega_0\setminus\Omega'_0)=0$, and

$$\limsup_{T\to\infty}\left|\int\mathcal{F}(x)\mu_T(\omega)(\,dx)-\int\mathcal{F}(x)\mu_0(\,dx)\right|<\varepsilon \tag{3.4}$$
$$\text{for all } \mathcal{F}\in\mathbf{F}(L)\quad\text{and } \omega\in\Omega'_0.$$

To prove relation (3.4) let us first observe that because of the uniform continuity of the functionals $\mathcal{F}\in\mathbf{F}_L$ on the compact set $\mathbf{K}$ the relation

$$\bigcup_{n=1}^{\infty}\mathbf{F}\left(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n}\right)=\mathbf{F}(L) \tag{3.5}$$

holds for all fixed $\varepsilon_0>0$, $\varepsilon>0$, $\eta>0$ and $L>0$.

Put $\delta=\frac{1}{n}$, consider the $\frac{1}{n}$-net $\mathbf{J}_{1/n}=\{x_1,\dots,x_r\}$ corresponding to it, and make a partition of the set $\mathbf{F}\left(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n}\right)$ into subclasses

$$\mathbf{F}\left(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n},j(1),\dots,j(r)\right)$$

with integers $|j(s)|\leqq(L+1)\eta^{-1}$, $s=1,\dots,r$, which consist of those functionals $\mathcal{F}\in\mathbf{F}\left(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n}\right)$ for which $\mathcal{F}(x_s)\in[j_s\eta,(j_s+1)\eta)$, $s=1,\dots,r$. If $\mathcal{F}_1$ and $\mathcal{F}_2$ belong to the same subclass $\mathbf{F}(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n},j(1),\dots,j(r))$, then $|\mathcal{F}_1(x)-\mathcal{F}_2(x)|<2\eta$ for all $x\in\mathbf{K}$ because of the module of continuity of these functionals on the set $\mathbf{K}$, and because of the relation $\mu_T(\omega)(\mathbf{K})\geqq 1-\dfrac{\varepsilon}{L}$ for all $\omega\in\Omega_0$,

$$\left|\int\mathcal{F}_1(x)\mu_T(\omega)(\,dx)-\mathcal{F}_2(x)\mu_T(\omega)(\,dx)\right|<\varepsilon+2\eta.$$

Let us choose an arbitrary functional $\mathcal{F}$ from all non-empty sets

$$\mathbf{F}(L,\varepsilon_0,\varepsilon,\eta,\frac{1}{n},j(1),\dots,j(r)).$$

We get by applying formula (3.2) for these functionals $\mathcal{F}$ and the previous estimation a weakened version of relation (3.4) on a set $\omega\in\Omega_0''(n)\subset\Omega_0$ such that $P(\Omega_0\setminus\Omega_0''(n))=0$, where $\mathbf{F}(L)$ is replaced by $\mathbf{F}\left(L,\varepsilon_0,\varepsilon.\eta,\frac{1}{n}\right)$, and the upper bound ε by $\varepsilon+2\eta$. Then we get, by applying this relation for all $n=1,2,\dots$ together with relation (3.5) the weakened version of (3.4) for all $\omega\in\bigcap_{n=1}^{\infty}\Omega_0''(n)$ and $\mathcal{F}\in\mathbf{F}(L)$ with upper bound $\varepsilon+2\eta$ instead of ε. Finally, we get formula (3.4) in its original form by letting $\eta\to 0$.

It is not difficult to see that relation (3.4) implies the weak convergence $\mu_T(\omega)$ to μ_0 for almost all $\omega\in\Omega$. Indeed, let us fix a number $L>0$ and $\varepsilon>0$. Then we get, by applying relation (3.4) for all $\varepsilon_0(n)=n^{-1}$, $n=1,2,\dots$ that there exists a set $\Omega_0(n)$, $P(\Omega_0(n))=1-\dfrac{1}{n}$, such that relation (3.4) holds for all $\omega\in\Omega_0(n)$. This implies that relation (3.4) holds for all $\omega\in\Omega=\bigcup_{n=1}^{\infty}\Omega_0(n)$, i.e. on a set of probability 1. Then, since relation (3.4) holds for all $L>0$ and

$\varepsilon > 0$ with probability 1 we get by letting $L \to \infty$ and $\varepsilon \to 0$ in this relation that the sequences of measures $\mu_T(\omega)$ converge weakly to the measure μ_0 for almost all $\omega \in \Omega$.

To complete the proof of Theorem 1 still we have to show that in the case of a Wiener or a stable process the generalized Ornstein–Uhlenbeck process corresponding to it is ergodic. This follows from a natural modification of the zero–one law for sums of independent identically distributed random variables to processes with independent and stationary increments which can be found for instance in Feller's book [4], Chapter 4, Section 7, Theorem 3. The continuous time version of this result which can be proved similarly, also holds. It states that if $X(t)$, $t \geqq 0$, is a stable process with some parameter α, $0 < \alpha \leqq 2$, and a set $\mathbf{A}$ is measurable with respect to the (tail) σ-algebra $\mathcal{F}$ which is the intersection $\mathcal{F} = \bigcap\limits_{T>0} \mathcal{F}_T$, where $\mathcal{F}_T = \sigma\{X(t,\cdot)\colon t \geqq T\}$, then $\mathbf{A}$ has probability zero or one. The same result holds if the set $\mathbf{A}$ is measurable with respect to the σ-algebra $\bigcap\limits_{T>0} \mathcal{F}'_T$, where $\mathcal{F}'_T = \sigma\{X(t,\cdot)\colon t \leqq T\}$. (This result follows for instance from the observation that $t^{-2/\alpha} X\left(\frac{1}{t}, \omega\right)$ is also a stable process.) These relations are equivalent to the statement that the generalized Ornstein–Uhlenbeck process $Z(t)$ corresponding to this stable process has trivial σ-algebra at infinity and minus infinity, i.e. all sets which are measurable with respect to the σ-algebra generated by the random variables $t \geqq T$ (or $t \leqq T$) for all $-\infty < T < \infty$ have probability zero or one. This is a property which is actually stronger than the ergodicity of the process.

Proof of Theorem 2. Theorem 2 will be proved by means of formula (3.2) with an appropriately defined functional $\mathcal{F}$ in the space $C([0,1])$ or $D([0,1])$. Let us define the functional $\mathcal{F} = \mathcal{F}_{\varepsilon,\delta}$ with some $\varepsilon > 0$ and $\delta > 0$ as

$$\mathcal{F}_{\varepsilon,\delta}(x) = I\left(\sup_{1-\varepsilon \leqq s,t \leqq 1} \rho(x_s(\cdot), x_t(\cdot)) \geqq \delta\right),$$

where the function x_t is defined in (2.4′), and $\rho(\cdot,\cdot)$ is the metric introduced in Section 2. We claim that under the conditions of Theorem 2

$$\lim_{\varepsilon \to 0} E\mathcal{F}_{\varepsilon,\delta}(X_1(\cdot,\omega)) = 0 \tag{3.6}$$

for all $\delta > 0$.

Let us also observe that by relation (3.2)

$$\lim_{T \to \infty} \mu_T(\omega)\left(\sup_{1-\varepsilon \leqq s,t \leqq 1} \rho(x_s, x_t) > \delta\right) = \lim_{T \to \infty} \int \mathcal{F}_{\varepsilon,\delta}(x)\, d\mu_T(\omega)(x)$$
$$= E\mathcal{F}_{\varepsilon,\delta}(X_1(\cdot,\omega))$$

for all $\varepsilon > 0$ and $\delta > 0$ and almost all ω, where the function x_t was defined in formula (2.4′). Then we get relation (2.6) with the help of formula (3.6), by letting $\varepsilon \to 0$ in the last formula. Hence to prove relation (2.6) it is enough to prove formula (3.6).

If $X_1(\cdot,\omega) \in C([0,1])$, then this relation follows from the observation that for all $\eta > 0$ there is a compact set $\mathbf{K}_\eta$ in $C([0,1])$ such that $P(X_1(\cdot,\omega) \in \mathbf{K}_\eta) \geqq 1-\eta$, and for all $\delta > 0$ there exists an $\varepsilon = \varepsilon(\eta) > 0$ such that $|x(u)-x(v)| < \delta$ if $x \in \mathbf{K}_\eta$, and $|u-v| \leqq \varepsilon$. There is also a constant $L > 0$ such that $\sup\limits_{x\in\mathbf{K}_\eta} |x(u)| \leqq L$. Since these relations hold for all $\delta > 0$ and appropriate $L > 0$ they imply that $\lim\limits_{\varepsilon\to 0} \sup\limits_{x\in\mathbf{K}_\eta, 1-\varepsilon\leqq t\leqq 1} \rho(x_t,x) = 0$. This means that for sufficiently small $\varepsilon > 0$, $\mathcal{F}_{\varepsilon,\delta}(X_1(\cdot,\omega)) = 0$ if $X_1(\cdot,\omega) \in \mathbf{K}_\eta$, i.e. in the case when an event of probability greater than $1-\eta$ occurs. Hence relation (3.6) holds in this case. The situation in the space $D([0,1])$ is more sophisticated. In this case formula (2.5) also has to be applied.

Since all functions $x(t)$ in the space $D([0,1])$ have a limit as $t \to 1-0$ it follows from relation (2.5) that for all $\delta > 0$

$$P\left(\lim_{\varepsilon\to 0} \sup_{1-\varepsilon\leqq t\leqq 1} |X(t,\omega) - X(1,\omega)| \geqq \frac{\delta}{2}\right) = 0.$$

Hence there is a set $\mathbf{K} = \mathbf{K}_\eta$ in the space $D([0,1])$ such that $P(X_1(\cdot,\omega) \in \mathbf{K}) \geqq 1-\eta$, the closure of the set $\mathbf{K}$ is compact, and for all $x \in \mathbf{K}$ $\lim\limits_{\varepsilon\to 0} \sup\limits_{1-\varepsilon\leqq t\leqq 1} |x_t - x| < \frac{\delta}{2}$, where the function x_t was defined in (2.4′). There is a finite $\frac{\delta}{5}$-net in $\mathbf{K}$, i.e. a finite set $\mathbf{J} = \{x^{(1)},\dots,x^{(s)}\}$, $x^{(r)} \in \mathbf{K}$, $r = 1,\dots,s$, in such a way that for all $x \in \mathbf{K}$ there is some $x^{(r)} \in \mathbf{J}$ such that $\rho(x, x^{(r)}) \leqq \frac{\delta}{5}$. Then to prove formula (2.6) it is enough to show that for all $x^{(r)} \in \mathbf{J}$ there is some $\bar\varepsilon > 0$ such that $\rho(x_t^{(r)}, x^{(r)}) \leqq \frac{\delta}{4}$ for all $1-\bar\varepsilon \leqq t \leqq 1$. Indeed, if this statement holds, then for arbitrary $x \in \mathbf{K}$ there is some $x^{(r)} \in \mathbf{J}$ such that $\rho(x, x^{(r)}) \leqq \frac{\delta}{5}$. Then $\rho(x_s,x_t) \leqq \rho(x_s, x_s^{(r)}) + \rho(x_t, x_t^{(r)}) + \rho(x_s^{(r)}, x_t^{(r)})$. Let us also observe that because of the definition of the functions x_t for sufficiently small $\bar\varepsilon > 0$ for all $x \in D([0,1])$, $1-\bar\varepsilon \leqq t \leqq 1$ and $x^{(r)} \in \mathbf{J}$ the inequality $\rho(x_t, x_t^{(r)}) \leqq \frac{5}{4}\rho(x, x^{(r)})$ holds, and $\rho(x_s^{(r)}, x_t^{(r)}) \leqq \rho(x_s^{(r)}, x^{(r)}) + \rho(x_t^{(r)}, x^{(r)})$. The above inequalities imply that $\rho(x_s,x_t) \leqq \delta$ for $1-\bar\varepsilon \leqq s,t \leqq 1$ if $x \in \mathbf{K}$. Hence $\mathcal{F}_{\varepsilon,\delta}(X_1(\cdot,\omega)) = 0$ with $\varepsilon = \bar\varepsilon$ if $X_1(\cdot,\omega) \in \mathbf{K}$. Then formula (3.6) follows from the relation $P(X_1(\cdot,\omega) \in \mathbf{K}) \geqq 1-\eta$.

Thus to complete the proof of formula (2.6) it is enough to show that for an arbitrary function $x \in D([0,1])$ such that $\lim\limits_{u\to 1-0} |x(u) - x(1)| < \frac{\delta}{2}$ the relation $\lim_{\varepsilon\to0}\rho(x_t,x)<\frac{\delta}{2}$ holds. (This relation means in particular that the limit exists.) To prove this relation let us define for all $\frac{1}{2}\leqq t<1$ the mapping $\lambda_t(u)$ of the interval $[0,1]$ into itself as $\lambda_t(u)=tu$ for $0\leqq u\leqq t^*(t)$ with $t^*(t)=1-\sqrt{1-t}$, and define $\lambda_t(u)$ in the remaining interval $(t^*(t),1]$ also linearly, i.e. let $\lambda_t(u)=(\sqrt{1-t}+t)u+1-t-\sqrt{1-t}$ for $t^*(t)\leqq u\leqq 1$. Then $\lim_{t\to1}\sup_{u\neq v}\log\left|\frac{\lambda_t(u)-\lambda_t(v)}{u-v}\right|=0$. Because of the definition of the metric $\rho=d_0$ it is enough to show that

$$\lim_{t\to1}\sup_{0\leqq u\leqq1}|x_t(u)-x(\lambda_t(u))|=\lim_{u\to1}|x(u)-x(1)|<\frac{\delta}{2}.$$

It is known that for an $x\in D([0,1])$ function $\sup_{0\leqq u\leqq1}|x(u)|<\infty$ (see e.g. Billingsley's book [1]). Hence

$$\begin{aligned}\sup_{0\leqq u\leqq t^*(t)}|x_t(u)-x(\lambda_t(u))|&\leqq(t^{-1/\alpha}-1)\sup_{0\leqq u\leqq1}|x(u)|\\&\leqq\text{const.}\,(t^{-1/\alpha}-1)\to0\quad\text{if } t\to1-0.\end{aligned}$$

Similarly, since a function $x\in D([0,1])$ has a right-hand side limit in the point 1, $\sup_{t^*(t)\leqq u<1}|x_t(u)-x(\lambda_t(u))|\to0$ as $t\to1-0$. Finally, in the point $u=1$ $\lambda_t(1)=1$, and $\lim_{t\to1-0}|x_t(1)-x(\lambda_t(1))|=\left|x(1)-\lim_{t\to1-0}x(t)\right|<\frac{\delta}{2}$. These relations imply that $\lim_{t\to1-0}\rho(x_t,x)=\lim_{t\to1-0}|x(t)-x(1)|<\frac{\delta}{2}$. Theorem 2 is proved.

Proof of Lemma 1. We have to prove that for arbitrary $\delta>0$

$$\lim_{\varepsilon\to0}P\left(\sup_{1-\varepsilon\leqq t\leqq1}|X(t,\omega)-X(1,\omega)|>\delta\right)=0.$$

Because of the stationary increment and self-similarity property of the process $X(t,\omega)$ with parameter $\alpha>0$ yields that

$$\begin{aligned}P\left(\sup_{1-\varepsilon\leqq t\leqq1}|X(t,\omega)-X(1,\omega)|>\delta\right)&=P\left(\sup_{0\leqq t\leqq\varepsilon}|X(t,\omega)-X(\varepsilon,\omega)|>\delta\right)\\&=P\left(\sup_{0\leqq t\leqq1}|X(t,\omega)-X(1,\omega)|>\delta\varepsilon^{-1/\alpha}\right).\end{aligned}$$

Then tending with $\varepsilon \to 0$ we get that $\delta\varepsilon^{-1/\alpha} \to \infty$, and the required property holds.

To prove Theorem 3 first we formulate and prove the following technical Lemma:

LEMMA B. *Let $(M, \mathcal{M}, \rho)$ be a separable, complete metric space such that $\mathcal{M}$ is the σ-algebra generated by the open sets of this space. Let two sequences of probability measures μ_N and $\bar{\mu}_N$, $N = 1, 2, \dots$, be given on the space $(M, \mathcal{M}, \rho)$ such that the measures μ_N weakly converge to a probability measure μ_0 on $(M, \mathcal{M}, \rho)$ as $N \to \infty$, and*

$$(3.7) \qquad \liminf_{N\to\infty} (\bar{\mu}_N(\mathbf{F}^\varepsilon) - \mu_N(\mathbf{F})) \geqq 0 \quad \textit{for all closed sets } \mathbf{F} \in \mathcal{M} \textit{ and } \varepsilon > 0,$$

where $\mathbf{A}^\varepsilon = \{x\colon \rho(x, \mathbf{A}) < \varepsilon\}$ denotes the ε-neighbourhood of a set $\mathbf{A} \in \mathcal{M}$. Then the measures $\bar{\mu}_N$ converge weakly to the same limit measure μ_0 as $N \to \infty$. Moreover, condition (3.7) can be slightly weakened. It is enough to assume that it holds for all compact sets $\mathbf{K} \in \mathcal{M}$ and $\varepsilon > 0$.

PROOF OF LEMMA B. The weak convergence of the measures $\bar{\mu}_N$ to μ_0 as $N \to \infty$ is equivalent to the relation $\liminf_{N\to\infty} \bar{\mu}_N(\mathbf{G}) \geqq \mu_0(\mathbf{G})$ for all open sets $\mathbf{G} \in \mathcal{M}$. For all open sets $\mathbf{G} \in \mathcal{M}$ and $\varepsilon > 0$ there exists a compact set $\mathbf{K} = \mathbf{K}_\varepsilon \in \mathcal{M}$ such that $\mathbf{K} \subset \mathbf{G}$ and $\mu_0(\mathbf{K}) \geqq \mu_0(\mathbf{G}) - \varepsilon$. Then there exists some $\eta > 0$ such that also the η-neighbourhood of $\mathbf{K}$ satisfies the relation $\mathbf{K}^\eta \subset \mathbf{G}$. Consider the $\eta/2$ neighbourhood $\mathbf{K}^{\eta/2}$ of the set $\mathbf{K}$. Since $\mathbf{G}$ contains the $\eta/2$ neighbourhood of the closure of $\mathbf{K}^{\eta/2}$, and the measures μ_N converge weakly to the measure μ_0 as $N \to \infty$ we can write with the help of relation (3.7) that $\liminf_{N\to\infty} \bar{\mu}_N(\mathbf{G}) \geqq \liminf_{N\to\infty} \mu_N(\mathbf{K}^{\eta/2}) \geqq \mu_0(\mathbf{K}^{\eta/2}) \geqq \mu_0(\mathbf{G}) - \varepsilon$. Since the last relation holds for all $\varepsilon > 0$ and open sets $\mathbf{G}$, it implies the convergence of the measures $\bar{\mu}_N$ to μ_0 as $N \to \infty$.

To complete the proof of Lemma B let us observe that because of the compactness (convergence) of the measures μ_N in the weak convergence topology for all $\varepsilon > 0$ there is a compact set $\mathbf{K} \in \mathcal{M}$ such that $\mu_N(\mathbf{K}) > 1 - \varepsilon$ for all $N = 1, 2, \dots$. Then for a closed set $\mathbf{F} \in \mathcal{M}$ the set $\mathbf{F} \cap \mathbf{K}$ is also compact, and $\liminf_{N\to\infty} (\bar{\mu}_N(\mathbf{F}^\varepsilon) - \mu_N(\mathbf{F})) \geqq \liminf_{N\to\infty} (\bar{\mu}_N((\mathbf{F} \cap \mathbf{K})^\varepsilon) - \mu_N(\mathbf{F} \cap \mathbf{K})) - \varepsilon \geqq -\varepsilon$. Since this relation holds for all $\varepsilon > 0$, it is enough to assume relation (3.7) for compact sets $\mathbf{K}$.

Now we introduce the notion of good coupling we shall use later and formulate a simple consequence of Lemma B.

DEFINITION of good coupling. Let two sequences of probability measures μ_N and $\bar{\mu}_N$, $N = 1, 2, \dots$, be given on a separable complete metric space $(M, \mathcal{M}, \rho)$, where $\mathcal{M}$ denotes the σ-algebra generated by the topology induced by the metric ρ. These two sequences of measures have a good coupling if for all $\varepsilon > 0$ and $\delta > 0$ there is a sequence of probability measures $P_N^{\varepsilon,\delta}$, $N = 1, 2, \dots$, on the product space $(M \times M, \mathcal{M} \times \mathcal{M}, \bar{\rho})$,

$\bar{\rho}((x_1,y_1),(x_2,y_2))=\rho(x_1,x_2)+\rho(y_1,y_2)$ which satisfies the following properties.

(i) The marginal distributions of $P_N^{\varepsilon,\delta}$ are μ_N and $\bar{\mu}_N$, i.e. $P_N^{\varepsilon,\delta}(\mathbf{A}\times M)=\mu_N(\mathbf{A})$ and $P_N^{\varepsilon,\delta}(M\times\mathbf{A})=\bar{\mu}_N(\mathbf{A})$ for all $\mathbf{A}\in\mathcal{M}$, and $n=1,2,\dots$.

(ii) $\limsup\limits_{N\to\infty} P_N^{\varepsilon,\delta}(\{(x,y)\colon\,\rho(x,y)>\varepsilon\})\leqq\delta$.

COROLLARY OF LEMMA B. *If two sequences of probability measures μ_N and $\bar{\mu}_N$, $N=1,2,\dots$, on a complete separable metric space $(M,\mathcal{M},\rho)$ have a good coupling, and the sequence of measures μ_N converge weakly to a probability measure μ_0, then the measures $\bar{\mu}_N$ converge weakly to the same measure μ_0.*

PROOF OF THE COROLLARY. Fix an $\varepsilon>0$. For all $\delta>0$ we can write

$$\liminf_{N\to\infty}(\bar{\mu}_N(\mathbf{F}^\varepsilon)-\mu_N(\mathbf{F}))\geqq-\limsup_{N\to\infty}P_N^{\varepsilon,\delta}(\{(x,y)\colon\,\rho(x,y)>\varepsilon\})\geqq-\delta.$$

We get the statement of the Corollary by letting $\delta\to0$.

PROOF OF THEOREM 3. We shall prove the weak convergence of the measures $\bar{\mu}_N(\omega)$ for almost all ω with the help of Lemma B with the choice of $\mu_{B_N}(\omega)$ as μ_N and $\hat{\mu}_N(\omega)$ as $\bar{\mu}_N$. Then (for almost all ω) the measures μ_N converge weakly to μ_0, and it is enough to show that for almost all $\omega\in\Omega$

$$\begin{aligned}(3.8)\qquad &\liminf_{N\to\infty}(\hat{\mu}_N(\omega)(\mathbf{F}^\varepsilon)-\mu_{B_N}(\omega)(\mathbf{F}))\geqq0\quad\text{for all closed sets}\\ &\qquad\mathbf{F}\subset D([0,1])\ \text{or}\ \mathbf{F}\subset C([0,1])\ \text{and}\ \varepsilon\geqq0.\end{aligned}$$

Let us recall that for arbitrary measurable set $\mathbf{B}\subset D([0,1])$ (or $\mathbf{B}\subset C([0,1])$)

$$\mu_{B_N}(\omega)(\mathbf{B})=\bar{\lambda}_{B_N}\{s\colon\,s\in[1,B_N],\ X_s(\cdot,\omega)\in\mathbf{B}\}$$

and

$$\begin{aligned}\bar{\mu}_N(\omega)(\mathbf{B})=\bar{\lambda}_{B_N}\{s\colon&\text{ there is some }1\leqq j<k_N\text{ such that}\\ &B_{j,N}\leqq s<B_{j+1,N},\quad\text{and }X_{B_{j,N}}(\cdot,\omega)\in\mathbf{B}\},\end{aligned}$$

where the measure $\bar{\lambda}_T$ was defined in the formulation of Lemma 1.

For a pair of numbers $\varepsilon>0$ and $\eta>0$ define the set

$$\mathbf{A}(\varepsilon,\eta)=\left\{x\in D([0,1])\colon\sup_{1-\varepsilon<s\leqq t\leqq1,|s-t|\leqq\eta}\rho(x_s,x_t)\leqq\varepsilon\right\}.$$

Given some $\varepsilon>0$ and $\delta>0$ fix some $\eta=\eta(\omega,\varepsilon,\delta)>0$ and $N_0=N_0(\omega,\varepsilon,\delta)$ in such a way that $\mu_{B_N}(\omega)(\mathbf{A}(\varepsilon,\eta))>1-\delta$ for $N\geqq N_0$. By Theorem 2 such a choice of η and N_0 is possible for almost all $\omega\in\Omega$. Then we can choose, since the numbers $B_{k,j}$ satisfy condition (2.7), some number $j_0=j_0(\eta)$ and

$N_1 \geqq N_0$ in such a way that $\dfrac{B_{k+1,N}}{B_{k,N}} \leqq 1+\frac{\eta}{2}$, if $N \geqq N_1$ and $j_0 \leqq k < N$, and $\dfrac{\log B_{j_0,N}}{\log B_N} < \delta$ if $N \geqq N_1$. Then for all $N \geqq N_1$

$$\begin{aligned}\hat\mu_N(\omega)(\mathbf{F}^\varepsilon) &\geqq \hat\mu_N(\omega)(X_{B_{k,N}}(\cdot,\omega)\in \mathbf{F}^\varepsilon,\ \text{for some } k\geqq j_0)\\ &= \bar\lambda_{B_N}(\{s\colon \text{there is some } j_0\leqq j<k_N \text{ such that}\\ &\qquad\qquad B_{j,N}\leqq s< B_{j+1,N} \text{ and } X_{B_{j,N}}(\cdot,\omega)\in\mathbf{F}^\varepsilon\})\\ &\geqq \bar\lambda_{B_N}(\{s\colon B_{j_0,N}\leqq s<B_N \text{ and } X_s(\cdot,\omega)\in \mathbf{F}\cap\mathbf{A}(\varepsilon,\eta)\}).\end{aligned}$$

The last inequality in this relation holds, because, in the case when $X_s(\cdot,\omega)\in \mathbf{F}\cap\mathbf{A}_N$ and $s\in[B_{j,N},B_{j+1,N})$ with some $j_0\leqq j<k_N$ (observe that the relation $[B_{j_0,N},B_N)=\bigcup\limits_{j=j_0}^{k_N-1}[B_{j,N},B_{j+1,N})$ holds), then $X_{B_{j,N}}(\cdot,\omega)\in\mathbf{F}^\varepsilon$, and this implies that all points $s\in(B_{j,N},B_{j+1,N}]$ are contained in the set whose $\bar\lambda_T$ measure is considered in the previous expression. To see the validity of this statement observe that with the notation $x=X_s(\cdot,\omega)$, $x\in D([0,1])$ $X_{B_{j,N}}(\cdot,\omega)=x_u$ with $u=\dfrac{B_{j,N}}{s}$, which satisfies the inequality $1-\eta\leqq\dfrac{1}{1+\frac{\eta}{2}}\leqq u\leqq 1$, where the function x_u is defined in formula (2.4′). Hence $x\in\mathbf{A}(\varepsilon,\eta)\cap\mathbf{F}$ implies that $x_u\in\mathbf{F}^\varepsilon$, as we claimed. Then we get that

$$\begin{aligned}\hat\mu_N(\omega)(\mathbf{F}^\varepsilon) &\geqq \bar\lambda_{B_N}(s\colon s\in[1,B_N),\ \text{and } X_s(\cdot,\omega)\in\mathbf{F})\\ &\qquad -\bar\lambda_{B_N}([1,B_{j_0,N}))-\mu_{B_N}(D([0,1])\setminus\mathbf{A}(\varepsilon,\eta))\\ &\geqq \bar\lambda_{B_N}(s\colon s\in[1,B_N),\ \text{and } X_s(\cdot,\omega)\in\mathbf{F})-2\delta=\mu_{B_N}(\mathbf{F})-2\delta,\end{aligned} \tag{3.9}$$

because $\mu_{B_N}(D([0,1])\setminus\mathbf{A}(\varepsilon,\eta))\leqq\delta$ and

$$\bar\lambda_{B_N}([1,B_{j_0,N}))=\frac{1}{\log B_N}\int_1^{B_{j_0,N}}\frac{1}{t}\,dt=\frac{\log B_{j_0,N}}{\log B_N}\leqq\delta.$$

Letting $\delta\to 0$ in formula (3.9) we get formula (3.8). This implies the first part of Theorem 3.

We prove the second statement of Theorem 3 with the help of the Corollary of Lemma B, where $\hat\mu_N(\omega)$ plays the role of μ_N and $\bar\mu_N(\omega)$ the role of $\bar\mu_N$. We define the measure $P_N^\varepsilon=P_N(\omega)$ on the space $D([0,1])\times D([0,1])$ independently of the parameter ε in the following way: The measure $P_N(\omega)$ is concentrated on the trajectories $(X_{B_j,N}(\cdot,\omega),\bar X_{B_j,N}(\cdot,\omega))$, and

$$P_N(\omega)(X_{B_j,N}(\cdot,\omega),\bar X_{B_j,N}(\cdot,\omega))=\frac{1}{\log B_N}\log\frac{B_{j+1,N}}{B_{j,N}}.$$

Such a coupling can be constructed e.g. in the following way: For all $N = 1, 2, \dots$ let A_N denote the set $\mathbf{A}_N = \{1, \dots, k_N\}$, $\mathcal{A}_N$ the σ-algebra consisting of all subsets of $\mathbf{A}_N$, and define the probability measure ν_N, $\nu_N(j) = \frac{1}{\log B_N} \log \frac{B_{j+1,N}}{B_{j,N}}$, $1 \leqq j < k_N$ on $(\mathbf{A}_N, \mathcal{A}_N)$. Then for all $\omega \in \Omega$ define the random variable $\xi_\omega(j) = (X_{B_j,N}(\cdot,\omega), \bar{X}_{B_j,N}(\cdot,\omega))$, $1 \leqq j \leqq k_N$, on the probability space $(\mathbf{A}_N, \mathcal{A}_N, \nu_N)$, and let $P_N(\omega)$ be the distribution of the random variables ξ_ω in the space $D([0,1]) \times D([0,1])$.

The marginal distributions of the measures $P_N(\omega)$ are $\hat{\mu}_N(\omega)$ and $\bar{\mu}_N(\omega)$. Hence by Corollary of Lemma B it is enough to prove that for almost all ω the relation

$$\lim_{N\to\infty} P_N(\omega)(\mathbf{A}_N(\varepsilon,\omega)) = 0 \tag{3.10}$$

holds with

$$\mathbf{A}_N(\varepsilon,\omega) = \left\{(X_{B_j,N}(\cdot,\omega), \bar{X}_{B_j,N}(\cdot,\omega)) \colon \rho(X_{B_j,N}(\cdot,\omega), \bar{X}_{B_j,N}(\cdot,\omega)) > \varepsilon\right\}$$

for all $\varepsilon > 0$. Since the measures $\hat{\mu}_N$ are compact for all $\eta > 0$ there is a compact set $\mathbf{K} = \mathbf{K}(\eta) \subset D([0,1])$ such that $\hat{\mu}_N(\mathbf{K}) > 1 - \eta$ for all $N = 1, 2, \dots$, and formula (3.10) can be reduced to the statement

$$\lim_{N\to\infty} P_N(\omega)(\mathbf{A}_N(\varepsilon,\omega) \cap (\mathbf{K} \times D([0,1]))) = 0 \tag{3.11}$$

for arbitrary compact set $\mathbf{K} \subset D([0,1])$. Moreover, this statement can be reduced to a slightly weaker statement. To formulate it let us define for all $\eta > 0$ and $N = 1, 2, \dots$ the number $\bar{j}(N) = \bar{j}(N,\eta)$ as $\bar{j}(N) = \max\{j \colon \log B_{j,N} \leqq \eta \log B_N\}$. Because of condition (2.7) imposed on the numbers $B_{j,k}$ in Theorem 3 $\bar{j}(N) \to \infty$ as $N \to \infty$. Because of the definition of the measures $\hat{\mu}_N(\omega)$ and the number $\bar{j}(N)$ the inequality $\hat{\mu}_N(\omega)\left\{\bigcup_{j\colon j \leqq \bar{j}(N)} X_{B_j,N}(\cdot,\omega)\right\} \leqq \eta$ holds.

Define the set

$$\begin{aligned}\mathbf{A}_N^\eta(\varepsilon,\omega) = \{(X_{B_j,N}(\cdot,\omega), \bar{X}_{B_j,N}(\cdot,\omega)) \colon \bar{j}(N,\eta) \leqq j \leqq k_N, \\ \rho(X_{B_j,N}(\cdot,\omega), \bar{X}_{B_j,N}(\cdot,\omega)) > \varepsilon\}\,.\end{aligned}$$

Then $\hat{\mu}_N(\omega)(\mathbf{A}_N(\varepsilon,\omega) \setminus \mathbf{A}_N^\eta(\varepsilon,\omega)) \leqq \eta$, and relation (3.11) can be reduced to the relation

$$\lim_{N\to\infty} P_N(\omega)(\mathbf{A}_N^\eta(\varepsilon,\omega) \cap (\mathbf{K} \times D([0,1]))) = 0 \tag{3.11'}$$

by letting $\eta \to 0$.

We claim that for an arbitrary compact set $\mathbf{K} \subset D([0,1])$, $\varepsilon > 0$ and $\eta > 0$ there is some $N_0 = N_0(\mathbf{K}, \varepsilon, \eta, \omega)$ such that for all $N \geqq N_0$ and $j \geqq \bar{j}(N)$ the relation $X_{B_{j,N}}(\cdot,\omega)\in\mathbf{K}$ implies that $\rho(X_{B_{j,N}}(\cdot,\omega),\bar X_{B_{j,N}}(\cdot,\omega))<\varepsilon$, hence the set $\mathbf{A}_N^\eta(\varepsilon,\omega)\cap(\mathbf{K}\times D([0,1]))$ is empty for large enough N. This statement clearly implies relation (3.11′).

To prove this statement let us observe that the trajectory $\bar X_{B_{j,N}}(\cdot,\omega)$ is obtained as a discretization of the trajectory $X_{B_{j,N}}(\cdot,\omega)$ of the following type: There is a partition $0=t_{j,0,N}<t_{j,1,N}<\cdots<t_{j,j,N}=1$ of the interval $[0,1]$ such that $\bar X_{B_{j,N}}(t,\omega)=X_{B_{j,N}}(t_{j,l-1,N},\omega)$ if $t_{j,l-1,N}\leqq t<t_{j,l,N}$, $1\leqq l\leqq j$, and $\bar X_{B_{j,N}}(1,\omega)=X_{B_{j,N}}(1,\omega)$. The numbers $t_{j,l,N}$ could be given explicitly as $t_{j,l,N}=\dfrac{B_{l-1,N}}{B_{j,N}}$, but we do not need their explicit form. What we need is the fact that conditions (2.7) and (2.8) imposed on the numbers $B_{j,N}$ imply that $\lim\limits_{\hat\jmath\to\infty}\limsup\limits_{N\to\infty}\sup\limits_{N\geqq j\geqq\hat\jmath}\sup\limits_{1\leqq l\leqq j}(t_{j,l,N}-t_{j,l-1,N})=0$. This relation holds since for all $\eta>0$ there exist some $\hat\jmath_1=\hat\jmath_1(\eta)$, $\hat\jmath_2=\hat\jmath_2(\eta)$ and $N_0=N_0(\eta)$ in such a way that $\dfrac{B_{l,N}}{B_{l-1,N}}\leqq 1+\frac{\eta}{2}$ if $\hat\jmath_1\leqq l\leqq N$ and $N\geqq N_0$, and $\eta B_{\hat\jmath_2,N}\geqq B_{\hat\jmath_1,N}$ if $N\geqq N_0$. Then for all $N\geqq j\geqq\hat\jmath_2$ and $N\geqq N_0$ $t_{j,l,N}-t_{j,l-1,N}\leqq\dfrac{B_{l,N}-B_{l-1,N}}{B_{l,N}}\leqq\eta$ for $j\geqq l\geqq\hat\jmath_1$, and $t_{j,l,N}-t_{j,l-1,N}\leqq\dfrac{B_{\hat\jmath_1,N}}{B_{\hat\jmath_2,N}}\leqq\eta$ if $1\leqq l\leqq\hat\jmath_1$. The width of the partitions considered above tends to zero if $\hat\jmath=\hat\jmath(N)\to\infty$, as we claimed. Indeed, the previous calculations imply that it is less than η for $\hat\jmath\geqq\hat\jmath_2(\eta)$.

We claim that this relation implies that

$$\lim_{N\to\infty}\sup_{j:\ j\geqq\hat\jmath(N),\,X_{j,N}(\cdot,\omega)\in\mathbf{K}}\rho(X_{j,N}(\cdot,\omega),\bar X_{j,N}(\cdot,\omega))=0$$

for all compact sets $\mathbf{K}\subset D([0,1])$, and this relation implies formula (3.11′) and hence the second part of Theorem 3.

Let us define the following function $g(x,\delta)$ for $x\in D([0,1])$ and $\delta>0$:

$$g(x,\delta)=\sup_{\substack{0=t_0<t_1<\cdots<t_s=1\\ t_j-t_{j-1}\leqq\delta,\,j=1,\dots,s}}\rho(x,\bar x_{t_0,\dots,t_s}),\tag{3.12}$$

where

$$\bar x_{t_0,\dots,t_s}(t)=x(t_{j-1})\text{ if }t_{j-1}\leqq t<t_j,\ j=1,\dots,s\quad\text{and}\quad\bar x_{t_0,\dots,t_s}(1)=x(1).$$

We shall prove the following Lemma C which is probably well-known among experts, but whose explicit formulation we did not find in the literature.

LEMMA C. *For all functions $x\in D([0,1])$ $\lim\limits_{\delta\to0}g(x,\delta)=0$. Moreover, for all compact sets $\mathbf{K}\subset D([0,1])$*

$$\lim_{\delta\to0}\sup_{x\in\mathbf{K}}g(x,\delta)=0.$$

Then to finish the proof of Theorem 3 it is enough to show that

$$\lim_{\delta\to 0}\sup_{x\in\mathbf{K}} g(x,\delta)=0$$

for all compact sets $\mathbf{K}\subset D([0,1])$, where the function $g(x,\delta)$ is defined in formula (3.12), and this is the content of Lemma C.

REMARK. Condition (2.8) about the properties of the numbers $B_{j,N}$ in Theorem 3 could have been dropped with the help of a slightly more complicated analysis. We could exploit that an upper bound on $t_{j,l,N}-t_{j,l-1,N}$ only for such triples (j,l,N), $1\leqq l\leqq j\leqq N$ is needed which also satisfy the relation $j\geqq \bar{\jmath}(N,\eta)$, $\bar{\jmath}(N,\eta)=\max\{j\colon\ \log B_{j,N}\leqq\eta\log B_N\}$; and $\lim\limits_{N\to\infty} B_{\bar{\jmath},N}=\infty$ for such $\bar{\jmath}$. But since this is a condition which is automatically satisfied in the cases interesting for us we did not omit it.

PROOF OF LEMMA C. It is known (see e.g. Billingsley's book [1] formulas (14.6) and (14.7)) that for all $\eta>0$ there is some $\alpha=\alpha(\eta)>0$ and a partition $0=u_0<u_1<\dots<u_r=1$ of the interval $[0,1]$ such that for $u_j-u_{j-1}>\alpha$, and $\sup\limits_{1\leqq j\leqq r}\ \sup\limits_{u_{j-1}\leqq s,t<u_j}|x(s)-x(t)|<\eta$. Let us consider an arbitrary partition $0=t_0<t_1<\dots<t_s=1$ of the interval $[0,1]$ such that $\sup\limits_{1\leqq j\leqq s}|t_j-t_{j-1}|<\alpha\eta$. We claim that in this case $\rho(x,\bar{x}_{t_1,\dots,t_s})\leqq\eta$. Since this relation holds for all $\eta>0$, it implies the first statement of Lemma C.

To prove this statement let us consider the partition $0=T_0<T_1<\cdots<T_r$, such that the interval $[T_j,T_{j+1})$ is the union of those intervals $[t_l,t_{l+1})$ for which $t_l\in[u_j,u_{j+1})$. Let $\lambda(\cdot)$ be that mapping of the interval $[0,1]$ into itself which maps the interval $[u_j,u_{j+1})$ linearly to the interval $[T_j,T_{j+1})$. Then $\sup\limits_{0\leqq u\leqq 1}|x(\lambda(u))-\bar{x}_{t_1,\dots,t_s}(u)|\leqq\eta$, and also $\sup\limits_{t\neq s}\log\left|\dfrac{\lambda(t)-\lambda(s)}{t-s}\right|\leqq\eta$. Hence $\rho(x,\bar{x}_{t_1,\dots,t_s})\leqq\eta$, as we claimed. This implies the first statement of Lemma C.

The second, more general statement follows in the same way. We only have to observe that the number $\alpha=\alpha(\eta)$ can be chosen as the same number for all $x\in\mathbf{K}$ in a compact set $\mathbf{K}\in D([0,1])$. This follows from the characterization of compact sets in the space $D([0,1])$. (See relation (14.33) in Theorem 14.3 in the book of Billingsley [1].)

PROOF OF THEOREM 4. Let us construct the following coupling of the random broken lines $\bar{S}_N(\cdot,\omega)$ and $T_N(\cdot,\omega)$ which are defined with the help of the random variables $\bar{S}_k(\omega)$ and $T_k(\omega)$, $k=1,2,\dots$, in formula (2.13). Let $P_N^{\varepsilon,\delta}(\omega)$, $\omega\in\Omega$, be a measure on $D([0,1])\times D([0,1])$ concentrated on the pairs, $(\tilde{S}_k(\cdot,\omega), T_k(\cdot,\omega))$ in such a way that

$$P_N^{\varepsilon,\delta}(\omega)(\tilde{S}_k(\cdot,\omega), T_k(\cdot,\omega)) = \mu_N(\omega)(T_k(\cdot,\omega)) = \frac{1}{\log \frac{B_N}{B_1}} \log \frac{B_{k+1}}{B_k}, \quad 1 \leqq k < N.$$

(The parameters $\varepsilon > 0$ and $\delta > 0$ in the definition are the same ε and δ which appear in formula (2.13).)

Then the marginal distributions of $P_N^{\varepsilon,\delta}(\omega)$ are the distributions $\mu_N(\omega)$ and $\bar{\mu}_N(\omega)$ appearing in the definition of the almost sure functional limit theorem. By the Corollary of Lemma B it is enough to prove that

$$\limsup_{N\to\infty} P_N^{\varepsilon,\delta}(\omega)\{(x,y)\colon \rho(x,y) > \varepsilon\} < \delta$$

for almost all $\omega \in \Omega$. Since $\rho(x,y) \leqq d(x,y)$ with $d(x,y) = \sup\limits_{0\leqq u\leqq 1} |x(u) - y(u)|$,

$$P_N^{\varepsilon,\delta}(\omega)\{(x,y)\colon \rho(x,y) > \varepsilon\} \leqq \frac{2}{\log B_N} \sum_{k=1}^{N-1} \log \frac{B_{k+1}}{B_k} I(d(\tilde{S}_k(\cdot,\omega), T_k(\cdot,\omega)) > \varepsilon)$$

for sufficiently large N. For a number N choose the number $\bar{n} = \bar{n}(N)$ such that $2^{\bar{n}-1} \leqq B_N < 2^{\bar{n}}$. Then $N \leqq N(\bar{n})$, and $\log B_N \geqq \bar{n} - 1$. Hence

$$P_N^{\varepsilon,\delta}(\omega)\{(x,y)\colon \rho(x,y) > \varepsilon\} \leqq \frac{1}{\bar{n}-1} \sum_{k=1}^{N(\bar{n})} \log \frac{B_{k+1}}{B_k} I\left(\left\{ \frac{\sup\limits_{1\leqq j\leqq k} |\bar{S}_j(\omega) - T_j(\omega)|}{A_k} > \varepsilon \right\}\right)$$

with this $\bar{n} = \bar{n}(N)$. As $\bar{n}(N)$ tends to infinity as $N \to \infty$ relation (2.13) implies that the lim sup of the right-hand side of the last expression is less than δ for almost all ω as $N \to \infty$. Theorem 4 is proved.

Proof of Theorem 5. Let $\mu_N(\omega)$ and $\bar{\mu}_N(\omega)$ denote the probability measures on the space $D([0,1])$ defined by the partial sums $S_k(\omega) = \sum\limits_{j=1}^{k} \xi_j(\omega)$ through formulas (2.11) and (2.12) with weight functions B_n, $A_n = B_n^{1/\alpha}$ and $\bar{B}_n$, $\bar{A}_n = \bar{B}_n^{1/\alpha}$, respectively. Let us also introduce the random polygons $S'_n(\omega)$ and the random measures $\bar{\mu}'_N(\omega)$ which are defined with the help of the partial sums $S_k(\omega)$ with the new weight functions $\bar{B}_n$ and the original sequence $A_n = B_n^{1/\alpha}$ by formulas (2.11) and (2.12). First we show the following (weaker) version of Theorem 5. For almost all $\omega \in \Omega$ the sequence

$\bar{\mu}'_N(\omega)$ has the same limit μ_0 if $N\to\infty$ as the sequence of measures $\mu_N(\omega)$. By Lemma B to prove this statement it is enough to show that for arbitrary compact set $\mathbf{K}\subset D([0,1])$ and $\varepsilon>0$

$$\limsup_{N\to\infty}(\bar{\mu}'_N(\omega)(\mathbf{K}^\varepsilon)-\mu_N(\omega)(\mathbf{K}))\geqq 0 \quad \text{for almost all } \omega\in\Omega, \tag{3.13}$$

where $\mathbf{K}^\varepsilon=\{x\colon \rho(x,\mathbf{K})\leqq\varepsilon\}$ is the ε-neighbourhood of the set $\mathbf{K}$. Define for all $k=1,2,\dots,$ and $\omega\in\Omega$ the set

$$\mathbf{A}_k=\mathbf{A}_k(\omega)=\bigcup_{j=1}^{k}\{S_j(\cdot,\omega)\}\subset D([0,1]).$$

Then we have

$$\lim_{N\to\infty}\mu_N(\omega)(\mathbf{A}_k(\omega))=0 \tag{3.14}$$

for arbitrary fixed $k>0$. Hence to prove relation (3.13) it is enough to show that for arbitrary $\varepsilon>0$ there is some $N_0=N_0(\varepsilon)$ such that

$$\rho(S_N(\cdot,\omega),\bar{S}'_N(\cdot,\omega))\leqq\varepsilon$$

if $N\geqq N_0$ and $S_N(\cdot,\omega)\in\mathbf{K}$. Indeed, this relation implies a modified version of (3.13), where $\mu_N(\omega)(\mathbf{K})$ is replaced by $\mu_N(\omega)(\mathbf{K}\setminus\mathbf{A}_{N_0})$. Then relation (3.13) follows from this statement and (3.14) if we let $N_0\to\infty$.

The above statement holds, but we must be careful in its proof. It follows immediately from the conditions of Theorem 5 that

$$\lim_{N\to\infty}d(S_N(\cdot,\omega),\bar{S}_N(\cdot,\omega))=0,$$

if the metric $\rho=d_0$ applied in this paper is replaced by the following metric $d(\cdot,\cdot)$ in the space $D([0,1])$: The relation $d(x,y)\leqq\varepsilon$, $x,y\in D([0,1])$, holds, if there is a strictly monotone increasing continuous function $\lambda(t)$ which is a homeomorphism of the interval $[0,1]$ into itself, and $\sup_{0\leqq t\leqq 1}|\lambda(t)-t|\leqq\varepsilon$, $\sup_{0\leqq t\leqq 1}|y(t)-x(\lambda(t))|\leqq\varepsilon$. The metric d induces the same topology as the metric $\rho=d_0$ in the space $D([0,1])$, but it has the unpleasant property that the space $D([0,1])$ is not a complete metric space with this metric. A detailed discussion about the relation between the metrics $d(\cdot,\cdot)$ and $d_0(\cdot,\cdot)$ is contained in the book of Billingsley [1].

In the proof we have to overcome the following difficulty. The natural transformation $\lambda(\cdot)$ for which $S_N(\lambda(\cdot,\omega))$ is close to $\bar{S}'_N(\cdot,\omega)$ is the map which transforms the point $\frac{B_k}{B_N}$ to the point $\frac{\bar{B}_k}{\bar{B}_N}$, and is linear between these points. This transformation shows that the corresponding trajectories are close in the $d(\cdot,\cdot)$ metric, but it supplies no good estimate for the distance in the $d_0(\cdot,\cdot)$ metric.

We recall the following result from Billingsley's book [1] (see Lemma 2 in Section 14): If $d(x,y)\leqq\delta^2$, $0<\delta\leqq 1/4$, then $\rho(x,y)=d_0(x,y)\leqq 4\delta+w'_x(\delta)$, where the inequality $w'_x(\delta)\leqq\varepsilon$ for a function $x\in D([0,1]$ means that there exist some numbers $0=t_0<t_1<\cdots<t_s=1$ such that $t_j-t_{j-1}\geqq\varepsilon$, and $\sup\limits_{t_{j-1}\leqq u,v<t_j}|x(u)-x(v)|\leqq\varepsilon$ for all $j=1,2,\dots$.

We have $\lim\limits_{\delta\to 0}w'_x(\delta)=0$ for arbitrary $x\in D([0,1])$. Moreover,

$$\lim_{\delta\to 0}\sup_{x\in\mathbf{K}}w'_x(\delta)=0$$

for arbitrary compact set $\mathbf{K}\subset D([0,1])$. (See Theorem 14.3 in Billingsley's book [1].) Given some $\delta>0$ and a compact set $\mathbf{K}\subset D([0,1])$ choose a number $0<\eta<1/4$ such that $5\eta<\delta$ and a number $\bar\eta>0$ such that $w'_x(\bar\eta)<\eta$ if $x\in\mathbf{K}$. Then there is an index $N_0=N_0(\eta,\bar\eta)$ such that $d(S_N(\cdot,\omega),\bar S'_N(\cdot,\omega))\leqq\min(\eta^2,\bar\eta^2)$, if $N\geqq N_0$. The above relations imply that $\rho(S_N(\cdot,\omega),\bar S_N(\cdot,\omega))\leqq 4\min(\eta,\bar\eta)+w'_{S_N(\cdot,\omega)}(\bar\eta)\leqq\delta$, if $N\geqq N_0$ and $S_N(\cdot,\omega)\in\mathbf{K}$. As we have pointed out, relation (3.13) is a consequence of this statement and relation (3.14). This implies the modified version of Theorem 5 with the modified weights $\bar B_n$ and the original weights A_n.

To complete the proof of Theorem 5 in its original form we compare the measures $\bar\mu_N(\omega)$ and $\bar\mu'_N(\omega)$. It is enough to show that for almost all $\omega\in\Omega$ and all $\varepsilon>0$ (and $\delta=\varepsilon$) the sequences of measures $\bar\mu_N(\omega)$ and $\bar\mu'_N(\omega)$ have a good coupling. We make the following coupling of these measures (independently of the parameter ε). Put

$$P_N(\omega)(\bar S_k(\cdot,\omega),\bar S'_k(\cdot,\omega))=\bar\mu_N(S_k(\cdot,\omega))=\frac{1}{\log\dfrac{\bar B_N}{\bar B_1}}\log\frac{\bar B_{k+1}}{\bar B_k},\quad 1\leqq k<N.$$

Observe that $\bar S'_k(\cdot,\omega)=\dfrac{\bar A_k}{A_k}\bar S_k(\cdot,\omega)$, and $\lim\limits_{k\to\infty}\dfrac{\bar A_k}{A_k}=1$. This relation implies that for arbitrary $\delta>0$

$$\lim_{N\to\infty}P_N(\omega)\left(\left\{(S_k(\omega),\bar S_k(\omega))\colon\rho(\bar S_k(\cdot,\omega),\bar S'_k(\cdot,\omega))>\delta\sup_{0\leq t\leq 1}|\bar S_k(\cdot,\omega)|\right\}\right)=0.$$

On the other hand, the measures $\bar\mu_N(\omega)$, $N=1,2,\dots$, are compact for almost all $\omega\in\Omega$, hence for almost all $\varepsilon>0$ there exists a $K=K(\omega)$ such that

$$\bar\mu_N\left(\left\{x\colon\sup_{0\leq t\leq 1}|x(t)|>K\right\}\right)<\varepsilon\quad\text{for all } N=1,2,\dots.$$

Applying the previous estimate with $\delta = \dfrac{\varepsilon}{K}$ we get that

$$\limsup_{N\to\infty} P_N(\omega)(\{(x,y)\colon (x,y) = (\bar S_k(\cdot,\omega), \bar S'_k(\cdot,\omega)),$$
$$\text{with some } 1 \leqq k \leqq N,\ \rho(\bar S_k(\cdot,\omega), \bar S'_k(\cdot,\omega)) > \varepsilon\}) < \varepsilon.$$

Since this relation holds for arbitrary $\varepsilon > 0$ the Corollary of Lemma B implies Theorem 5 in its original form.

ACKNOWLEDGEMENT. I would like to thank István Berkes for many useful discussions on this subject.

REFERENCES

[1] BILLINGSLEY, P., *Convergence of probability measures*, John Wiley & Sons, Inc., New York–London–Sydney–Toronto, 1968. *MR* **38** #1718

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

[3] DOBRUSHIN, R. L., Gaussian and their subordinated self-similar random generalized fields, *Ann. Probab.* **7** (1979), 1–28. *MR* **80e**:60069

[4] FELLER, W., *An introduction to probability theory and its applications*, Vol. II, Second edition, John Wiley & Sons, Inc., New York–London–Sydney–Toronto, 1971. *MR* **42** #5292

[5] FISHER, A., Convex invariant means and a pathwise central limit theorem, *Adv. in Math.* **63** (1987), 213–246. *MR* **88g**:60058

[6] LACEY, M. AND PHILIPP, W., A note on the almost sure central limit theorem, *Statist. Probab. Lett.* **9** (1990), 201–205. *MR* **91e**:60100

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

(Received May 4, 1998)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

major@math-inst.hu

and

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
BOLYAI KOLLÉGIUMA
AMERIKAI ÚT 96
H–1145 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 305–316

A CONDITIONED LAW OF LARGE NUMBERS FOR MARKOV ADDITIVE CHAINS

A. MEDA and P. NEY

To the memory of A. Rényi

Abstract

Let $Y_n = (Y_{n_1}, Y_{n_2})$, $n = 1, 2, \dots$ be a sequence of $\mathbb{R}^d$-valued random variables with $Y_{n_i} \in \mathbb{R}^{d_i}$, $i = 1, 2$, $d_1 + d_2 = d$, and assume that

$$\lim n^{-1} \log E \exp\langle \alpha, Y_n \rangle = \Lambda(\alpha) \leqq \infty, \quad \alpha \in \mathbb{R}^d,$$

exists, and is strictly convex and essentially smooth. ($\langle \cdot, \cdot \rangle$ is inner product.) Then Y_{n_1}/n converges exponentially with respect to the conditional probability measures $\mathbb{P}(\cdot | Y_{n_2}/n \in C \subset \mathbb{R}^{d_2})$, to a point which is specified in terms of Λ and C. This result is specialized to a conditional LLN's for Markov-additive chains.

1. Background

Let $X_1, X_2, \dots,$ be a sequence of independent identically distributed (i.d.d.) random variables taking values in a measurable space $(S, \mathcal{S})$, f a function on S, $\mathrm{S}_n = \sum_{i=1}^{n} f(X_i)$. Stimulated by the important papers of O. Vasicek [16], and especially I. Csiszár [1], a considerable literature has developed on the limit laws of $X_1, \dots, X_k$, conditioned on S_n/n. In [1], S_n/n is represented as the empirical measure of $X_1, \dots, X_n$, and it is shown that $(X_1, \dots, X_n \,|\, \mathrm{S}_n/n \in C)$, where C is a "completely convex" set of probability measures, are asymptotically "quasi-independent" with a limiting measure given by a so-called "I-projection". (These terms are defined in the above paper.) Recently Dembo and Kuelbs [3] have shown that for $f \colon S \to E =$ a separable Banach space, $(X_1, \dots, X_k | \mathrm{S}_n/n \in D)$ with D an open convex set, converge in a strong sense (that implies total variation norm convergence) to independent random variables with identifiable distributions. They allow $k = k(n) \to \infty$ with $k(n)$ depending on D. (Typically $k(n) = o(\frac{n}{\log n})$.) Their analysis used an extension to the Banach space setting (by Einmahl and Kuelbs [5]) of a dominating point construction introduced in the finite

1991 *Mathematics Subject Classification.* Primary 60F10, 60F05; Secondary 60K15.
Key words and phrases. Laws of large numbers, large deviations, Markov-additive chains.

0081–6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

dimensional case by Ney [10]. (See also the use of "exposed points", e.g. in Dembo and Zeitouni [4] p. 44.)

A related conditional law of large numbers (with $X_1, X_2, \ldots$ still i.i.d.), also resting on the dominating point concept, was the work of Nummelin [13], and Lehtonen and Nummelin [8], [9]. Say that a sequence of random variables $\{V_n\}$ converges exponentially to v, written $V_n \xrightarrow{\exp} v$, with respect to a sequence of probability measures P_n if given any $\epsilon > 0$, there exist $0 < b < \infty$ such that

$$P_n\{\|V_n - v\| > \epsilon\} \leqq e^{-bn}$$

for sufficiently large n.

Lehtonen and Nummelin proved that for functions $g : S \to \mathbb{R}^{d_1}$ and $u : S \to \mathbb{R}^{d_2}$ one has

$$n^{-1} \sum_{i=1}^{n} g(X_i) \xrightarrow{\exp} v \tag{1.1}$$

with respect to the conditioned measures

$$P\Big(\cdot \mid n^{-1} \sum_{i=1}^{n} u(X_i) \in C \Big), \tag{1.2}$$

with v being identified as a "dominating point" (as described below).

Extensions of the original set-up of Csiszár [1] to Markov chains have been carried out in Csiszár, Cover and Choi [2], and Schroeder [14]. These extensions are required even in the i.i.d. setting, if one wants to treat functions of the form $n^{-1} \sum_{i=1}^{n} g(X_i, X_{i+1})$, rather than $n^{-1} \sum_{i=1}^{n} g(X_i)$. Independence is lost in this case, but the Markov structure is retained (see e.g. [2]). In this note we will show that results like (1.1), (1.2) hold quite generally. We show first that analogs hold for "general" sequences of random variables satisfying the hypotheses of the Gärtner–Ellis Theorem. We then specialize this result to show how an explicit determination of the limit point in (1.1) can be made in the case of Markov-Additive (MA) chains (which are somewhat more general than ordinary Markov chains, and are defined below). Our conditions on the underlying Markov chain are less restrictive than in the above cited papers.

2. General sequences

For any $\Gamma \subset \mathbb{R}^d$, let Γ^0 = interior of Γ, $\overline{\Gamma}$ = closure of Γ, and $\partial\Gamma$ = boundary of Γ. Let $\langle \cdot, \cdot \rangle$ denote inner product. For any $F : \mathbb{R}^d \to [0, \infty]$, let

$\mathcal{D}(F) = \{x \in \mathbb{R}^d : F(x) < \infty\}$. Let $Y_1, Y_2, \ldots$ be $\mathbb{R}^d$-valued random variables with probability law $\mathcal{L}(Y_n) = \mu_n, n = 1, 2, \ldots$. Assume that

HYPOTHESIS (H1).

(i) $\lim \frac{1}{n} \log E e^{\langle \alpha, Y_n \rangle} = \Lambda(\alpha) \leqq \infty$ exists for $\alpha \in \mathbb{R}^d$, and

(ii) Λ is strictly convex and essentially smooth (see Rockafellar [17] p. 251).

From (ii) it follows that

$$O \in \mathcal{D}^0(\Lambda). \tag{2.1}$$

From Hypothesis (H1) we can draw several conclusions, which we summarize in Lemma 2.1 below. (This is part of the Lemma in Ney [10].) Let $\Lambda^*(v) = \sup_{\alpha \in \mathbb{R}^d} [\langle \alpha, v \rangle - \Lambda(\alpha)] =$ the convex conjugate of Λ. Condition (2.1) implies that the level sets of Λ^*, $L_a(\Lambda^*) = \{v : \Lambda^*(v) \leqq a\}$ are compact for $a \in [0, \infty)$, and is needed in the following lemma. Let ∇ denote gradient.

LEMMA 2.1. *Assume* (H1) *and let B be open and convex with* $[B \cap \mathcal{D}(\Lambda^*)]^0 \neq \emptyset$. *Then*

(i) $\inf[\Lambda^*(v) : v \in B]$ *is achieved at a unique point* $v_B \in \overline{B} \cap \mathcal{D}^0(\Lambda^*)$.

(ii) *The equation*

$$\nabla \Lambda(\alpha) = v_B \tag{2.2}$$

has a solution $\alpha_{v_B} \in \mathbb{R}^d$.

(iii) *If* $\alpha_{v_B} \neq 0$ *then* $v_B \in \partial B$, *and*

$$\langle (v - v_B), \alpha_{v_B} \rangle > 0 \quad \textit{for all } v \in B. \tag{2.3}$$

We call v_B the dominating point of (Λ, B). We will abbreviate $\inf[\Lambda^*(v) : v \in B] = \Lambda^*(B)$, and $\alpha_{v_B} = \alpha_B$. From (H1) we can also conclude that

LEMMA 2.2. *If* (H1) *holds and B is open and convex with* $[B \cap \mathcal{D}(\Lambda^*)]^0 \neq \emptyset$, *then*

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left\{\frac{Y_n}{n} \in B\right\} = -\Lambda^*(B) = -\Lambda^*(v_B). \tag{2.4}$$

This is just the Gärtner–Ellis Theorem specialized to convex sets. (See e.g. Dembo and Zeitouni [4], Theorem 2.3.6.)

With the above background, we can now turn to

LEMMA 2.3 (Conditional weak law of large numbers). *Assume* $\{Y_n\}$ *satisfies* (H1) *and that* $B \subset \mathbb{R}^d$ *is open and convex with* $[B \cap \mathcal{D}(\Lambda^*)]^0 \neq \emptyset$. *Then for all* $\epsilon > 0$

$$\mathbb{P}\left\{\left\|\frac{Y_n}{n} - v_B\right\| > \epsilon \,\middle|\, \frac{Y_n}{n} \in B\right\} \xrightarrow{\exp} 0 \ \textit{as}\ n \to \infty. \tag{2.5}$$

REMARK. Without the condition $Y_n/n \in B$,

$$\mathbb{P}\left\{\left\|\frac{Y_n}{n} - \nabla\Lambda(0)\right\| > \epsilon\right\} \xrightarrow{\exp} 0. \tag{2.6}$$

(See e.g. Ellis [6], Theorem II.6.3 or Dembo and Zeitouni [4].) But under (H1) $\Lambda^*(\nabla\Lambda(0)) = 0$, and $\Lambda^*(v) \geqq 0$ for all v. Hence $\Lambda^*(\nabla\Lambda(0)) = \inf_{v \in \mathbb{R}^d} \Lambda^*(v)$. Now if $\nabla\Lambda(0) \in \overline{B}$, then $\inf_{v \in \mathbb{R}^d} \Lambda^*(v) = \inf_{v \in B} \Lambda^*(v) = v_B = \nabla\Lambda(0)$ in (2.5), and comparing with (2.6) we see that in this case the conditioning in (2.5) has no effect. This should be contrasted with the case $\nabla\Lambda(0) \notin \overline{B}$, in which $v_B \neq \nabla\Lambda(0)$.

PROOF OF LEMMA 2.3. Let $\{\Lambda(\alpha); \alpha \in \mathbb{R}^d\}$ be as defined in (H1)(i). Let $\|x\| = \max\{|x_i|, 1 \leqq i \leqq d\}$ for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, and let

$$B^{(\epsilon)} = B \cap \{v \in \mathbb{R}^d : \|v - v_B\| > \epsilon\}.$$

Let

$$B_i = \{v \in \mathbb{R}^d : (v_i - v_{B,i}) > \epsilon\} \cap B,$$

and

$$B_{d+i} = \{v \in \mathbb{R}^d : -(v_i - v_{B,i}) > \epsilon\} \cap B, \qquad i = 1, \dots, d,$$

where

$$v_{B,i} = \text{the i}^{th} \text{ coordinate of } v_B.$$

Then

$$B^{(\epsilon)} = \bigcup_{i=1}^{2d} B_i.$$

If $[B_i \cap \mathcal{D}(\Lambda^*)]^0 \neq \emptyset$, then by Lemma 2.1 there is a dominating point v_{B_i} with $\Lambda^*(v_{B_i}) = \inf_{v \in B_i} \Lambda^*(v)$. By the uniqueness of the dominating point and the construction of B_i,

$$\Lambda^*(B_i) = \Lambda^*(v_{B_i}) > \Lambda^*(v_B) = \Lambda^*(B).$$

For $[B_i \cap \mathcal{D}(\Lambda^*)]^0 = \emptyset$, since B_i is open and Λ^* is strictly convex, one has $\Lambda^*(B_i) = \infty$, and trivially $\Lambda^*(B_i) > \Lambda^*(v_B) = \Lambda^*(B)$. Hence

$$\Lambda^*(B^{(\epsilon)}) > \Lambda^*(B). \tag{2.7}$$

Then by Lemma 2.2

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big\{ \frac{Y_n}{n} \in B^{(\epsilon)} \;\Big|\; \frac{Y_n}{n} \in B \Big\} \\
&\quad = \limsup_{n\to\infty} \frac{1}{n} \Big[\log \mathbb{P}\Big\{ \frac{Y_n}{n} \in B^{(\epsilon)} \Big\} - \log \mathbb{P}\Big\{ \frac{Y_n}{n} \in B \Big\} \Big] \\
&\quad \leqq \limsup_{n\to\infty} \frac{1}{n} \log \sum_{i=1}^{2d} \mathbb{P}\Big\{ \frac{Y_n}{n} \in B_i \Big\} + \Lambda^*(B) \\
&\quad \leqq - \min_{1\leq i\leq 2d} [\Lambda^*(B_i)] + \Lambda^*(B) \\
&\quad = -\Lambda^*(B^{(\epsilon)}) + \Lambda^*(B) < 0 \quad \text{by Lemma 2.1(i).}
\end{aligned}$$

This implies the Lemma. □

In some applications, as in [9], one wants a limit law for one function, conditioned on another function. We will state a general theorem of this kind. We use the notation that for any $v \in \mathbb{R}^d$, we write $v = (v_1, v_2)$ with $v_1 \in \mathbb{R}^{d_1}$, $v_2 \in \mathbb{R}^{d_2}$, $d_1 + d_2 = d$. Thus $Y_n = (Y_{n_1}, Y_{n_2})$, $Y_{n_1} \in \mathbb{R}^{d_1}$, $Y_{n_2} \in \mathbb{R}^{d_2}$. Also, for $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^d$,

$$\Lambda(\alpha_1, 0) = \Lambda^{(1)}(\alpha_1), \quad \Lambda(0, \alpha_2) = \Lambda^{(2)}(\alpha_2), \tag{2.8}$$

where

$$\Lambda^{(i)}(\alpha_i) = \lim \frac{1}{n} \log E e^{\langle \alpha_i, Y_{n_i} \rangle}, \qquad i = 1, 2.$$

Write

$$\Lambda^{(i)^*}(v_i) = \sup_{\alpha_i \in \mathbb{R}^{d_i}} [\langle \alpha_i, v_i \rangle - \Lambda^{(i)}(\alpha_i)], \quad v_i \in \mathbb{R}^{d_i}.$$

One can check that if $\{Y_n\}$ satisfies (H1) on $\mathbb{R}^d$ then also $\{Y_{n_i}, \Lambda^{(i)}\}$ satisfy (H1) in $\mathbb{R}^{d_i}$. Hence if C is open and convex with $[C \cap \mathcal{D}(\Lambda^{(2)^*})]^0 \neq \emptyset$, then the unique dominating point v_C exists and

$$\nabla \Lambda^{(2)}(\alpha_2) = v_C, \ \alpha_2 \in \mathbb{R}^{d_2}, \ v_C \in \mathbb{R}^{d_2} \tag{2.9}$$

has a solution denoted by $\alpha_C \in \mathbb{R}^{d_2}$.

We can now state

THEOREM 1. *Assume* (H1), *and let* $C \subset \mathbb{R}^{d_2}$ *be open and convex, with* $[C \cap \mathcal{D}(\Lambda^{(2)^*})]^0 \neq \emptyset$. *Let* $v_C \in \mathbb{R}^{d_2}$ *be the dominating point for* $(\Lambda^{(2)}, C)$, *and* α_C *be as defined in* (2.9). *Then*

$$\mathbb{P}\Big\{ \Big\| \frac{Y_{n_1}}{n} - v_1 \Big\| > \epsilon \;\Big|\; \frac{Y_{n_2}}{n} \in C \Big\} \xrightarrow{\exp} 0, \tag{2.10}$$

where

$$v_1 = (\nabla\Lambda(0, \alpha_C))_1 \,. \tag{2.11}$$

REMARK. As in the remark following Lemma 2.3, if $\nabla\Lambda^{(2)}(0) \in \overline{C}$, then the conditioning has no effect. The interesting case is $\nabla\Lambda^{(2)}(0) \notin \overline{C}$.

APPLICATION. Let $\{X_n, n = 1, 2, \dots\}$ be a sequence of random variables on a measurable space $(S, \mathcal{S})$, let $g : S \to \mathbb{R}^{d_1}$, $u : S \to \mathbb{R}^{d_2}$, $f = (g, u) : S \to \mathbb{R}^d$, $Y_n = \sum_{i=1}^{n} f(X_i)$, $Y_{n_1} = \sum_{i=1}^{n} g(X_i)$, $Y_{n_2} = \sum_{i=1}^{n} u(X_i)$. Assume that $\{Y_n\}$ satisfies (H1). Then

$$\frac{1}{n}\sum_{i=1}^{n} g(X_i) \xrightarrow{\text{exp}} (\nabla\Lambda(0, \alpha_C))_1 \tag{2.12}$$

with respect to

$$\mathbb{P}\Big\{ \cdot \,\Big|\, \frac{1}{n}\sum_{i=1}^{n} u(X_i) \in C \Big\}. \tag{2.13}$$

This formulation is relevant in the MA case below.

PROOF OF THEOREM 1. Take C as specified in the theorem and let $B = \mathbb{R}^{d_1} \times C$. Then (Λ, B) satisfy (H1) and we can apply Lemma 2.3. The conditioning in (2.5) becomes $\left\{\frac{Y_n}{n} \in B\right\} = \left\{\frac{Y_{n_2}}{n} \in C\right\}$, as required in (2.10). We will identify v_B in (2.5) as

$$v_B = \nabla\Lambda(0, \alpha_C). \tag{2.14}$$

Then

$$\left(\frac{Y_n}{n} - v_B\right)_1 = \frac{Y_{n_1}}{n} - (\nabla\Lambda(0, \alpha_C))_1$$

as required.

Thus it remains to prove (2.14). To this end we have

$$\begin{aligned}
\Lambda^*(v_B) &= \Lambda^*(B) \ \text{ by the property of the dominating point,} \\
&= \Lambda^*(\mathbb{R}^{d_1} \times C) = \inf_{c \in C} \Lambda^*(\mathbb{R}^{d_1} \times \{c\}) \\
&= \inf_{c \in C} \inf_{\{v : \pi(v) = c\}} \Lambda^*(v), \quad \text{where } \pi(v_1, v_2) = v_2, \qquad (2.15) \\
&= \inf_{c \in C} \Lambda^{(2)^*}(c) \ \text{ by the contraction principle,} \\
&= \Lambda^{(2)^*}(C) = \Lambda^{(2)^*}(v_C).
\end{aligned}$$

Now recall by (2.8) that $\Lambda(0,\alpha_2)=\Lambda^2(\alpha_2)$ for $\alpha_2\in\mathbb{R}^{d_2}$, and that by (2.9) there exists $\alpha_C\in\mathbb{R}^{d_2}$ such that $\nabla\Lambda^2(\alpha_C)=v_C$. Hence $\nabla\Lambda(0,\alpha_C)$ exists, with

$$(\nabla\Lambda(0,\alpha_C))_2=\nabla\Lambda^2(\alpha_C)=v_C. \tag{2.16}$$

Also clearly if for some $v\in\mathbb{R}^d$, $\alpha_v\in\mathbb{R}^d$, $\nabla\Lambda(\alpha_v)$ exists and $=v$, then

$$\Lambda^*(v)=\langle\alpha_v,v\rangle-\Lambda(\alpha_v). \tag{2.17}$$

Now substituting $v=\nabla\Lambda(0,\alpha_C)$ in (2.17), and applying (2.8), (2.15), (2.16), we get

$$\begin{aligned}\Lambda^*(\nabla\Lambda(0,\alpha_C))&=\langle(0,\alpha_C),\nabla\Lambda(0,\alpha_C)\rangle-\Lambda(0,\alpha_C)\\&=\langle\alpha_C,\nabla\Lambda^{(2)}(\alpha_C)\rangle-\Lambda^{(2)}(\alpha_C)\\&=\Lambda^{(2)^*}(\nabla\Lambda^{(2)}(\alpha_C))=\Lambda^{(2)^*}(v_C)=\Lambda^*(v_B).\end{aligned}$$

By the uniqueness of the dominating point

$$\nabla\Lambda(0,\alpha_C)=v_B.$$

□

3. Application to Markov additive chains

To define the MA chain let $\{X_n;n=0,1,\dots\}$ be a Markov chain (MC) taking values in a measurable space $(S,\mathcal{S})$, irreducible with respect to some measure φ on $(S,\mathcal{S})$, and let $\{\xi_n;n=0,1,\dots\}\subset\mathbb{R}^d$ be an adjoined sequence such that $\{(X_n,\xi_n)\}$ is itself a MC on $(S\times\mathbb{R}^d,\mathcal{S}\times\mathcal{R}^d)$ ($\mathcal{R}^d=$ Borel sets on $\mathbb{R}^d$), with transition function

$$P\{(X_{n+1},\xi_{n+1})\in A\times\Gamma\mid X_n=x,\xi_n=s\}=p(x,A\times\Gamma) \tag{3.1}$$

for $A\times\Gamma\in\mathcal{S}\times\mathcal{R}^d$, $x\in S$. Note that the right side is assumed to be independent of s. Let $S_n=\sum_{i=0}^{n}\xi_i$. Then $\{(X_n,S_n);n=0,1,\dots\}$ is called an MA chain. Examples are

$$S_n=\sum_{i=1}^{n}f(X_i),$$

or

$$\mathrm{S}_n=\sum_{i=1}^{n}h(X_{i-1},X_i),$$

for some $f:S\to\mathbb{R}^d$, $h:S\times S\to\mathbb{R}^d$.

Let

$$\widehat{P}(\alpha) = \{\widehat{p}(x, A; \alpha)\} = \left\{ \int e^{\langle \alpha, s \rangle} p(x, A \times ds), \alpha \in \mathbb{R}^d \right\}. \tag{3.2}$$

Thus $\widehat{P}(\alpha)$ is a (non-stochastic) irreducible kernel on $(S, \mathcal{S})$. Let $R(\alpha)$ be its convergence parameter, which always exists under the φ-irreducibility condition (see [12]). Under suitable further hypothesis, $\lambda(\alpha) = R^{-1}(\alpha)$ is an eigenvalue of $\widehat{P}(\alpha)$, with left eigenmeasure $\{l(A; \alpha) : A \in \mathcal{S}\}$ and right eigenfunction $\{r(x; \alpha) : x \in S\}$. An (unnecessarily strong) sufficient condition for this is the existence of a measure ν on $(S \times \mathbb{R}^d, \mathcal{S} \times \mathcal{R}^d)$ such that for some $0 < a \leqq b < \infty$, $m_0 > 0$

$$a\nu(A \times \Gamma) \leqq p^{m_0}(x, A \times \Gamma) \leqq b\nu(A \times \Gamma). \tag{H2}$$

(See Lemma 3.1 of Iscoe, Ney, Nummelin [7], hereafter referred to as [INN].)

There are weaker conditions for the existence of λ, l and r (see e.g. Ney and Nummelin [11]), but we will assume (H2) for definiteness and since it simplifies some arguments. Also note that (H2) implies irreducibility of $\{X_n\}$ with $\varphi(A) = \nu(A \times \mathbb{R}^d)$.

Let

$$H = \text{ the convex hull of the support of } \nu(S \times \cdot\,). \tag{3.3}$$

We will assume that

$$\mathcal{D}(\lambda) \textit{ is open, and } H^0 \neq \emptyset, \tag{H3}$$

where $\mathcal{D}(\lambda) = \{\alpha : R(\alpha) > 0\}$ by definition.

Now one can check that

$$\widehat{p}^n(x, A; \alpha) = E_x[e^{\langle \alpha, S_n \rangle}; X_n \in A].$$

The following lemma summarizes some relevant parts of Lemma 3.4 and Theorem 5.1 (and its proof) from INN. From this lemma we can then conclude that the measures

$$\mathbb{P}_x\{S_n \in \cdot\,, X_n \in A\}, \quad A \in \mathcal{S} \quad \text{with} \quad \varphi(A) > 0,$$

satisfy the conditions of hypothesis (H1).

LEMMA 3.1. *Assume* (H2) *and* (H3). *Then*

(i) $\Lambda = \log \lambda$ *(as defined above) is analytic and strictly convex on* $\mathcal{D}(\Lambda)$.
(ii) Λ *is essentially smooth on* $\mathcal{D}(\Lambda)$.
(iii) $\nabla \Lambda(\alpha) = v$ *has a solution* $\alpha_v \subset \mathcal{D}(\Lambda)$ *for all* $v \in H^0$.
(iv) $\lim \frac{1}{n} \log E_x[e^{\langle \alpha, S_n \rangle}; X_n \in A] = \Lambda(\alpha)$ *for* $x \in S$, $A \in \mathcal{S}$ *with* $l(A, 0) > 0$, *and* $\alpha \in \mathcal{D}(\Lambda)$.
(v) $\mathcal{D}^0(\Lambda^*) = H^0$.

REMARK. Under (H2), $l(A, 0) > 0$ is equivalent to $\varphi(A) > 0$. From now on, to avoid trivialities, we assume $\varphi(A) > 0$ for the sets A referred to in Lemmas 3.2, 3.3 and Theorem 2.

Thus the hypothesis (H1) is satisfied and we get from Lemmas 2.1 and 2.2:

LEMMA 3.2. *Assume* (H2) *and* (H3), *and let* B *be open and convex, with* $(B \cap H)^0 \neq \emptyset$. *Then*

(i) $\inf[\Lambda^*(v); v \in B]$ *is achieved at a unique point* $v_B \in \overline{B} \cap H^0$ *(the dominating point),*

(ii) $\nabla\Lambda(\alpha) = v_B$ *has solution* $\alpha_{v_B} \in \mathbb{R}^d$,

(iii) $\lim \frac{1}{n} \log \mathbb{P}\Big\{\frac{S_n}{n} \in B, X_n \in A\Big\} = -\Lambda^*(B) = -\Lambda^*(v_B)$.

(Note that $\mathcal{D}^0(\Lambda^*) = H^0$ (see e.g. the proof of Theorem 5.2 of INN).)

Replacing the measure $\mathbb{P}\Big\{Y_n \in \cdot \;\Big|\; \frac{Y_n}{n} \in B\Big\}$ in Lemma 2.3 by

$$\mathbb{P}_x\Big\{S_n \in \cdot \;\Big|\; \frac{S_n}{n} \in B, X_n \in A\Big\}$$

we also obtain

LEMMA 3.3. *Assume* (H2) *and* (H3) *and let* B *be open and convex with* $B \cap H^0 \neq \emptyset$. *Then*

$$\mathbb{P}_x\Big\{\Big\|\frac{S_n}{n} - v_B\Big\| > \epsilon \;\Big|\; \frac{S_n}{n} \in B, X_n \in A\Big\} \xrightarrow{\text{exp}} 0.$$

The proof is exactly the same as Lemma 2.3, except that (2.8) is replaced by

$$\begin{aligned}\limsup \frac{1}{n} \log \mathbb{P}_x\Big\{\frac{S_n}{n} \in B^{(\epsilon)} \;\Big|\; \frac{S_n}{n} \in B, X_n \in A\Big\}\\ = \limsup \frac{1}{n}\Big[\log \mathbb{P}_x\Big\{\frac{S_n}{n} \in B^{(\epsilon)}, X_n \in A\Big\} - \log \mathbb{P}_x\Big\{\frac{S_n}{n} \in B, X_n \in A\Big\},\end{aligned}$$

with Lemma (3.2)(iii) then applied to the above expression.

To state an analog of Theorem 1, we again use the convention that for any $v \in \mathbb{R}^d$, we write $v = (v_1, v_2)$ with $v_1 \in \mathbb{R}^{d_1}$, $v_2 \in \mathbb{R}^{d_2}$, $d_1 + d_2 = d$, so v_1 is the first d_1 coordinates of v.

Now consider the MA chain $\Big(X_n, S_n = \sum_1^n \xi_i\Big)$, and write $\xi_n = (\xi_{n1}, \xi_{n2}) \in \mathbb{R}^d$, with $\xi_{n1} \in \mathbb{R}^{d_1}$, $\xi_{n2} \in \mathbb{R}^{d_2}$, and write $S_{ni} = \sum_{k=1}^n \xi_{ki}$, $i = 1, 2$. Then $\{X_n, S_{n1}\}$ is an MA chain with transition function

$$p^{(1)}(x, A \times \Gamma_1) = p(x, A \times (\Gamma_1 \times \mathbb{R}^{d_2})) \text{ for } \Gamma_1 \in \mathcal{R}^{d_1}, \tag{3.4}$$

and transform kernel

$$\widehat{P}^{(1)}(\alpha_1) = \{\widehat{p}^{(1)}(x, A; \alpha_1)\} = \{\widehat{p}(x, A; (\alpha_1, 0))\},\ \alpha_1 \in \mathbb{R}^{d_1},\ (\alpha_1, 0) \in \mathbb{R}^d. \tag{3.5}$$

Similarly $\{X_n, S_{n2}\}$ is an MA chain with

$$p^{(2)}(x, A\times\Gamma_2) = p(x, A\times\mathbb{R}^{d_1}\times\Gamma_2),\ \Gamma_2\in\mathcal{R}^{d_2}, \tag{3.6}$$

and

$$\widehat{p}^{(2)}(x, A;\alpha_2) = \widehat{p}(x, A;(0,\alpha_2)),\ \alpha_2\in\mathbb{R}^{d_2},\ (0,\alpha_2)\in\mathbb{R}^d.$$

We want to prove that

$$\frac{S_{n1}}{n}\xrightarrow{\exp} v \tag{3.7}$$

with respect to the conditioned measures

$$\mathbb{P}_x\Big\{\cdot\ \Big|\ \frac{S_{n2}}{n}\in C, X_n\in A\Big\}, \qquad C\subset\mathbb{R}^{d_2}, \tag{3.8}$$

and to identify the dependence of v on C.

To this end, note first that under (H2), $p^{(i)}$, $i=1,2$ will satisfy

$$a\nu^i(A\times\Gamma_i)\le p^{(i)^{m_0}}(x, A\times\Gamma_i)\le b\nu^i(A\times\Gamma_i),\ \Gamma_i\subset\mathbb{R}^{d_i}, \tag{3.9}$$

where

$$\begin{aligned}\nu^1(A\times\Gamma_1) &= \nu(A\times(\Gamma_1\times\mathbb{R}^{d_2})),\\ \nu^2(A\times\Gamma_2) &= \nu(A\times(\mathbb{R}^d\times\Gamma_2)).\end{aligned} \tag{3.10}$$

Then there will exist $\lambda^i(\alpha_i)$, $l^i(\cdot,\alpha_i)$, $r^i(\cdot,\alpha_i)$, $\alpha_i\in\mathbb{R}^{d_i}$, $i=1,2$. Let $\Lambda^{(i)} = \log\lambda^{(i)}$, and define $H^{(i)}$ analogously to (3.3).

Note that

$$\Lambda(\alpha_1, 0) = \Lambda^{(1)}(\alpha_1), \qquad \Lambda(0,\alpha_2) = \Lambda^{(2)}(\alpha_2), \tag{3.11}$$

and

$$(\nabla\Lambda(\alpha_1,0))_1 = \nabla\Lambda^{(1)}(\alpha_1), \qquad (\nabla\Lambda(0,\alpha_2))_2 = \nabla\Lambda^{(2)}(\alpha_2). \tag{3.12}$$

One can check that if $C\subset\mathbb{R}^{d_2}$ is open and convex with $C\cap H^{(2)^0}\ne 0$, then $(\Lambda^{(2)}, H^{(2)}, S_{n2})$ satisfy the hypotheses and conclusions of Lemmas 3.1 and 3.2. Hence there again exists a unique dominating point v_C and solution $\alpha_C\in\mathbb{R}^{d_2}$ of $\nabla\Lambda(\alpha_2) = v_C$. Taking $B=\mathbb{R}^{d_1}\times C$, and arguing as in Theorem 1, we conclude that

THEOREM 2. *Let $\{(X_n, S_n); n=1,2,\dots\}$ be an MA chain satisfying* (H2) *and* (H3), *and let C, v_C, α_C be as given above. Then for every $\epsilon>0$*

$$\mathbb{P}_x\Big\{\Big\|\frac{S_{n1}}{n} - v_1\Big\| > \epsilon\ \Big|\ \frac{S_{n2}}{n}\in C, X_n\in A\Big\}\xrightarrow{\exp} 0 \text{ as } n\to\infty,$$

where

$$v_1 = (\nabla\Lambda(0, \alpha_C))_1.$$

SPECIAL CASE. Take $\xi_n = f(X_n)$, $f\colon S \to \mathbb{R}^d$, and take $f(s) = (g(s), u(s))$, $s \in S$, $g \in \mathbb{R}^{d_1}$, $u \in \mathbb{R}^{d_2}$. Then

$$\mathbb{P}_x\Big\{\Big\|\frac{1}{n}\sum_{i=1}^{n} g(X_i) - v_1\Big\| > \epsilon \,\Big|\, \frac{1}{n}\sum_{i=1}^{n} u(X_i) \in C, X_n \in A\Big\} \xrightarrow{\exp} 0 \text{ as } n \to \infty.$$

REFERENCES

[1] CSISZÁR, I., Sanov property, generalized I-projection and a conditional limit theorem, *Ann. Probab.* **12** (1984), 768–793. *MR* **86h**:60065

[2] CSISZÁR, I., COVER, T. M. and CHOI, B. S., Conditional limit theorems under Markov conditioning, *IEEE Trans. Information Theory* **33** (1987), 788–801. *MR* **89c**: 60074

[3] DEMBO, A. and KUELBS, J., A Gibbs conditioning principle for certain infinite dimensional statistics, *University of Wisconsin Technical Report*, 1997.

[4] DEMBO, A. and ZEITOUNI, O., *Large deviation techniques and applications*, Jones and Bartlett Publishers, Boston, MA, 1993. *MR* **95a**:60034

[5] EINMAHL, U. and KUELBS, J., Dominating points and large deviations for random vectors, *Probab. Theory Related Fields* **105** (1996), 529–543. *MR* **97k**:60008

[6] ELLIS, R., *Entropy, large deviations, and statistical mechanics*, Grundlehren der math. Wissenschaften, **271**, Springer-Verlag, New York – Berlin, 1985. *MR* **87d**: 82008

[7] ISCOE, I., NEY, P. and NUMMELIN, E., Large deviations of uniformly recurrent Markov additive processes, *Adv. in Appl. Math.* **6** (1985), 373–412. *MR* **88b**:60077

[8] LEHTONEN, T. and NUMMELIN, E., On the convergence of empirical distributions under partial observations, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **13** (1988), 219–223. *MR* **90e**:60035

[9] LEHTONEN, T. and NUMMELIN, E., Level I theory of large deviations in the ideal gas, *Internat. J. Theoret. Phys.* **29** (1990), 621–635. *MR* **91f**:82033

[10] NEY, P., Dominating points and the asymptotics of large deviations for random walk on $\mathbb{R}^d$, *Ann. Probab.* **11** (1983), 158–167. *MR* **85b**:60026

[11] NEY, P. and NUMMELIN, E., Markov additive processes. I. Eigenvalue properties and limit theorems, *Ann. Probab.* **15** (1987), 561–592. *MR* **88h**:60056

[12] NUMMELIN, E., *General irreducible Markov chains and nonnegative operators*, Cambridge Tracts in Mathematics, **83**, Cambridge University Press, Cambridge – New York, 1984. *MR* **87a**:60074

[13] NUMMELIN, E., A conditional weak law of large numbers, *Stability problems for stochastic models* (Sukhumi, 1987), Lecture Notes in Math., 1412, Springer, Berlin, 1989, 259–262. *MR* **91f**:60057

[14] SCHROEDER, C., I-projection and conditional limit theorems for discrete parameter Markov processes, *Ann. Probab.* **21** (1993), 721–758. *MR* **94g**:60054

[15] 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

[16] Vasicek, O. A., A conditional law of large numbers, *Ann. Probab.* **8** (1980), 142–147. *MR* **81f**:60044

[17] Rockafellar, R. T., *Convex analysis*, Princeton Math. Series, No. 28, Princeton University Press, Princeton, NJ, 1970. *MR* **43** #445

(Received May 16, 1997)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN–MADISON
VAN VLECK HALL
480 LINCOLN DRIVE
MADISON, WI 53706
U.S.A.

recep@math.wisc.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 317–332

COVERAGE PROBABILITIES OF RÉNYI CONFIDENCE BANDS

Z. MEGYESI

Dedicated to the memory of Alfréd Rényi

Abstract

The applicability of Rényi confidence bands, as extended by S. Csörgő [6], are investigated by an extensive computer simulation study. Some new bands are also proposed.

1. Introduction

Theorems 5 and 6 of Rényi's paper [8] provide the possibility of drawing confidence contours and bands to the continuous distribution function $F(\cdot)$ or to the survival function $1-F(\cdot)$ the width of which is proportional to the natural unbiased estimator of the function to be estimated. Such bands are called *Rényi confidence bands.* M. Csörgő pointed out in [2] that $F(\cdot)$ in the denominator of these theorems of Rényi can be replaced by the sample distribution function $F_n(\cdot)$ if the supremum of the relative error of $F(\cdot)$ is taken over the set $\{x : p \leqq F_n(x)\}$ rather than $\{x : p \leqq F(x)\}$, for any fixed $p \in (0,1)$. E. Csáki showed in [1] that Rényi's Theorem 5 remains true if the fixed p is changed to a sequence p_n such that $p_n \in (0,1)$ and $p_n \to 0$ provided $np_n \to \infty$. M. Csörgő, S. Csörgő, L. Horváth and D. M. Mason [3] proved that *both* theorems of Rényi remain true if $\{p_n\}_{n=1}^{\infty}$ is a sequence such that $0 < p_n \leqq p$ for some $p \in (0,1)$ and $np_n \to \infty$; thus even $p_n \to 0$ is not necessary.

S. Csörgő proved in [6], Theorem 1, that $F(\cdot)$ can be replaced by $F_n(\cdot)$ and $\{x : p_n \leqq F(x)\}$, the set over which the supremum is taken, by $\{x : p_n \leqq F_n(x)\}$ under much more general conditions than in the paragraph above. Also, his Theorem 2 shows the existence of certain narrowed and combined versions. Later on we use the notations and definitions of [6], so our paper should be regarded as a continuation of [6]. We use some terminology consistently throughout: we write *confidence band* if it can be used for the estimation of the unknown $F(\cdot)$ or $1-F(\cdot)$, (i.e., it is determined by the sample itself) and at every point of its support it gives *both* lower and upper bounds for $F(\cdot)$ or $1-F(\cdot)$, respectively, and *confidence contour* if the statistics can be used for estimation, but give either only lower or upper bound on the function to be estimated. When a band can be drawn only if

1991 *Mathematics Subject Classification.* Primary 62G05.

Key words and phrases. Asymptotic confidence bands, distribution and survival functions, coverage probabilities, simulation.

a simple goodness-of-fit null hypothesis specifies $F(\cdot)$, it will be referred to as a *test band*. This is the case for the test contour and band (1) and (2) below, which arise from Rényi's original theorems in the case of $p_n \equiv p$. We write *one-sided* band if the band is motivated by the problem of estimation of either the left or right tail of the distribution, and *two-sided* band if it can be used to estimate *both* tails. Almost all of our bands are one-sided, but, for example, (8) and (9) are two-sided. We formulate all of our results in detail in the form motivated by the problem of estimating the right tail, or more precisely, the survival function $1-F(\cdot)$; the left-tail versions are briefly touched upon at the end of Section 6.

The results mentioned above are all limit theorems, with the exception of Csáki's formulae in [1]; the latter give the exact actual coverage probabilities of test contours. It would be very useful for any application to know for what sample sizes n of practical statistical use and — using the notations of [6] — for what choice of k_n or p_n are the actual and nominal coverage probabilities tolerably close and when do the actual coverage probabilities reach the nominal one. The aim of this paper is to describe the results of an extensive computer simulation investigating these questions.

Let $\alpha \in (0,1)$ be a fixed number and let y_α, z_α and w_α denote the unique values for which $K(y_\alpha) = L(z_\alpha) = 2\Phi(w_\alpha) - 1 = 1 - \alpha$, where $K(\cdot)$ is the Kolmogorov distribution function, $L(\cdot)$ is the distribution function of the absolute supremum of a standard Wiener process on the interval $[0,1]$ and $\Phi(\cdot)$ stands for the normal distribution function. The critical values y_α and w_α can be obtained from many textbooks, for z_α we refer the reader to [4]. However, for the sake of greater precision, these values were also computed by our programs directly by using the formulae for the corresponding functions. Let us define, for any $p_n \in (0,1)$,

$$\gamma^+_{p_n}(\alpha) := 1 + w_\alpha \sqrt{\frac{1-p_n}{np_n}}, \quad c^+_{p_n} := 1 + z_\alpha \sqrt{\frac{1-p_n}{np_n}}, \quad c^-_{p_n} := 1 - z_\alpha \sqrt{\frac{1-p_n}{np_n}}.$$

If $p_n \equiv p$ then Rényi's [8] Theorem 5 implies that

$$\mathbf{P}\left\{ \frac{1-F_n(x)}{\gamma^+_{p_n}(\alpha)} \leqq 1 - F(x),\ F(x) \leqq 1 - p_n \right\} \to 1 - \alpha, \tag{1}$$

and his Theorem 6 implies that

$$\mathbf{P}\left\{ \frac{1-F_n(x)}{c^+_{p_n}(\alpha)} \leqq 1 - F(x) \leqq \frac{1-F_n(x)}{c^-_{p_n}(\alpha)},\ F(x) \leqq 1 - p_n \right\} \to 1 - \alpha. \tag{2}$$

From (3) and (4) of [6] we know that (1) and (2) hold for any sequence $p_n \in (0,p)$ for any fixed $p \in (0,1)$, as long as $np_n \to \infty$. Theorem 1 of [6] also implies that

$$\mathbf{P}\left\{ \frac{1-F_n(x)}{\gamma^+_{k_n}(\alpha)} \leqq 1 - F(x),\ x \leqq X_{n-k_n,n} \right\} \to 1 - \alpha, \tag{3}$$

where

$$\gamma^+_{n,k_n} = 1 + w_\alpha \sqrt{\frac{1-\frac{k_n}{n}}{k_n}},$$

for a sequence of integers $\{k_n\}_{n=1}^\infty$ such that $1 \leqq k_n \leqq np$, $n \geqq 1/p$, for some $p \in (0,1)$. Note that (1) is a test contour, while (3) can be used for estimation as well, and if we choose $p_n = k_n/n$ then (1) and (3) are the same on the common part of their supports. We call them *analogous* in this sense. We use this term for bands, too. The formulae (11) and (12) of [6] imply that

$$\mathbf{P}\left\{ \frac{1-F_n(x)}{c^+_{n,k_n}(\alpha)} \leqq 1 - F(x) \leqq \frac{1-F_n(x)}{c^-_{n,k_n}(\alpha)},\ \ x \leqq X_{n-k_n,n} \right\} \to 1-\alpha, \tag{4}$$

$$\mathbf{P}\left\{ c^-_{n,k_n}(\alpha)[1-F_n(x)] \leqq 1 - F(x) \leqq c^+_{n,k_n}(\alpha)[1-F_n(x)],\ \ x \leqq X_{n-k_n,n} \right\} \to 1-\alpha, \tag{5}$$

by the above conditions on $\{k_n\}_{n=1}^\infty$. Hence (4) and the original Rényi confidence band (2) are analogous. One can expect that

$$\mathbf{P}\left\{ c^-_{p_n}(\alpha)[1-F_n(x)] \leqq 1 - F(x) \leqq c^+_{p_n}(\alpha)[1-F_n(x)],\ \ F(x) \leqq 1 - p_n \right\} \to 1-\alpha \tag{6}$$

holds under the above conditions on $\{p_n\}_{n=1}^\infty$; here (6) is the test version of (5). We shall prove in the next section that this is indeed the case.

From the first part of Theorem 2 of [6] we have that

$$\mathbf{P}\left\{ \frac{1-F_n(x)}{c^+_{n,k_n}(\alpha)} \leqq 1 - F(x) \leqq c^+_{n,k_n}(\alpha)[1-F_n(x)],\ \ x \leqq X_{n-k_n,n} \right\} \to 1-\alpha. \tag{7}$$

That theorem also shows that if $k_n/n \to 0$ then from (7) and from the corresponding left-tail version we can construct a *two-sided* confidence band

$$\mathbf{P}\left\{ L^{(\alpha)}_{n,k_n}(x) \leqq F(x) \leqq U^{(\alpha)}_{n,k_n}(x),\ \ X_{k_n,n} \leqq x \leqq X_{n-k_n,n} \right\} \to 1-\alpha. \tag{8}$$

For the upper and lower contours $U^{(\alpha)}_{n,k_n}$ and $L^{(\alpha)}_{n,k_n}$ we refer the reader to [6].

The first and most important fact we have to mention is that (1)–(8) are *distribution-free*: for any meaningful p_n, k_n, n, α the actual coverage probability does not depend on $F(\cdot)$, as the first step of the proofs in [6] shows. This provides the possibility of studying them by investigating only the particular case of the Uniform(0,1) distribution, with all our findings being universally applicable for any continuous distribution function $F(\cdot)$. Thus we may and do assume that the underlying distribution is the Uniform(0,1), i.e., $F(x) = x$ identically on [0,1]. Obviously it is enough to watch only the order statistics to see whether $F(\cdot)$ lies in the band determined by $F_n(\cdot)$.

We concentrate on the investigation of the *bands*, we study the contours mostly to check our simulation. We used several thousand hours of running time on IBM Pentium personal computers.

2. Coverage probabilities of one-sided bands

Our investigations cover the cases when the sample size n is between 10 and 2000, and $\alpha \in \{0.1, 0.05, 0.01\}$; these are the most important and customarily typical cases for practical use. All our qualitative and quantitative statements are for these three situations. Rényi confidence bands are for tail estimation, so we may restrict ourselves to the values $p_n \in (0, \frac{1}{2})$ and $k_n \in \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$. (The upper and lower integer parts of x are denoted by $\lceil x \rceil$ and $\lfloor x \rfloor$, respectively.)

We generated 100000 samples of the appropriate sizes and we obtained the actual coverage probabilities as the means of these 100000 Bernoulli-trials. Thus the error of our simulated values are not more than

$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	
0.00156	0.00113	0.00052	with 90% probability
0.00186	0.00135	0.00062	with 95% probability
0.00244	0.00177	0.00081	with 99% probability

Table 1

From the formulae (1)–(8) one can see that the width of these bands and the distance of the contours from the unknown $F(\cdot)$ decrease as p_n or k_n grow, so the behaviour of the actual coverage probabilities is far from obvious.

Let us first investigate the narrowest band, (7). In this particular case, as k_n grows, the actual coverage probability also grows as expected. In spite of this, the actual coverage probabilities are quite far from the nominal ones even for k_n close to $\lfloor \frac{n}{2} \rfloor$. For example, if $n = 100$ then the actual coverage probabilities are not more than 0.8656, 0.92 and 0.9728 instead of 0.9, 0.95 and 0.99. The situation is, of course, better for larger samples, but even for $n = 1000$, the actual coverage probabilities reach only 0.8868, 0.9397 and 0.9856, respectively. That is, the actual coverage probabilities of (7) converge to the nominal one from below, and the rate of convergence is quite slow. For sample sizes between 10 and 2000 they are not close enough, so the band (7) is not suitable for practical use.

Let us now look at the bands (4) and (5). If the true distribution function $F(\cdot)$ lies in the band in (7) then it lies both in those in (4) and (5), thus we can expect a better behaviour for these two bands.

Indeed, the actual coverage probabilities grow in these cases as k_n grows, but in the case of (4) they get close to the nominal one and they reach it only for samples of very large sizes. In the case of (5) the actual coverage probabilities increase rapidly with k_n and even for relatively small values they reach the nominal coverage probability and remain above it. (We deal with the question what "relatively small" means in Section 6.)

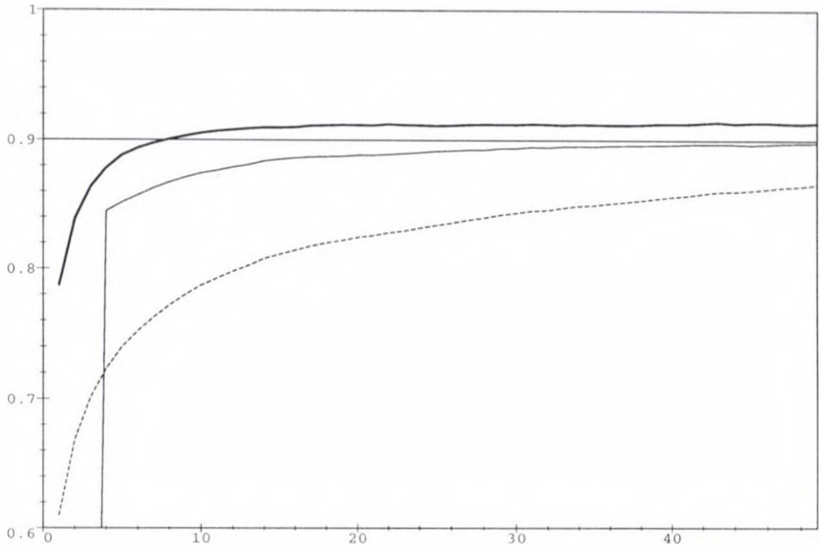

Figure 1. The actual coverage probabilities of (4), (5), and (7), shown by thin, medium wide and broken lines, respectively, in the case of $n=100$ and $\alpha=0.1$, as functions of k_n.

We have to mention about the band (4), and this is also true for the test band (2), that for very small k_n or p_n, $c^-_{n,k_n}(\alpha)$ or $c^-_{p_n}(\alpha)$ become negative, respectively, so the upper contours of these bands vanish. This occurs if

$$k_n \leqq \frac{z_\alpha^2}{1+\frac{z_\alpha^2}{n}} \quad \text{or} \quad p_n \leqq \frac{z_\alpha^2}{n+z_\alpha^2},$$

respectively. In these cases we considered the actual coverage probabilities to be 0.

The width of (4) is always larger than that of (5). If we denote the widths of these bands, which are proportional to $1-F_n(\cdot)$ in both cases, by $d^{(4)}_{n,k_n}(\alpha)[1-F_n(x)]$ and $d^{(5)}_{n,k_n}(\alpha)[1-F_n(x)]$, respectively, then

$$d^{(4)}_{n,k_n}(\alpha)=\frac{2z_\alpha\sqrt{k_n\left(1-\frac{k_n}{n}\right)}}{k_n-z_\alpha^2\left(1-\frac{k_n}{n}\right)} \quad \text{and} \quad d^{(5)}_{n,k_n}(\alpha)=2z_\alpha\sqrt{\frac{1-\frac{k_n}{n}}{k_n}},$$

yielding that,

$$d^{(4)}_{n,k_n}(\alpha)=d^{(5)}_{n,k_n}(\alpha)\frac{k_n}{k_n-z_\alpha^2\left(1-\frac{k_n}{n}\right)}.$$

The actual coverage probabilities of the Kolmogorov band were also always recorded in the course of our investigations. We did this partly with the aim of checking of our simulation. It is well-known about the Kolmogorov–Smirnov bands and contours from experience that for each finite n their actual coverage probabilities are above the corresponding nominal ones. Our simulation agreed with this for the case of the Kolmogorov band. The empirical observation was proved as a mathematical fact by Massart [7] for the Smirnov confidence contour.

Our simulation showed that the actual coverage probabilities of Rényi's original test band (2) exceed the nominal one for every meaningful choice of n and p_n, i.e., the (2) band is, like the Kolmogorov band, *conservative* in the above sense.

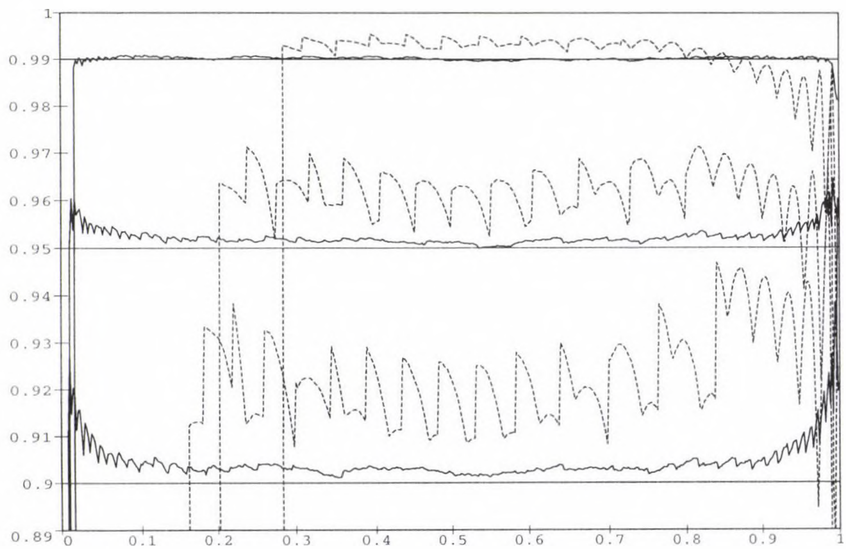

Figure 2. The actual coverage probabilities of the test band (2) (as function of p_n), in the case of $n=20$ (broken line) and $n=500$ (continuous line).

We prove (6) before investigating it. Let us denote the actual coverage probability of (6) by $\pi_n(\alpha)$. By using a simple idea of §3 of Rényi [9] we have

$$\pi_n(\alpha)=\mathbf{P}\Big\{c^-_{p_n}(\alpha)[1-F_n(x)]\leqq 1-F(x)\leqq c^+_{p_n}(\alpha)[1-F_n(x)],\ F(x)\leqq 1-p_n\Big\}$$

$$=\mathbf{P}\left\{\frac{1}{c^+_{p_n}(\alpha)}\leqq\frac{1-F_n(x)}{1-F(x)}\leqq\frac{1}{c^-_{p_n}(\alpha)}\ ,\ F(x)\leqq 1-p_n\right\}$$

$$=\mathbf{P}\left\{\frac{-z_\alpha\sqrt{\frac{1-p_n}{np_n}}}{c^+_{p_n}(\alpha)}\leqq\frac{F(x)-F_n(x)}{1-F(x)}\leqq\frac{z_\alpha\sqrt{\frac{1-p_n}{np_n}}}{c^-_{p_n}(\alpha)}\ ,F(x)\leqq 1-p_n\right\},$$

that is,

$$\mathbf{P}\left\{\sqrt{\frac{np_n}{1-p_n}}\sup_{F(x)\leqq 1-p_n}\frac{|F_n(x)-F(x)|}{1-F(x)}\leqq\frac{z_\alpha}{1+z_\alpha\sqrt{\frac{1-p_n}{np_n}}}\right\}\leqq\pi_n(\alpha)$$

$$\leqq\mathbf{P}\left\{\sqrt{\frac{np_n}{1-p_n}}\sup_{F(x)\leqq 1-p_n}\frac{|F_n(x)-F(x)|}{1-F(x)}\leqq\frac{z_\alpha}{1-z_\alpha\sqrt{\frac{1-p_n}{np_n}}}\right\},$$

where formula (4) in [6] shows that the upper and lower bounds both converge to $1-\alpha$ as long as $np_n\to\infty$, by the continuity of the limiting distribution function. Thus (6) is proved.

The behaviour of the actual coverage probabilities of (6) is very similar to that of (4): they grow as long as p_n grows, but they reach the nominal coverage probability only for samples of very large sizes. Thus (6) produces worse results than (2), so it cannot be recommended for practical use in simple goodness-of-fit tests.

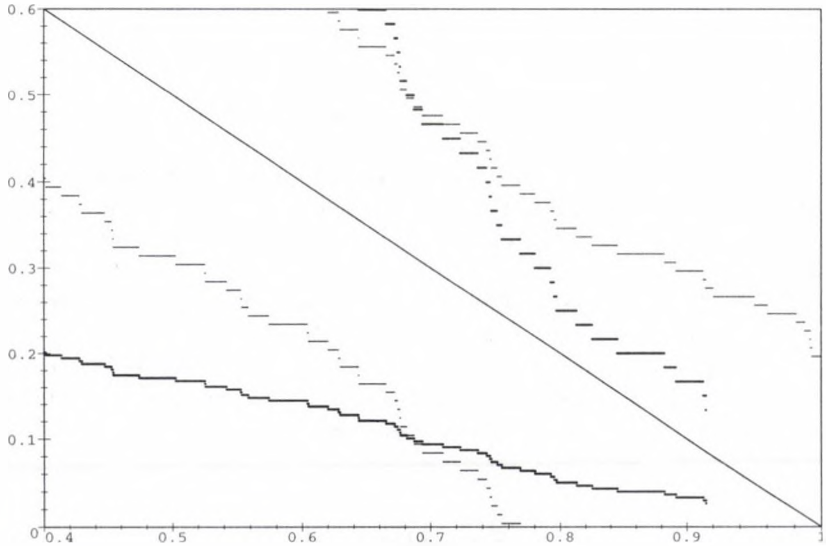

Figure 3. A typical view of (5) at $n=100$, $k_n=9$, $\alpha=0.1$, in comparison to the Kolmogorov band, which is depicted in thin lines. As Rényi confidence bands are for tail estimation, the figure concentrates on the right tail.

However, it is interesting to compare the behaviour of (2), (6), (4) and (5). If we choose $p_n=k_n/n$, the expectation of the support length of the band in (4) and (5) is almost the same as the support length of the band in (6) and (2). (This error results from the way we defined (4) and (5) and vanishes asymptotically.) The test band (2) is the test analogue of (4), while (6) is that of (5). By comparing (4) and (5) one can see that (5), the narrower band, is proved to be surprisingly far better, but the test band (6), which is the test analogue of (5) was beaten by (2). This shows that even for a simple goodness-of-fit test, for which all our bands can be used, there can be a drastic difference between analogous bands with the same nominal coverage probabilities.

This fact points to the significance of (5), which can be regarded as an almost always conservative all-purpose narrowed version of the original conservative Rényi confidence band (2).

3. Confidence contours

The investigation of Rényi's original test contour (1) was the most adequate way to check our simulation, because the exact values can be computed by Corollary 1.4.2 of Csáki's paper [1]. The simulated and computed probabilities were close indeed, their difference followed an approximative normal distribution with an expectation near to 0 and with variance according to the size of the Monte-Carlo. (This comparison was made in fact for the corresponding left-tail versions.)

Figure 4 shows the actual coverage probabilities of (1) at $n=100$, $\alpha=0.1$. The diagram is interesting in itself. One can see that the actual coverage probabilities grow rapidly with p_n and they get close to the nominal coverage probability very fast, then the actual ones oscillate for quite a long time near to the nominal one, later they reach it, and even while decreasing they remain above it.

The band (3), which is analogous to (1), produces far worse results. In this case the actual coverage probabilities of (3) grow monotonically with k_n, but they do not reach the nominal one. The relationship between (1) and (3) is similar to that between (2) and (4), as discussed at the end of the previous section.

4. Two-sided bands

In Section 1 above, (8) is the only example of a two-sided band. It is obtained by combining (7) and the corresponding left-tail version. We have seen that the actual and nominal coverage probabilities remain far apart in the case of (7) for sample sizes of practical use. This property of (7), which makes it unusable for any practical application is inherited by (8). So, let us try to draw a two-sided confidence band from the best one-sided band available, that is, (5).

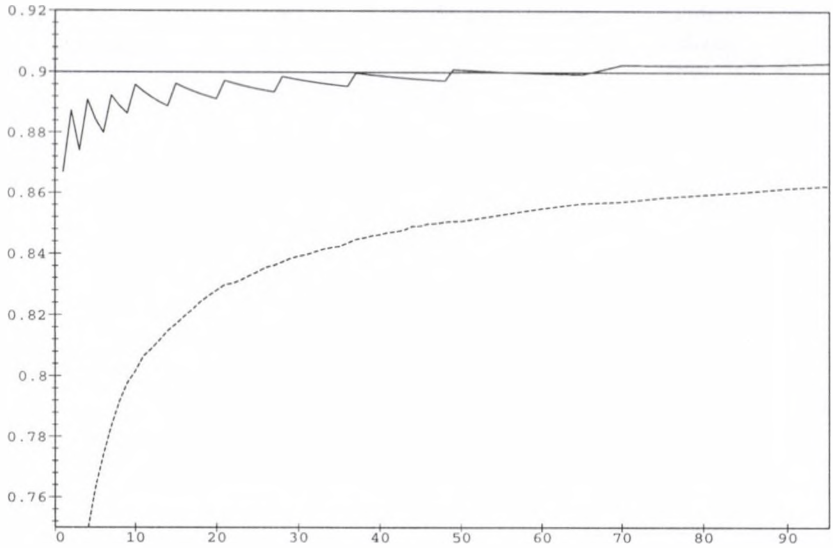

Figure 4. The actual coverage probabilities of (1) and (3), based on computed and simulated values, respectively in solid and broken lines, choosing $p_n = k_n/n$ for (1) in the case of $n = 200$, $\alpha = 0.1$.

Let z_α^* be the unique value for which $L(z_\alpha^*) = \sqrt{1-\alpha}$ and let

$$^*c_{n,k_n}^-(\alpha) := 1 - z_\alpha^* \frac{\sqrt{1-\frac{k_n}{n}}}{\sqrt{k_n}} \quad \text{and} \quad ^*c_{n,k_n}^+(\alpha) := 1 + z_\alpha^* \frac{\sqrt{1-\frac{k_n}{n}}}{\sqrt{k_n}}\,.$$

The formula (18) of [6] and the formula (5) of this paper imply that

$$(*) \qquad \mathbf{P}\Big\{ {}^*c_{n,k_n}^-(\alpha)F_n(x) \leqq F(x) \leqq {}^*c_{n,k_n}^+(\alpha)F_n(x)\,, \quad x \geqq X_{k_n,n} \Big\} \to \sqrt{1-\alpha}\,;$$

$$(**) \qquad \mathbf{P}\Big\{ 1 - {}^*c_{n,k_n}^+(\alpha)[1-F_n(x)] \leqq F(x) \leqq 1 - {}^*c_{n,k_n}^-(\alpha)[1-F_n(x)]\,, \quad x \leqq X_{n-k_n,n} \Big\} \to \sqrt{1-\alpha}\,.$$

If $F_n(x) \leqq 1/2$ then the upper contour of $(*)$ is beneath that of $(**)$ and the lower contour of $(*)$ is above that of $(**)$. This motivates the choices of the contours of a two-sided band shown below:

$$^*L_{n,k_n}^{(\alpha)}(x) := \begin{cases} {}^*c_{n,k_n}^-(\alpha)F_n(x), & X_{k_n,n} \leqq x < X_{\lceil n/2 \rceil,n}, \\ 1 - {}^*c_{n,k_n}^+(\alpha)[1-F_n(x)], & X_{\lceil n/2 \rceil,n} \leqq x \leqq X_{n-k_n,n}, \end{cases}$$

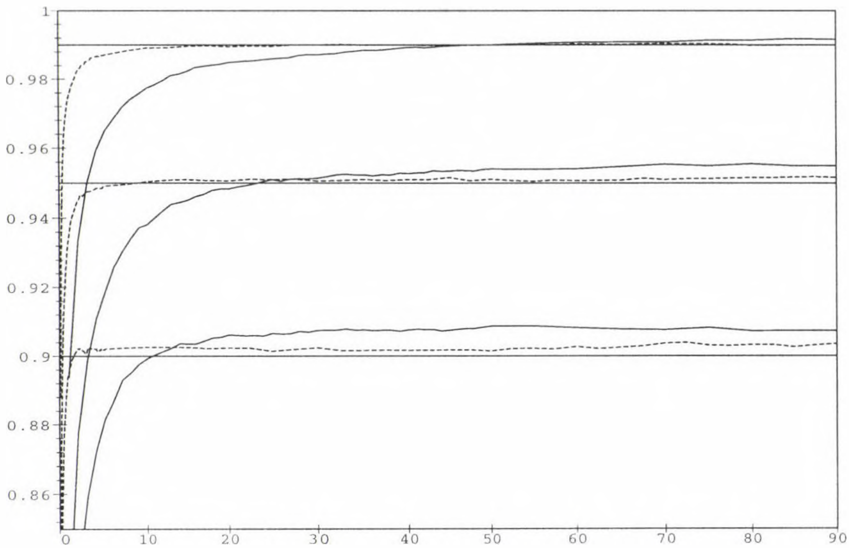

Figure 5. The coverage probabilities of (9) in the cases of $n=200$ (continuous line) and $n=2000$ (broken line). In the second case 1/10 of the real values of k_n are shown on the vertical axis.

$$^*U_{n,k_n}^{(\alpha)}(x):=\begin{cases} ^*c_{n,k_n}^{+}(\alpha)F_n(x), & X_{k_n,n}\leqq x<X_{\lceil n/2\rceil,n},\\ 1-{}^*c_{n,k_n}^{-}(\alpha)[1-F_n(x)], & X_{\lceil n/2\rceil,n}\leqq x\leqq X_{n-k_n,n}.\end{cases}$$

Then for any sequence of integers $\{k_n\}_{n=1}^{\infty}$ such that $1\leqq k_n\leqq np$ and $n\geqq 1/p$ for some $p\in(0,1)$, and $k_n\to\infty$ and $k_n/n\to 0$ we have that

$$\mathbf{P}\Big\{{}^*L_{n,k_n}^{(\alpha)}(x)\leqq F(x)\leqq {}^*U_{n,k_n}^{(\alpha)}(x),\ X_{k_n,n}\leqq x\leqq X_{n-k_n,n}\Big\}\to 1-\alpha. \tag{9}$$

This statement can be proved in the same way as (8), which is denoted by (21) in [6], by using the formulae (12) and (18) of [6]. The width of (9) in the intervals $x\in\big[X_{k_n,n},\,X_{\lceil n/2\rceil,n}\big)$ and $\big[X_{\lceil n/2\rceil,n},\,X_{n-k_n,n}\big]$ is proportional to $F_n(x)$ and $1-F_n(x)$, respectively. Let us denote this width by $d_{n,k_n}^{(9)}(\alpha)F_n(x)$ and $d_{n,k_n}^{(9)}(\alpha)[1-F_n(x)]$, respectively, where

$$d_{n,k_n}^{(9)}(\alpha)=2z_\alpha^*\sqrt{\frac{1-\frac{k_n}{n}}{k_n}}\,.$$

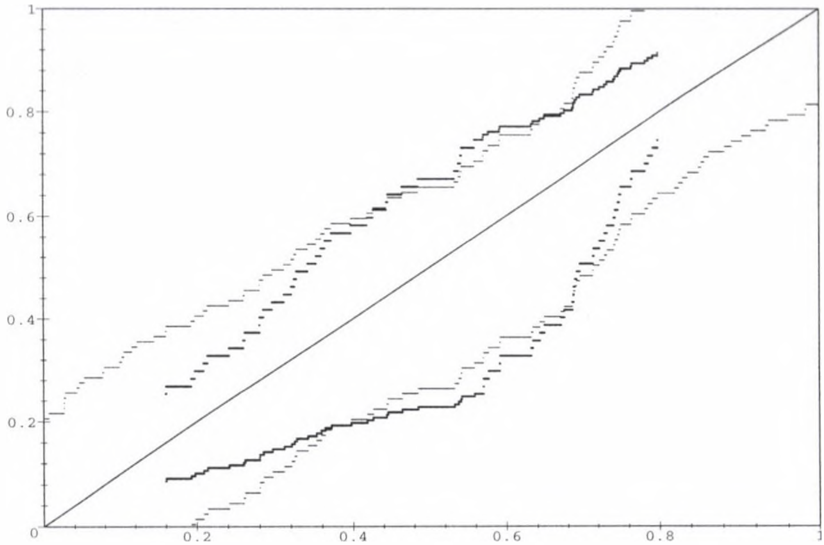

Figure 6. A typical view of (9), in the case of $n=100$, $k_n=17$, $\alpha=0.1$, in comparison to the Kolmogorov band (thin lines).

We expected (9) to inherit the good properties of (5). Indeed, the actual coverage probabilities grow together with k_n and they reach the nominal one quite fast and later remain above it. Figure 5 shows the coverage probabilities of (9) at samples of two different sizes.

5. On a narrowed Kolmogorov band

S. Csörgő and L. Horváth published the following result in [5], which narrows the Kolmogorov confidence band:

$$\begin{aligned}(10)\quad &\mathbf{P}\Bigg\{F_n(x)-\frac{y_\alpha}{n^{1/2}}+\frac{y_\alpha^2}{nF_n(x)+y_\alpha n^{1/2}}\leqq F(x)\\ &\leqq F_n(x)+\frac{y_\alpha}{n^{1/2}}-\frac{y_\alpha^2}{n[1-F_n(x)]+y_\alpha n^{1/2}},\ -\infty<x<\infty\Bigg\}\to 1-\alpha\,.\end{aligned}$$

Our investigations were extended to this band. The simulation showed that by narrowing the Kolmogorov band it loses its conservatism and the convergence of the actual coverage probabilities occurs from below, and they remain far from the nominal one. The actual coverage probabilities are shown in Figure 7.

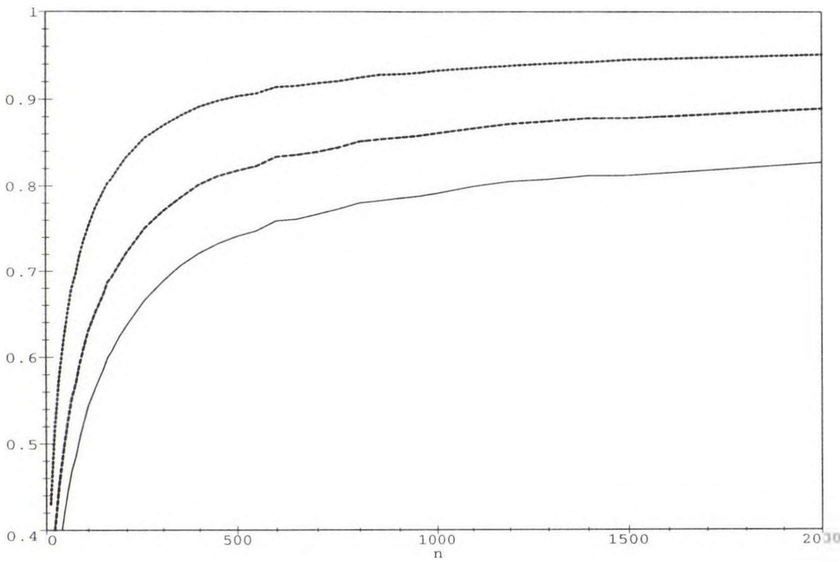

Figure 7. The coverage probabilities of (10), for $\alpha = 0.1$, $\alpha = 0.05$ and $\alpha = 0.01$, shown in continuous, broken, and dense broken lines, respectively.

The authors of [5] write:

> "We have conducted a small scale Monte Carlo simulation to check the applicability of the Theorem [which yields (10)] at $n = 50$. We generated 40 samples of size 50 from the uniform (0,1) distribution and constructed the bands [...] with $1 - \alpha = 0.9$. Maybe we were lucky, but $F(t) = t$ went out of the band only once."

This simulation was probably wrong, given that the actual coverage probability is only 0.443 in the above case, based on our results from 100000 samples. If their simulation is correct, the probability of the event described above is about $3.61 \cdot 10^{-13}$; i.e., less than 10000th part of the probability of winning the jackpot in the Hungarian National Lottery, where one has to pick 5 numbers out of 90.

The negative results obtained for (7) and (10) provide reason for some scepticism concerning the practical use of narrowing of conservative bands by mixing them; it seems easy to loose the conservatism and it may occur that the actual coverage probabilities remain far from the nominal one for sample sizes of practical use.

6. Rules of thumb

In this section we describe the conditions of applicability of some previously discussed bands of "good behaviour". For easier use we give some "rules of thumb".

We have noted about Rényi's original test band (2) that it is conservative for any $p_n \in \left(\frac{z_\alpha^2}{n+z_\alpha^2}, \frac{1}{2}\right)$ for sample sizes between 10 and 2000, i.e., (2) can be used for simple goodness-of-fit tests without any further consideration.

The actual coverage probabilities grow rapidly with k_n in case of (5) and (9), then they reach the nominal one and remain above it. We call the k_n for which the nominal coverage probability is achieved a *point of conservatism* and we denote it by $\kappa_n(\alpha)$. When we want to emphasize the band whose point of conservatism is being discussed, we write $\kappa_n^{(5)}(\alpha)$ or $\kappa_n^{(9)}(\alpha)$, respectively.

The first general observation resulting from the simulation is that for any fixed n, $\kappa_n(\alpha)$ grows as α decreases. In the case of $\alpha = 0.1$, the actual coverage probabilities grow quite rapidly for values of k_n close to $\kappa_n(0.1)$, so the exact values of $\kappa_n(0.1)$ can be determined quite accurately. For $\alpha = 0.05$ the growth of the actual coverage probabilities for k_n close to $\kappa_n(0.05)$ is much slower (cf. Figure 5), so the determined value of $\kappa_n(0.05)$ has a larger error resulting the random fluctuation of the simulation. This is even more true for $\kappa_n(0.01)$.

We prefer to give some simple inequalities instead of lengthy tables of points of conservatism. These rules are constructed for practical use, and they do not state anything concerning the asymptotic behaviour of $\kappa_n(\alpha)$. The "rules of thumb" for (5) are:

$$\kappa_n^{(5)}(0.1) \leqq 4\,n^{0.2}$$
$$\kappa_n^{(5)}(0.05) \leqq 1.5\,n^{0.55}$$
$$\kappa_n^{(5)}(0.01) \leqq 0.8\,n^{0.8}.$$

For example if $\alpha = 0.1$ and we have a sample of size 270, then the actual coverage probability of (5) is at least 0.9 if $n/2 \geqq k_n \geqq 4 \cdot 270^{0.2}$, that is, if $135 \geqq k_n \geqq 13$. These values are valid for $n \geqq 200$, but they are not sharp enough, so we have the following table of $\kappa_n^{(5)}(\alpha)$ for this small values of n.

	10	20	30	40	50	60	70	80	90
0.9	3	4	5	6	6	7	7	8	8
0.95	3	5	7	8	10 (9)	11	12	13	14
0.99	3	6	9	12	14	17 (16)	19	22 (21)	24 (23)

	100	120	140	160	180	200
0.9	9 (8)	9	10	10	11 (10)	11
0.95	15	17	19 (18)	21 (20)	23 (22)	24 (23)
0.99	26 (25)	31 (30)	34 (33)	40 (38)	44 (42)	46 (44)

Table 2

The above values are observed from 2500000 samples of the appropriate sizes. If there are two values in a cell, this means that the value of $\kappa_n^{(5)}(\alpha)$ could not be determined unambiguously. If k_n equals to the larger one, the actual coverage probability reaches the nominal one with a probability of at least 0.9, but even if k_n equals to the smaller one, which is in brackets, this will happen with probability of at least 0.1.

The inequalities for (9) are even easier to remember:

$$\kappa_n^{(9)}(0.1) \leqq 0.7\, n^{0.7}$$
$$\kappa_n^{(9)}(0.05) \leqq 0.6\, n^{0.8}$$
$$\kappa_n^{(9)}(0.01) \leqq 0.5\, n^{0.9}.$$

These are sharp for small n as well, so it is not necessary to give an additional table. One can see that (9) becomes conservative somewhat later than (5).

There is no point in giving such rules for the test contour (1), because the exact coverage probabilities can be computed by the formulae of Csáki's [1]. We mention that if one only requires an actual coverage probability of 0.885 instead 0.9, (1) satisfies this for $p_n > 7/n$ if $10 \leqq n \leqq 400$ and for $p_n > 10/n$ if $400 < n \leqq 2000$. In fact, one can give similar rules for any other value less than the nominal coverage probability (cf. Figure 4).

Finally, we spell out the equivalent left-tail versions of the bands discussed in this paper. For test bands it is obvious that for any choice of $0 < c_1 < 1 < c_2 < \infty$ and any $p_n \in (0,1)$ and n

$$\mathbf{P}\Big\{c_1 F_n(x) \leqq F(x) \leqq c_2 F_n(x),\ F(x) \geqq p_n\Big\} =$$

$$= \mathbf{P}\Big\{c_1[1 - F_n(x)] \leqq 1 - F(x) \leqq c_2[1 - F_n(x)],\ F(x) \leqq 1 - p_n\Big\}.$$

For confidence bands we have

$$\mathbf{P}\Big\{c_1 F_n(x) \leqq F(x) \leqq c_2 F_n(x),\ x > X_{k_n+1,n}\Big\} =$$

$$= \mathbf{P}\Big\{c_1[1 - F_n(x)] \leqq 1 - F(x) \leqq c_2[1 - F_n(x)]\,,\ \ x \leqq X_{n-k_n,n}\Big\}.$$

since $F_n(\cdot)$ is *right*-continuous.

Notice that in [6] the corresponding left and right-tail versions are not equivalent in this sense, however, for (8) and (9) we followed the style of [6]. We did so with the aim of the compatibility with [6]. Of course, one can write these two bands so as to be equivalent on the tails, then asymptotically everything remains the same, and the actual coverage probabilities get slightly better.

7. Conclusion

Summarizing our investigations, we can state that the original Rényi test band (2) can be used for any choice of $p_n \in \left(\frac{z_\alpha^2}{n+z_\alpha^2}, \frac{1}{2}\right)$ for simple goodness-of-fit tests against alternative hypotheses describing deviations on the tails. When one is interested in the behaviour of the unknown $F(\cdot)$ on *one* tail, then (5) can be used for this purpose, or, more importantly for estimation purposes if $k_n \geqq \kappa_n^{(5)}(\alpha)$. When one wants to estimate the unknown distribution function on *both* tails, (9) is suitable for this purpose if $k_n \geqq \kappa_n^{(9)}(\alpha)$. All of our statements are for $n \in \{10, \dots, 2000\}$ and $\alpha \in \{0,1; 0,05; 0,01\}$.

It remains for me to discharge the pleasant duty of expressing my thanks to Professor Sándor Csörgő who drew my attention to the topic discussed in this paper and helped with many valuable suggestions to improve the paper.

REFERENCES

[1] Csáki, E., Vizsgálatok az empirikus eloszlásfüggvényről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **23** (1977), 239–327. *MR* **57** #4415. [English translation: Investigations concerning the empirical distribution function, *Selected Translations in Mathematical Statistics and Probability* **15**, American Mathematical Society, Providence, RI, 1981, 229–317.]

[2] Csörgő, M., Some Rényi type limit theorems for empirical distribution functions, *Ann. Math. Statist.* **36** (1965), 322–326. *MR* **31** # 4130

[3] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M., Weighted empirical and quantile processes, *Ann. Probab.* **14** (1986), 31–85. *MR* **87e**:60041

[4] Csörgő, S. and Horváth, L., The baboons come down from the trees quite normally, *Mathematical statistics and applications*, Vol. B (Bad Tatzmannsdorf, 1983), Proceedings of the Fourth Pannonian Symposium on Mathematical Statistics, (W. Grossman, G. Ch. Pflug, I. Vincze and W. Wertz, eds.), D. Reidel, Dordrecht, 1985, 95–106. *MR* **87d**:62002

[5] Csörgő, S. and Horváth, L., Confidence bands from censored samples, *Canad. J. Statist.* **14** (1986), 131–144. *MR* **87i**:62089

[6] Csörgő, S., Rényi confidence bands, *Studia Sci. Math. Hungar.* **34** (1998), 71–87.

[7] MASSART, P., The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality, *Ann. Probab.* **18** (1990), 1269–1283. *MR* **91i**:60052

[8] RÉNYI, A., On the theory of order statistics, *Acta Math. Acad. Sci. Hungar.* **4** (1953), 191–231. *MR* **15**, 885e. [Hungarian version: A rendezett minták elméletéről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **3** (1953), 467–503. *MR* **15**, 544d.]

[9] RÉNYI, A., A rendezett minták elméletének egy problémaköréről, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **18** (1968), 23–30. *MR* **39** #3630. [English version: On some problems in the theory of order statistics, *Bull. Inst. Internat. Statist.* **42** (1969), 165–176. *MR* **41** #2859.]

(Received June 11, 1997)

SZEGEDI EGYETEM
BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
H–6720 SZEGED
HUNGARY

megyesiz@server.math.u-szeged.hu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 333–344

ASYMPTOTICS OF PERIODIC PERMANENTS

T. F. MÓRI

To the memory of Professor Alfréd Rényi

Abstract

A limit theorem is proved, as $n \to \infty$, for the permanent of the $mn \times mn$ matrix tiled with n^2 copies of a fixed positive $m \times m$ matrix.

1. Introduction

Let $\mathbf{A} = (a_{ij})$ be a fixed $m \times m$ matrix with positive entries. We want to study the asymptotic behaviour of the permanent of the $(nm) \times (nm)$ matrix consisting of $n \times n$ blocks, each identical with $\mathbf{A}$. Formally, let

$$\mathrm{P}_n(\mathbf{A}) = \mathrm{Per} \left.\begin{bmatrix} \mathbf{A} & \mathbf{A} & \dots & \mathbf{A} \\ \mathbf{A} & \mathbf{A} & \dots & \mathbf{A} \\ \dots & \dots & \dots & \dots \\ \mathbf{A} & \mathbf{A} & \dots & \mathbf{A} \end{bmatrix}\right\} n \text{ blocks.}$$

Why study that? One possible motivation comes from the following observation. When the elements of $\mathbf{A}$ are all equal: $a_{ij} = a$, we have $\mathrm{Per}\,(\mathbf{A}) = a^m m!$, that is, $a = (\mathrm{Per}\,(\mathbf{A})/(m)!)^{1/m}$. This suggests that $\Phi(\mathbf{A}) =: (\mathrm{Per}\,(\mathbf{A})/(m)!)^{1/m}$ could be used as a "mean value" of the elements of a square matrix $\mathbf{A}$. $\Phi(\mathbf{A})$ is easily seen to lie between the geometrical and the arithmetical means of the elements, and, in addition, it mirrors somehow the matrix structure, too. Unfortunately,

$$\Phi(\mathbf{A}) = \Phi\left(\begin{bmatrix} \mathbf{A} & \mathbf{A} \\ \mathbf{A} & \mathbf{A} \end{bmatrix}\right)$$

fails in general, though such an identity might be required. Therefore, one might want to modify the definition of $\Phi(\mathbf{A})$ to $\lim\limits_{n\to\infty} (\mathrm{P}_n(\mathbf{A})/(mn)!)^{1/mn}$, provided the latter exists. The existence of the limit follows from a general theorem of Girko [4]; still it is of interest to determine its concrete value for these periodic matrices. Girko's method was based on the fact that a permanent can always be represented as the second moment of a random

1991 *Mathematics Subject Classification.* Primary 15A15.
Key words and phrases. Permanent, local CLT, saddle point method.

0081–6906/98/$ 5.00

determinant. This time we use another, elementary method to obtain more precise asymptotic results for $\mathrm{P}_n(\mathbf{A})$.

Permuting the rows and columns we can see that

$$\mathrm{P}_n(\mathbf{A}) = \mathrm{Per}\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1m} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2m} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_{m1} & \mathbf{A}_{m2} & \dots & \mathbf{A}_{mm} \end{bmatrix}, \tag{1.1}$$

where all elements of the $n \times n$ matrix $\mathbf{A}_{ij}$ are equal to a_{ij}.

The general term in the expansion of $\mathrm{P}_n(\mathbf{A})$ is of the form

$$\prod_{i=1}^{m}\prod_{j=1}^{m} a_{ij}^{n_{ij}},$$

where n_{ij} denotes the number of elements taken from $\mathbf{A}_{ij}$. One can see that

$$\sum_{k=1}^{m} n_{ik} = n, \quad \text{and} \quad \sum_{k=1}^{m} n_{kj} = n \quad \text{for every} \quad i, j, \ 1 \leqq i \leqq m, \ 1 \leqq j \leqq m.$$

How many terms of this form are there? Let us first choose the n_{ij} rows of the elements to be taken from $\mathbf{A}_{ij}$, then the corresponding columns, finally we assign a column to each row. In that way we obtain

$$\begin{aligned} \mathrm{P}_n(\mathbf{A}) &= \sum_{\mathbf{N}} \prod_{i=1}^{m}\left(\frac{n!}{\prod_j n_{ij}!}\right) \prod_{j=1}^{m}\left(\frac{n!}{\prod_i n_{ij}!}\right) \prod_{i=1}^{m}\prod_{j=1}^{m} n_{ij}! \prod_{i=1}^{m}\prod_{j=1}^{m} a_{ij}^{n_{ij}} \\ &= (n!)^{2m} \sum_{\mathbf{N}} \prod_{i=1}^{m}\prod_{j=1}^{m} \frac{a_{ij}^{n_{ij}}}{n_{ij}!}, \end{aligned} \tag{1.2}$$

where the summation runs over $m \times m$ matrices $\mathbf{N} = (n_{ij})$ with nonnegative integer entries, such that its row sums and column sums are all equal to n.

Such a sum can be estimated similarly to what is usually done in the proof of the Moivre–Laplace theorem [3, Ch. VII]. One first selects the maximal term of the sum, then nearby terms can be estimated through their ratio to the maximal term, yielding a sum asymptotically equivalent to a Gauss integral. And this is just what we are going to do in the next section.

This plan of work has a drawback: it relies on the positivity of $\mathbf{A}$. It happens several times — even in the most interesting applications — that permanents of 0–1 matrices are considered. For instance, the number of 1-factors (matchings) in a bipartite graph with 2-colorization $\{U, V\}$, $U = \{u_1, \dots, u_n\}$, $V = \{v_1, \dots, v_n\}$ is equal to the permanent of the $n \times n$ matrix $\mathbf{A} = (a_{ij})$, where a_{ij} denotes the number of (u_i, v_j)-edges [5, Problem 4.21]. For his result [4] Girko needed the same restriction. That is not surprising. Things may become messy with zero or negative entries. As an illustrative example, in Section 3, we give a complete discussion of the most simple case of 2×2 matrix $\mathbf{A}$. Our tools will be generating functions and saddle point method; they do not seem to help in the general case.

2. A limit theorem

In order to formulate our result we first need the following fact.

LEMMA. *There exists a unique* $m \times m$ *matrix* $\mathbf{T} = (t_{ij})$ *such that for every* $i = 1, 2, \ldots, m$, *and* $j = 1, 2, \ldots, m$

$$\sum_{k=1}^{m} t_{kj} = m, \qquad \sum_{k=1}^{m} t_{ik} = m \quad \textit{and} \quad t_{ij} = c_i d_j a_{ij}$$

for some positive constants c_i *and* d_j.

PROOF. Let $\hat{\mathbf{A}} = \mathbf{A} / \sum_i \sum_j a_{ij}$ and define the linear family

$$\mathcal{L} = \Big\{ \hat{\mathbf{T}} = (\hat{t}_{ij})_{m \times m} : \hat{t}_{ij} \geqq 0, \ \sum_k \hat{t}_{kj} = \sum_k \hat{t}_{ik} = \frac{1}{m}, \forall i, \forall j \Big\}.$$

The correspondig exponential family of distributions (see [2]) is

$$\mathcal{E} = \{ \mathbf{Q} = (q_{ij})_{m \times m} : q_{ij} = c_i d_j \hat{a}_{ij} \quad \text{for some} \quad c_i > 0, d_j > 0 \},$$

and its closure $\mathrm{cl}\,\mathcal{E}$ is of the same form with $c_i \geqq 0$, $d_j \geqq 0$.

Let us minimize the Kullback–Leibler divergence

$$D(\hat{\mathbf{T}} \| \hat{\mathbf{A}}) = \sum_i \sum_j \hat{t}_{ij} \log(\hat{t}_{ij} / \hat{a}_{ij}) \tag{2.1}$$

over $\mathcal{L}$. It is well known [2] that the minimum is attained for a unique probability distribution $\hat{\mathbf{T}}$ called the I-projection of $\hat{\mathbf{A}}$ on $\mathcal{L}$, and this is the only element of $\mathcal{L} \cap \mathrm{cl}\,\mathcal{E}$. In addition, since for $\hat{\mathbf{A}}$ there exists a distribution in $\mathcal{L}$ with the same support, the I-projection $\hat{\mathbf{T}}$ falls into $\mathcal{E}$ itself. Now, let $\mathbf{T} = m^2 \hat{\mathbf{T}}$. □

REMARK 1. For numerical computations it is useful to know that $\mathbf{T}$ can be obtained as the limit of a natural iteration procedure. Starting from $\mathbf{A}$, in each step we divide the rows by the row means ($= 1/m$ times the row sums), then the columns by the column means and so on, alternately (iterative proportional fitting procedure, IPFP). This is so because $\mathcal{L} = \mathcal{L}_1 \cap \mathcal{L}_2$, where

$$\mathcal{L}_1 = \Big\{ \hat{\mathbf{T}} = (\hat{t}_{ij})_{m \times m} : \hat{t}_{ij} \geqq 0, \ \sum_k \hat{t}_{ik} = \frac{1}{m}, \forall i \Big\},$$

$$\mathcal{L}_2 = \Big\{ \hat{\mathbf{T}} = (\hat{t}_{ij})_{m \times m} : \hat{t}_{ij} \geqq 0, \ \sum_k \hat{t}_{kj} = \frac{1}{m}, \forall j \Big\},$$

and IPFP, as described above, is just an alternating sequence of I-projections onto $\mathcal{L}_1$ and $\mathcal{L}_2$, resp. (apart from a constant factor), see [2].

Let $\mathbf{M}$ denote the following $m \times (m-1)$ matrix

$$\mathbf{M} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ \hline -1 & \dots & -1 \end{bmatrix}, \tag{2.2}$$

and $\mathbf{G} = \mathbf{M} \otimes \mathbf{M}$ (Kronecker product).

The main result of the present note is the following theorem.

THEOREM 1. *Let* $\mathbf{T}$ *be the matrix constructed to* $\mathbf{A}$ *in the Lemma above. Define* $\Delta = [\operatorname{diag}(\operatorname{vec}(\mathbf{T}))]^{-1}$. *Then*

$$\mathrm{P}_n(\mathbf{A}) \sim (nm)! \left(\frac{\det(\mathbf{G}'\mathbf{G}) \det(\Delta)}{\det(\mathbf{G}'\Delta\mathbf{G})} \right)^{1/2} \left(\frac{\operatorname{Per}(\mathbf{A})}{\operatorname{Per}(\mathbf{T})} \right)^n \tag{2.3}$$

as $n \to \infty$.

REMARK 2. Some of the components in (2.3) can be expressed in a more explicit way, e.g.

$$\frac{\operatorname{Per}(\mathbf{A})}{\operatorname{Per}(\mathbf{T})} = \left(\prod_{i=1}^m \prod_{j=1}^m \frac{a_{ij}}{t_{ij}} \right)^{1/m} = \left(\prod_{i=1}^m c_i \prod_{j=1}^m d_j \right)^{-1},$$

$$\det(\mathbf{G}'\mathbf{G}) = \det(\mathbf{M}' \otimes \mathbf{M}'\mathbf{M} \otimes \mathbf{M}) = \det(\mathbf{M}'\mathbf{M} \otimes \mathbf{M}'\mathbf{M}) = m^{2m-2},$$

$$\det(\Delta) = \left(\prod_{i=1}^m \prod_{j=1}^m t_{ij} \right)^{-1}.$$

However, formula (2.3) has the advantage that in the case where all entries of $\mathbf{A}$ are 1's it reduces immediately to $(nm)!$.

PROOF OF THEOREM 1. In (1.1), let us multiply the rows and columns through $\mathbf{A}_{ij}$ by c_i and d_j, resp. In that way we obtain that

$$\mathrm{P}_n(\mathbf{A}) = \left(\prod_{i=1}^m c_i \prod_{j=1}^m d_j \right)^{-1} \operatorname{Per} \begin{bmatrix} \mathbf{T} & \mathbf{T} & \dots & \mathbf{T} \\ \mathbf{T} & \mathbf{T} & \dots & \mathbf{T} \\ \dots & \dots & \dots & \dots \\ \mathbf{T} & \mathbf{T} & \dots & \mathbf{T} \end{bmatrix},$$

hence it suffices to deal with the case $\mathbf{A} = \mathbf{T}$.

Now consider (1.2). Define $\mathbf{N}^0 = (n_{ij}^0)_{m \times m}$ with nonnegative integer entries in such a way that its row sums and column sums be all equal to n, and $n_{ij}^0 = \frac{n}{m} t_{ij} + O(1)$ as n tends to infinity, $1 \leqq i \leqq m$, $1 \leqq j \leqq m$.

For a general term of the sum in the right-hand side of (1.2) let

$$S(\mathbf{N}) = \prod_{i=1}^m \prod_{j=1}^m \frac{t_{ij}^{n_{ij}}}{n_{ij}!}.$$

Introduce s_{ij} by $n_{ij} = n_{ij}^0 + s_{ij}\sqrt{n_{ij}^0}$. Clearly,

$$\sum_{k=1}^{m} s_{ik}\sqrt{n_{ik}^0} = \sum_{k=1}^{m} s_{kj}\sqrt{n_{kj}^0} = 0 \tag{2.4}$$

holds for every i and j. Suppose

$$|s_{ij}| \leqq K = K(n) = o(n^{1/6}). \tag{2.5}$$

Then, using Stirling's formula, we can write

$$\begin{aligned}
&\log \frac{S(\mathbf{N})}{S(\mathbf{N}^0)} \\
&= \log\left(\prod_{i=1}^{m}\prod_{j=1}^{m} t_{ij}^{n_{ij}-n_{ij}^0}\frac{n_{ij}^0!}{n_{ij}!}\right) = \sum_{i=1}^{m}\sum_{j=1}^{m}\sqrt{n_{ij}^0}s_{ij}\log t_{ij} \\
&\quad + \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\left(n_{ij}^0 + \frac{1}{2}\right)\log n_{ij}^0 - n_{ij}^0 - \left(n_{ij} + \frac{1}{2}\right)\log n_{ij} + n_{ij} + O(K/\sqrt{n})\right).
\end{aligned}$$

The double sum in the last line can be treated as follows.

$$\begin{aligned}
&\sum_{i=1}^{m}\sum_{j=1}^{m}\left((n_{ij}^0 - n_{ij})\log n_{ij}^0 - \left(n_{ij} + \frac{1}{2}\right)\log\frac{n_{ij}}{n_{ij}^0} + (n_{ij} - n_{ij}^0) + O(K/\sqrt{n})\right) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m}\Bigg(-\sqrt{n_{ij}^0}s_{ij}\left(\log\frac{n}{m} + \log t_{ij} + O\left(\frac{1}{n}\right)\right) \\
&\qquad - \left(n_{ij}^0 + \sqrt{n_{ij}^0}s_{ij} + \frac{1}{2}\right)\left(\frac{s_{ij}}{\sqrt{n_{ij}^0}} - \frac{s_{ij}^2}{2n_{ij}^0} + O(K/\sqrt{n})^3\right)\Bigg) + O(K/\sqrt{n}).
\end{aligned}$$

Hence,

$$\log\frac{S(\mathbf{N})}{S(\mathbf{N}^0)} = -\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} s_{ij}^2 + O(K^3/\sqrt{n}), \tag{2.6}$$

where the remainder is uniform.

Next, we shall estimate $S(\mathbf{N})$ for matrices $\mathbf{N}$ not satisfying (2.5). Clearly,

$$S(\mathbf{N}) = \frac{m^{2nm}}{(mn)!}\Pr(X_{ij} = n_{ij}, 1 \leqq i \leqq m, 1 \leqq j \leqq m),$$

where $\mathbf{X} = (X_{ij})_{m\times m}$ is a random matrix with polynomial distribution of order mn and parameters $\hat{\mathbf{T}} = (m^{-2}t_{ij})$. Therefore, applying Chernoff's bound on the tail of the binomial distribution [1, p. 236] we obtain

$$\begin{aligned}\sum_{\mathbf{N}:\ \exists(i,j),|s_{ij}|>K} S(\mathbf{N}) &\leqq \frac{m^{2nm}}{(mn)!}\Pr(\exists(i,j): |X_{ij} - n^0_{ij}| > K\sqrt{n}) \\ &\leqq \frac{m^{2nm}}{(mn)!}\sum_{i=1}^{m}\sum_{j=1}^{m}\Pr\Big(\Big|X_{ij} - \frac{n}{m}t_{ij}\Big| > K\sqrt{n}/2\Big) \\ &\leqq \frac{m^{2nm}}{(mn)!}2m^2\exp(-K^2/2m).\end{aligned}$$

On the other hand,

$$S(\mathbf{N}^0) = \frac{m^{2nm}}{(mn)!}\Pr(\mathbf{X} = \mathbf{N}^0) \geqq \text{const.}\frac{m^{2nm}}{(mn)!}n^{-m^2/2},$$

thus

$$\sum_{\mathbf{N}:\exists(i,j),|s_{ij}|>K} S(\mathbf{N}) = o(S(\mathbf{N}^0)), \tag{2.7}$$

whenever $K^2 - m^3\log n \to \infty$.

Comparing (2.6) and (2.7) we get

$$\begin{aligned}\mathrm{P}_n(\mathbf{A}) &\sim (n!)^{2m}\sum_{\mathbf{N}:\ \forall(i,j),|s_{ij}|\leqq K} S(\mathbf{N}) \\ &\sim (n!)^{2m}S(\mathbf{N}^0)\sum_{\mathbf{N}}\exp\Big(-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}s_{ij}^2\Big).\end{aligned} \tag{2.8}$$

In the rightmost sum $\mathbf{S} = (s_{ij})$ varies in the $(m-1)^2$-dimensional linear subspace $\mathcal{R}_n \subset \mathbb{R}^{m\times m}$ defined by (2.4). More precisely, since $\sqrt{n^0_{ij}}s_{ij}$ is integer, $\mathbf{S}$ runs over a lattice in $\mathcal{R}$. Let us compute the volume of an elementary cell. Such a parallelepipedon is spanned by the following vectors.

For $1 \leqq i \leqq m-1$, $1 \leqq j \leqq m-1$, let

$$\mathbf{E}_{ij} = \begin{array}{|ccccc} & & i & & m \\ 0 & & | & 0 & | \\ \hline & & (1) & \text{———} & (-1) \quad \leftarrow j \\ 0 & & | & 0 & | \\ \hline & & (-1) & \text{———} & (1) \quad \leftarrow m \end{array}$$

(that is, only four elements of $\mathbf{E}_{ij}$ differ from 0). A base of $\mathcal{R}_n$ can be obtained by dividing the elements of these $\mathbf{E}_{ij}$'s by the square roots of the corresponding elements of $\mathbf{N}^0$. In other words, by introducing $\mathbf{D}_n = [\operatorname{diag}(\operatorname{vec}(\mathbf{N}^0))]^{-1}$, we obtain a base of $\mathcal{R}_n$ in the form

$$\{\mathbf{D}_n^{1/2}\operatorname{vec}(\mathbf{E}_{ij})\colon 1 \leqq i \leqq m-1, 1 \leqq j \leqq m-1\}.$$

Let $\mathbf{G}$ denote the $m^2 \times (m-1)^2$ matrix with columns $\operatorname{vec}(\mathbf{E}_{ij})$ ordered lexicographically. It is easy to see that $\mathbf{G} = \mathbf{M} \otimes \mathbf{M}$ with $\mathbf{M}$ defined in (2.2). Thus, $\mathcal{R}_n$ is the column space of $\mathbf{D}^{1/2}\mathbf{G}$ and the volume of the parallelepipedon is equal to the product of non-zero characteristic values of that matrix, that is, to $(\det \mathbf{G}'\mathbf{D}_n\mathbf{G})^{1/2}$. Since $\mathbf{D}_n \sim \dfrac{m}{n}\Delta$, we have that

$$\det(\mathbf{G}'\mathbf{D}_n\mathbf{G}) \sim (m/n)^{(m-1)^2} \det(\mathbf{G}'\Delta\mathbf{G}).$$

Note that the sequence of subspaces $\mathcal{R}_n$ also converges to

$$\mathcal{R} = \Big\{(s_{ij}) \in \mathbb{R}^{m\times m}\colon \sum_{k=1}^{m} s_{ik}\sqrt{t_{ik}} = \sum_{k=1}^{m} s_{kj}\sqrt{t_{kj}} = 0\Big\}.$$

These and (2.8) together imply that

$$\mathrm{P}_n(\mathbf{A}) \sim S(\mathbf{N}^0)(n!)^{2m}(n/m)^{(m-1)^2/2}(\det \mathbf{G}'\Delta\mathbf{G})^{-1/2}\int_{\mathcal{R}} \exp\Big(-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} s_{ij}^2\Big)\,d\lambda,$$

where λ is the Lebesgue measure on (the Borel sets of) $\mathcal{R}$. This integral is well known to equal $(2\pi)^{(m-1)^2/2}$, for $\dim \mathcal{R} = \dim \mathcal{R}_n = (m-1)^2$.

The last step of the proof is the estimation of $S(\mathbf{N}^0)$. By the Stirling formula and some calculation one obtains

$$\begin{aligned}(n!)^{2m}\Big(\frac{2n\pi}{m}\Big)^{(m-1)^2/2} S(\mathbf{N}^0) &= (n!)^{2m}\Big(\frac{2n\pi}{m}\Big)^{(m-1)^2/2}\prod_{i=1}^{m}\prod_{j=1}^{m}\frac{t_{ij}^{n_{ij}^0}}{n_{ij}^0!}\\ &\sim (nm)!\,m^{m-1}\Big(\prod_{i=1}^{m}\prod_{j=1}^{m} t_{ij}\Big)^{-1/2}\prod_{i=1}^{m}\prod_{j=1}^{m}\Big(\frac{nt_{ij}}{mn_{ij}^0}\Big)^{n_{ij}^0}.\end{aligned}$$

Let us take the logarithm of the last double product.

$$\sum_{i=1}^{m}\sum_{j=1}^{m} n_{ij}^0 \log\Big(\frac{nt_{ij}}{mn_{ij}^0}\Big) = \sum_{i=1}^{m}\sum_{j=1}^{m} n_{ij}^0\Big(\frac{nt_{ij}}{mn_{ij}^0} - 1 + O(n^{-2})\Big) = O(n^{-1}),$$

that is, the product in question converges to 1. Thus, by Remark 2,

$$\mathrm{P}_n(\mathbf{A}) \sim (nm)!m^{m-1}\Big(\prod_{i=1}^{m}\prod_{j=1}^{m} t_{ij}\Big)^{-1/2}(\det \mathbf{G}'\Delta\mathbf{G})^{-1/2}$$
$$= (nm)!\Big(\frac{\det(\mathbf{G}'\mathbf{G})\det(\Delta)}{\det(\mathbf{G}'\Delta\mathbf{G})}\Big)^{1/2}$$

as claimed. □

REMARK 3. Consider the general term in the sum (1.2) and maximize it in $\mathbf{N}$. Let $\hat{\mathbf{T}} = \frac{1}{mn}\mathbf{N}$. Then by the Stirling formula we have

$$\log\Big(\prod_{i=1}^{m}\prod_{j=1}^{m}\frac{a_{ij}^{n_{ij}}}{n_{ij}!}\Big) = nm + nm\sum_{i=1}^{m}\sum_{j=1}^{m}\hat{t}_{ij}\Big(\log(\hat{a}_{ij}/\hat{t}_{ij})$$
$$+\log\Big(\sum_i\sum_j a_{ij}\Big) - \log(nm)\Big) + o(n)$$
$$= nm + nm\log\Big(\frac{\Sigma\Sigma a_{ij}}{nm}\Big) - nmD(\hat{\mathbf{T}}\|\hat{\mathbf{A}}) + o(n) \to \max!$$

Thus, $\hat{\mathbf{T}}$ is asymptotically equal to the I-projection of $\hat{\mathbf{A}}$ on $\mathcal{L}$. This explains how the reference term $S(\mathbf{N}^0)$ was selected in the proof.

REMARK 4. More precise calculation shows the rate of convergence: in (2.3) exact equality can be achieved with the right-hand side multiplied by a factor $\Big(1 + O\Big(\frac{1}{n}\Big)\Big)$.

3. The 2 × 2 case

Apparently, there exists no explicit formula for the matrix $\mathbf{T}$ that played central role in Theorem 1. In general we have a system of nonlinear equations for the multipliers c_i and d_j, namely,

$$(3.1)\qquad \sum_{k=1}^{m} a_{ik}d_k = m/c_i, \quad \text{and} \quad \sum_{k=1}^{m} a_{kj}c_k = m/d_j, \quad 1 \leqq i \leqq m,\ 1 \leqq j \leqq m.$$

In the simplest nontrivial particular case, i.e., where $m = 2$, equations (3.1) can be solved explicitly. Let

$$\mathbf{A} = \begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix},$$

then

$$\mathbf{T} = \begin{bmatrix} t & 2-t \\ 2-t & t \end{bmatrix} = \begin{bmatrix} \alpha c_1 d_1 & \beta c_1 d_2 \\ \gamma c_2 d_1 & \delta c_2 d_2 \end{bmatrix}, \quad \det \Delta = \frac{1}{t^2(2-t)^2} = \frac{1}{c_1^2 c_2^2 d_1^2 d_2^2 \alpha\beta\gamma\delta},$$

$$\mathbf{M} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{G}'\Delta\mathbf{G} = \frac{4}{t(2-t)}, \quad \frac{\mathrm{Per}\,(\mathbf{A})}{\mathrm{Per}\,(\mathbf{T})} = \frac{1}{c_1 c_2 d_1 d_2},$$

and

$$2 = t + (2-t) = (\sqrt{\alpha\delta} + \sqrt{\beta\gamma})\sqrt{c_1 c_2 d_1 d_2}.$$

Consequently,

$$\mathrm{P}_n(\mathbf{A}) \sim (2n)!(\alpha\beta\gamma\delta)^{-\frac{1}{4}} \left(\frac{\sqrt{\alpha\delta} + \sqrt{\beta\gamma}}{2}\right)^{2n+1}. \tag{3.2}$$

In Section 2 we only dealt with positive matrices $\mathbf{A}$. A naturally arising question is what can be said when positivity is not required. That seems hard in general, but in the 2×2 case it can be answered by applying the saddle-point method.

A key to the answer is the observation that

$$\mathrm{P}_n(\mathbf{A}) = (n!)^2 \sum_{k=0}^{n} \binom{n}{k}^2 (\alpha\delta)^k (\beta\gamma)^{n-k}$$

is just the coeffient of z^n in the polynomial

$$g(z) = (n!)^2 (z + \alpha\delta)^n (z + \beta\gamma)^n = (n!)^2 (z^2 + Pz + \alpha\beta\gamma\delta)^n, \tag{3.3}$$

where $P = \mathrm{Per}\,(\mathbf{A}) = \alpha\delta + \beta\gamma$. We shall also apply the notation $D = \det(A) = \alpha\delta - \beta\gamma$.

From (3.3) $\mathrm{P}_n(\mathbf{A})$ can be read off directly in some simple particular cases.

If $\alpha\beta\gamma\delta = 0$, then $\mathrm{P}_n(\mathbf{A}) = (n!)^2 \mathrm{P}^n \sim (2n)!\sqrt{n\pi}(P/4)^n$.

If $P = 0$, then $\mathrm{P}_n(\mathbf{A}) = 0$ for odd n, and for even n

$$\begin{aligned} \mathrm{P}_n(\mathbf{A}) &= (n!)^2 \binom{n}{n/2} (\alpha\beta\gamma\delta)^{n/2} \sim (2n)!\sqrt{2}(\alpha\beta)^n(-1)^{n/2} \\ &= (-1)^{n/2}(2n)!\sqrt{2}(D/4)^n. \end{aligned}$$

In the general case $\alpha\beta\gamma\delta \neq 0$ and $P \neq 0$. For the sake of simplicity, by the help of exchanging rows or columns, and multiplying them, if necessary, by -1, we can always achieve that α, β, δ, P and D are positive.

THEOREM 2. *Suppose $\alpha, \beta, \delta, P, D > 0$, $\gamma < 0$. Then*

$$\mathrm{P}_n(\mathbf{A}) = (2n)!(-\alpha\beta\gamma\delta)^{-\frac{1}{4}}\left(\frac{\alpha\delta+\beta\gamma}{4}\right)^{n+1/2}\left(\cos\left(\frac{2n+1}{2}\varphi\right)+\sin\left(\frac{2n+1}{2}\varphi\right)+o(1)\right),$$

where $\varphi = \arccos\dfrac{\alpha\delta+\beta\gamma}{\alpha\delta-\beta\gamma}$.

PROOF. In order to approximate $\mathrm{P}_n(\mathbf{A})$ we apply saddle-point method. By the Cauchy integral formula we have

$$\begin{aligned}\mathrm{P}_n(\mathbf{A}) &= (n!)^2\frac{1}{2\pi i}\oint\limits_{|z|=r} z^{-n-1}(z^2+Pz+\alpha\beta\gamma\delta)^n dz\\ &= (n!)^2\frac{1}{2\pi i}\oint\limits_{|z|=r}\left(z+P+\frac{\alpha\beta\gamma\delta}{z}\right)^n\frac{dz}{z}\\ &= (n!)^2\frac{1}{2\pi}\int_{-\pi}^{+\pi}\left(re^{it}+P+\frac{\alpha\beta\gamma\delta}{r}e^{-it}\right)^n dt.\end{aligned}$$

Introduce

$$\begin{aligned}R &=: \left|re^{it}+P+\frac{\alpha\beta\gamma\delta}{r}e^{-it}\right|^2\\ &= P^2+\left(r-\frac{\alpha\beta\gamma\delta}{r}\right)^2+2P\left(r+\frac{\alpha\beta\gamma\delta}{r}\right)\cos t+4\alpha\beta\gamma\delta\cos^2 t.\end{aligned}$$

In the case of $\gamma > 0$, r fixed, R becomes maximal when $\cos t = 1$, i.e., $t = 0$. Then $R = \left(r+P+\frac{\alpha\beta\gamma\delta}{r}\right)^2$, which is minimal for $r = \sqrt{\alpha\beta\gamma\delta}$. Thus the saddle-point is $z_0 = \sqrt{\alpha\beta\gamma\delta}$, and standard calculations, details omitted, lead to (3.2).

The case $\gamma < 0$ is more interesting. For sake of convenience we apply the notation $\vartheta = \sqrt{-\alpha\beta\gamma\delta}$. For $r > 0$ fixed let us take the maximum of R in $y = \cos t$ over the interval $[-1;1]$. The derivative is

$$R'(y) = \frac{\partial R}{\partial y} = 2P(r-\vartheta^2/r) - 8\vartheta^2 y.$$

If $R'(1) = 2P(r-\vartheta^2/r) - 8\vartheta^2 > 0$, then the maximum of R is at $y = 1$, and

$$\max R = R(1) = (P+r-\vartheta^2/r)^2 > (P+4\vartheta^2/P)^2 = D^4/P^2 > D^2.$$

If $R'(-1) = 2P(r-\vartheta^2/r) + 8\vartheta^2 < 0$, then the maximum of R is at $y = -1$, and

$$\max R = R(-1) = (P-r+\vartheta^2/r)^2 > (P+4\vartheta^2/P)^2 > D^2.$$

Finally, if $R'(1) \leqq 0 \leqq R'(-1)$, i.e., $P|r - \vartheta^2/r| \leqq 4\vartheta^2$, then the maximum of R is attained at the solution y_0 of $R'(y) = 0$. Thus,

$$y_0 = \frac{P(r + \vartheta^2/r)}{4\vartheta^2}, \quad \text{and} \quad R(y_0) = \frac{D^2(r - \vartheta^2/r)^2}{4\vartheta^2}.$$

This latter is minimal when $r = \vartheta$, then $R'(1) = -8\vartheta^2 < 0 < 8\vartheta^2 = R'(-1)$, $y_0 = 0$ and the minimum is $R(y_0) = D^2$. Hence the saddle-points are $z_0 = \pm i\vartheta$. Returning to the integral we can write

$$\begin{aligned}
P_n(\mathbf{A}) &= (n!)^2 \frac{1}{2\pi} \int_{-\pi}^{+\pi} (P + 2\vartheta i \sin t)^n dt \\
&= (n!)^2 \frac{1}{\pi} \operatorname{Re}\Bigl(\int_0^{\pi} (P + 2\vartheta i \sin t)^n dt\Bigr) \\
&= (n!)^2 \frac{1}{\pi} \operatorname{Re}\Bigl(\int_{-\pi}^{+\pi} (P + 2\vartheta i \cos t)^n dt\Bigr) \\
&= (n!)^2 \frac{1}{\pi} \operatorname{Re}\Bigl((P + 2\vartheta i)^n \int_{-\pi}^{+\pi} \Bigl(1 - \frac{2\vartheta i}{P + 2\vartheta i}(1 - \cos t)\Bigr)^n dt\Bigr).
\end{aligned} \tag{3.4}$$

Substitute $t = x/\sqrt{n}$. Then the last integral is asymptotically equal to

$$\sqrt{1/n} \int_{-\infty}^{+\infty} \exp\Bigl(-\frac{1}{1 - iP/2\vartheta} \frac{x^2}{2}\Bigr) dx = \sqrt{2\pi/n}\sqrt{1 - iP/2\vartheta},$$

where the last square root is chosen to have positive real part.

Since $P + 2\vartheta i = De^{i\varphi}$, from (3.4) we obtain

$$\begin{aligned}
\mathrm{P}_n(\mathbf{A}) &= (n!)^2 \frac{1}{\pi} \operatorname{Re}\Bigl((P + 2\vartheta i)^{n+1/2} (\pi/n\vartheta i)^{1/2} (1 + o(1))\Bigr) \\
&= (2n)! 2^{-2n} D^{n+1/2} \vartheta^{-1/2} \Bigl(\cos\Bigl(\Bigl(n + \frac{1}{2}\Bigr)\arg(P + 2\vartheta i) - \frac{\pi}{4}\Bigr) + o(1)\Bigr) \\
&= (2n)! (D/4)^{n+1/2} \vartheta^{-1/2} \Bigl(\cos\Bigl(\Bigl(n + \frac{1}{2}\Bigr)\varphi\Bigr) + \sin\Bigl(\Bigl(n + \frac{1}{2}\Bigr)\varphi\Bigr) + o(1)\Bigr).
\end{aligned}$$

The proof is completed. □

4. Back to the motivating problem

By Theorem 1 the modified mean of a positive matrix **A** could be defined as

$$\Phi(\mathbf{A}) =: \lim_{n\to\infty} \left(\mathrm{P}_n(\mathbf{A})/(mn)!\right)^{1/mn} = \left(\frac{\operatorname{Per}(\mathbf{A})}{\operatorname{Per}(\mathbf{T})}\right)^{1/m}$$
$$= \left(\prod_{i=1}^{m}\prod_{j=1}^{m}\frac{a_{ij}}{t_{ij}}\right)^{1/m^2} = \left(\prod_{i=1}^{m} c_i \prod_{j=1}^{m} d_j\right)^{-1/m}$$

(with the notations introduced in the Lemma of Section 2). Particularly, $\Phi(\mathbf{A}) = 1$ for all matrices with

$$\sum_{k=1}^{m} a_{kj} = m, \quad \sum_{k=1}^{m} a_{ik} = m, \quad 1 \leqq i \leqq m, \ 1 \leqq j \leqq m.$$

For 2×2 matrices $\mathbf{A} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ we obtain

$$\Phi(\mathbf{A}) = \frac{\sqrt{\alpha\delta} + \sqrt{\beta\gamma}}{2}.$$

Acknowledgement. This work was supported by the Hungarian National Foundation for Scientific Research Grants No. T-016384, T-016237. Author is indebted to G. J. Székely for proposing this problem.

REFERENCES

[1] Alon, N. and Spencer, J. H., *The probabilistic method*, With an appendix by Paul Erdős, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley, New York, 1992. *MR* **93h**:60002

[2] Csiszár, I., *I*-divergence geometry of probability distributions and minimization problems, *Ann. Probab.* **3** (1975), 146–158. *MR* **51** #2050

[3] Feller, W., *An introduction to probability theory and its applications*, Vol. I. 3rd ed., Wiley, New York, 1968. *MR* **37** #3604

[4] Girko, V. L., The method of random determinants for estimating the permanent, *Random Oper. Stochastic Equations* **3** (1995), 181–192. *MR* **96j**:65004

[5] Lovász, L., *Combinatorial problems and exercises*, North-Holland Publ. Co., Amsterdam – New York, 1979. *MR* **80m**:05001

(Received May 6, 1997)

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 34 (1998), 345–348

ALMOST SURE BEHAVIOUR OF SOME RANDOM SEQUENCES

V. V. PETROV

Dedicated to the memory of Alfréd Rényi

Abstract

This note examines the almost sure behaviour of sequences of random variables under conditions expressed in terms of characteristic functions. We obtain generalizations of some results of Chung and Erdős related to sums of independent identically distributed random variables.

1. Introduction

The set of functions $\psi(x)$ that are positive and non-decreasing in the region $x > x_0$ for some x_0 (depending on ψ) and such that the series $\sum 1/(n\psi(n))$ converges (diverges) will be denoted by Ψ_c (respectively, Ψ_d). For example, $x^p \in \Psi_c$ for every $p > 0$; $(\log x)^p \in \Psi_c$ if $p > 1$; $\log x \in \Psi_d$.

Chung and Erdős [1] proved that if $\{X_n\}$ is a sequence of independent random variables having a common distribution function with non-zero absolutely continuous component and if $\mathbf{E}X_1 = 0$, $\mathbf{E}|X_1|^5 < \infty$, then

$$\liminf_{n\to\infty} n^{1/2}\psi(n)|S_n| > 0 \quad \text{a.s.} \tag{1}$$

for every function $\psi \in \Psi_c$, but if $\psi \in \Psi_d$, then

$$\liminf_{n\to\infty} n^{1/2}\psi(n)|S_n| = 0 \quad \text{a.s.}. \tag{2}$$

Here $S_n = \sum_{j=1}^{n} X_j$.

In [3] and [4] (see also Theorem 6.20 in [5]) it was proved that if $\{X_n\}$ is a sequence of independent identically distributed random variables satisfying the Cramér condition

$$\limsup_{|t|\to\infty} \left|\mathbf{E}e^{itX_1}\right| < 1, \tag{C}$$

then

$$\lim_{n\to\infty} n^{1/2}\psi(n)|S_n| = \infty \quad \text{a.s.} \tag{3}$$

1991 *Mathematics Subject Classification.* Primary 60F15; Secondary 60F05.

Key words and phrases. Sequences of random variables, almost sure behaviour, characteristic functions.

0081–6906/98/$ 5.00 ©1998 *Akadémiai Kiadó, Budapest*

for every $\psi \in \Psi_c$; if the additional assumptions $\mathbf{E}X_1 = 0$ and $\mathbf{E}X_1^2 < \infty$ are satisfied, then (2) holds for every $\psi \in \Psi_d$. Note that there are no moment conditions in this proposition connected with (3).

In [6] relation (3) was proved for sums of arbitrary (not necessarily independent or identically distributed) random variables under some conditions expressed in terms of the characteristic functions of these sums; for a sequence of independent random variables sufficient conditions are given in terms of characteristic functions of summands. Unfortunately, the above mentioned generalization and strengthening of the Chung–Erdős theorem connected with (1) does not follow from results in [6]. The present note contains a more general result which is free of this disadvantage.

2. Results and proofs

Consider an arbitrary sequence of random variables $Y_1, Y_2, \dots$. We put $f_n(t) = \mathbf{E}e^{itY_n}$.

THEOREM 1. *Let $\psi \in \Psi_c$. Let $\{g(n)\}$ be a sequence of positive numbers. Suppose that the following condition is satisfied:*

$$\int\limits_{|t| \leqq \varepsilon n g(n) \psi(n)} |f_n(t)| dt = O(g(n)) \tag{4}$$

for some positive constant ε. Then

$$\lim_{n\to\infty} n g(n) \psi(n) |Y_n| = \infty \quad a.s.. \tag{5}$$

PROOF. For an arbitrary random variable X the Lévy concentration function $Q(X;\lambda)$ is defined by the equality

$$Q(X;\lambda) = \sup_x \mathbf{P}(x \leqq X \leqq x + \lambda).$$

By Esseen's inequality [2] (see also [5], Lemma 1.16) we have

$$Q(X;\lambda) \leqq A\lambda \int\limits_{|t| \leqq 1/\lambda} |f(t)| dt$$

for every $\lambda > 0$ where $f(t)$ is the characteristic function of X and A is an absolute positive constant. For our sequence of random variables $\{Y_n\}$ we obtain

$$Q(Y_n;\lambda) \leqq A\lambda \int\limits_{|t| \leqq 1/\lambda} |f_n(t)| dt$$

for every $\lambda > 0$.

Let $\psi \in \Psi_c$. Put $\lambda = L/(ng(n)\psi(n))$. Then

$$Q(Y_n;\lambda) \leqq \frac{AL}{ng(n)\psi(n)} \int_B |f_n(t)|dt, \tag{6}$$

where $B = \{t\colon |t| \leqq ng(n)\psi(n)/L\}$.

Let ε be a positive constant satisfying condition (4). If L is sufficiently large, $L > 1/\varepsilon$, then

$$\int_B |f_n(t)|dt = O(g(n)) \quad (n \to \infty).$$

It follows from (6) that

$$Q(Y_n;\lambda) \leqq \frac{C}{n\psi(n)}$$

for all sufficiently large n where C is a positive constant. We have

$$\mathbf{P}\Big(|Y_n| \leqq L/(2ng(n)\psi(n)) \leqq Q(Y_n; L/(ng(n)\psi(n))\Big) \\ \leqq C/(n\psi(n))$$

for all sufficiently large n. By the Borel-Cantelli lemma,

$$\mathbf{P}(|Y_n| \leqq L/(2ng(n)\psi(n)) \text{ i.o.}) = 0, \tag{7}$$

since $\psi \in \Psi_c$.

The only restriction on L is $L > 1/\varepsilon$ where ε is a positive constant satisfying condition (4). Therefore we conclude that (7) holds for an arbitrarily large L. The relation (5) follows. Theorem 1 is proved.

The following proposition is an immediate consequence of Theorem 1 in the case when $g(n) = n^{-1/2}$.

THEOREM 2. *If*

$$\int_{-\infty}^{\infty} |f_n(t)|dt = O(n^{-1/2})$$

then

$$\lim_{n\to\infty} n^{1/2}\psi(n)|Y_n| = \infty \quad a.s.$$

for every function $\psi \in \Psi_c$.

Theorem 2 is a generalization of a result in [6]. Theorem 2 does not imply the main result of [3] connected with relation (3). However, the latter result follows from Theorem 1 with $g(n)=n^{-1/2}$. It can be proved along the lines of the proof of Theorem 1 in [3].

REFERENCES

[1] CHUNG, K. L. and ERDŐS, P., On the lower limit of sums of independent random variables, *Ann. of Math.* (2) **48** (1947), 1003–1013. *MR* **9**, 292f

[2] ESSEEN, C. G., On the concentration function of a sum of independent random variables, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1968), 290–308. *MR* **37** #6974

[3] PETROV, V. V., On the lower limit of the modulus of sums of independent random variables, *Vestnik Leningrad. Univ. Mat. Mekh. Astronom.* **1978**, vyp. 2, 49–54 (in Russian). *MR* **80b**:60046

[4] PETROV, V. V., A remark on the lower limit of the absolute values of sums of independent random variables, *Litovsk. Mat. Sb.* **18** (1978), 115–119 (in Russian). *MR* **81a**:60045

[5] PETROV, V. V., *Limit theorems of probability theory*, Oxford Studies in Probability, 4, Oxford University Press, 1995. *MR* **96h**:60048

[6] PETROV, V. V., On the almost sure behaviour of sums of random variables, *Statist. Probab. Lett.* **24** (1995), 229–231. *MR* **96k**:60112

(Received May 9, 1997)

FACULTY OF MATHEMATICS AND MECHANICS
ST. PETERSBURG UNIVERSITY
BIBLIOTECHNAYA PL. 2
STARY PETERHOF
RU–198904 ST. PETERSBURG
RUSSIA

and

DEPARTMENT OF STATISTICS AND PROBABILITY
WELLS HALL
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824–1027
U.S.A.

petrov@stt.msu.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 349–378

BOUNDS ON PROBABILITIES AND EXPECTATIONS USING MULTIVARIATE MOMENTS OF DISCRETE DISTRIBUTIONS

A. PRÉKOPA

To the memory of Professor Alfréd Rényi

Abstract

The paper deals with the multivariate moment problems in case of discrete probability distribution. Assuming the knowledge of a finite number of multivariate moments, lower and upper bounds are provided for probabilities and expectations of functions of the random variables involved. These functions obey higher order convexity formulated in terms of multivariate divided differences. As special cases, the multivariate Bonferroni inequalities are derived. The bounds presented are given by formulas as well as linear programming algorithms. Numerical examples are presented.

1. Introduction

In this paper we present bounds on functionals of an unknown probability distribution under some moment information. Our functionals are expectations of higher order convex functions (see Popoviciu [17]) of random variables and probabilities of some events. Moments, at least some of them, are frequently easy to compute (even in experimental sciences, see, e.g., Wheeler and Gordon [24]) and the bounds that can be obtained on this ground are frequently very good, in the sense that the lower and upper bounds on some value are close to each other.

While the literature is rich in papers handling univariate moment problems of this kind, the multivariate case has not been studied enough until recently. The papers by Dulá [4], Kall [9], Kemperman and Skibinski [13] and Prékopa [21] can be mentioned as examples. Examples for more general moment problem formulations can be found, e.g., in the paper by Kemperman [12].

A few years ago the sharp Bonferroni inequalities of Dawson and Sankoff [3], Kwerel [14] and others, have been discovered as discrete moment problems by Samuels and Studden [23] and Prékopa [20]. In this case the random

1991 *Mathematics Subject Classification.* Primary 62H99, 90C05.

Key words and phrases. Multivariate moment problem, stochastic inequalities, linear programming, Lagrange interpolation.

Research supported by the Air Force–Grant numbers: AFORS-89-0512B, F49620-93-1-0041, and by the Hungarian Foundation for the Scientific Research (OTKA) Grant T 014102.

variables of which some of the moments are known are occurrences concerning event sequences and the moments are binomial rather than power moments.

Given the information that a random variable is discrete, where the support is also known, the application of the general moment problem (where the support is unrestricted) provides us with weaker bounds than the application of the discrete moment problem. In fact, in the latter case the set of feasible solutions is smaller than in the former case. Discrete random variables with known support are quite frequent in applications. Thus, research in discrete moment problems is important both from the point of view of theory and applications.

Research in connection with the multivariate discrete moment problem has been initiated by Prékopa [21]. This paper presents further and more important results in this respect.

Let $\xi_1, \dots, \xi_s$ be discrete random variables and assume that the support of ξ_j is a known finite set $Z_j = \{z_{j0}, \dots, z_{jn_j}\}$, where $z_{j0} < \cdots < z_{jn_j}$, $j = 1, \dots, s$. Then the support of the random vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_s)^T$ is part of the set $\mathbf{Z} = Z_1 \times \cdots \times Z_s$. We do not assume, however, the knowledge that which part of $\mathbf{Z}$ is the exact support of $\boldsymbol{\xi}$.

Let us introduce the notations

$$p_{i_1 \cdots i_s} = P(\xi_1 = z_{1i_1}, \dots, \xi_s = z_{si_s}) \quad 0 \leqq i_j \leqq n_j,\ j = 1, \dots, s \tag{1.1}$$

$$\mu_{\alpha_1 \cdots \alpha_s} = \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} p_{i_1 \cdots i_s}, \tag{1.2}$$

where $\alpha_1, \dots, \alpha_s$ are nonnegative integers. The number $\mu_{\alpha_1 \cdots \alpha_s}$ is called the $(\alpha_1, \dots, \alpha_s)$-order moment of the random vector $(\xi_1, \dots, \xi_s)$. The sum $\alpha_1 + \cdots + \alpha_s$ is called the total order of the moment.

We assume that the probabilities in (1.1) are unknown but known are some of the multivariate moments (1.2). We are looking for lower and upper bounds on the values

$$E\left[f(\xi_1, \dots, \xi_s)\right] \tag{1.3}$$

$$P\left(\xi_1 \geqq r_1, \dots, \xi_s \geqq r_s\right) \tag{1.4}$$

$$P\left(\xi_1 = r_1, \dots, \xi_s = r_s\right), \tag{1.5}$$

where f is some function defined on the discrete set $\mathbf{Z}$ and $r_j \in Z_j$, $j = 1, \dots, s$. The problems of bounding the probabilities (1.4) and (1.5) are special cases of the problem of bounding the expectation (1.3). In fact, if

$$f(z_1, \dots, z_s) = \begin{cases} 1, & \text{if } z_j \geqq r_j,\ j = 1, \dots, s \\ 0, & \text{otherwise,} \end{cases} \tag{1.6}$$

then (1.3) is equal to (1.4), and if

$$f(z_1, \ldots, z_s) = \begin{cases} 1, & \text{if } z_j = r_j,\ j = 1, \ldots, s \\ 0, & \text{otherwise,} \end{cases} \tag{1.7}$$

then (1.3) is equal to (1.5). In spite of this coincidence, the condition that we will impose on f, when bounding the expectation (1.3), does not always allow for the functions (1.6) and (1.7). Hence, separate attention has to be paid to the problems of bounding the probabilities.

As regards the moments (1.2), two different cases will be considered:

(a) there exist nonnegative integers $m_1, \ldots, m_s$ such that $\mu_{\alpha_1 \cdots \alpha_s}$ are known for $0 \leqq \alpha_j \leqq m_j$, $j = 1, \ldots, s$;

(b) there exists a positive integer m such that $\mu_{\alpha_1 \cdots \alpha_s}$ are known for $\alpha_1 + \cdots + \alpha_s \leqq m$, $\alpha_j \geqq 0$, $j = 1, \ldots, s$.

Case (b) is of course more practical than Case (a). If, e.g., we know all expectations, variances and covariances of the random variables $\xi_1, \ldots, \xi_s$, then Case (b) applies. If only the expectations and the covariances are known, then Case (a) applies. However, when the covariances are known then, in most cases, the variances are known, too.

We formulate the bounding problems as linear programming problems. For the sake of simplicity we will use the notation $f_{i_1 \cdots i_s} = f(z_{1i_1}, \ldots, z_{si_s})$. In both problems formulated below the decision variables are the $p_{i_1 \cdots i_s}$, all other entries are supposed to be known. In Case (a) the bounding problems are

$$\begin{aligned} &\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s} \\ &\text{subject to} \\ &\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} p_{i_1 \cdots i_s} = \mu_{\alpha_1 \cdots \alpha_s} \\ &\quad \text{for } 0 \leqq \alpha_j \leqq m_j,\ j = 1, \ldots, s \\ &p_{i_1 \cdots i_s} \geqq 0, \quad \text{all } i_1, \ldots, i_s. \end{aligned} \tag{1.8}$$

In Case (b) the bounding problems are

$$\begin{aligned} &\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s} \\ &\text{subject to} \\ &\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} p_{i_1 \cdots i_s} = \mu_{\alpha_1 \cdots \alpha_s} \\ &\quad \text{for } \alpha_j \geqq 0,\ j = 1, \ldots, s; \alpha_1 + \cdots + \alpha_s \leqq m \\ &p_{i_1 \cdots i_s} \geqq 0, \quad \text{all } i_1, \ldots, i_s. \end{aligned} \tag{1.9}$$

We reformulate these problems, using more concise notations. Let

$$A_j = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_{j0} & z_{j1} & \cdots & z_{jn_j} \\ \vdots & & \ddots & \\ z_{j0}^{m_j} & z_{j1}^{m_j} & \cdots & z_{jn_j}^{m_j} \end{pmatrix}, \qquad j = 1, \ldots, s$$

$$A = A_1 \otimes \cdots \otimes A_s,$$

where the symbol $\otimes$ refers to the tensor product. For example the tensor product of A_1 and A_2 equals

$$A_1 \otimes A_2 = \begin{pmatrix} A_1 & A_1 & \cdots & A_1 \\ z_{20} A_1 & z_{21} A_1 & \cdots & z_{2n_2} A_1 \\ \vdots & & \ddots & \\ z_{20}^{m_2} A_1 & z_{21}^{m_2} A_1 & \cdots & z_{2n_2}^{m_2} A_1 \end{pmatrix}.$$

Note that the tensor product is noncommutative but it has the associative property (see, e.g., Horn and Johnson [7]). We further introduce the notations:

$$\begin{aligned} \mathbf{b} &= E\left[(1, \xi_1, \ldots, \xi_1^{m_1}) \otimes \cdots \otimes (1, \xi_s, \ldots, \xi_s^{m_s})\right]^T \\ &= (\mu_{00\cdots0}, \mu_{10\cdots0}, \ldots, \mu_{m_10\cdots0}, \mu_{010\cdots0}, \mu_{11\cdots0}, \ldots)^T \\ \mathbf{p} &= (p_{i_1\cdots i_s},\ 0 \leqq i_1 \leqq m_1, \ldots, 0 \leqq i_s \leqq m_s)^T \\ \mathbf{f} &= (f_{i_1\cdots i_s},\ 0 \leqq i_1 \leqq m_1, \ldots, 0 \leqq i_s \leqq m_s)^T, \end{aligned}$$

where the ordering of the components in $\mathbf{p}$ and $\mathbf{f}$ coincides with that of the corresponding columns in the matrix $A = (a_{i_1\cdots i_s})$.

The optimum values of the linear programming problems

$$\begin{aligned} &\min(\max) \quad \mathbf{f}^T \mathbf{p} \\ &\text{subject to} \\ (1.10) \qquad &\qquad A\mathbf{p} = \mathbf{b} \\ &\qquad\ \ \mathbf{p} \geqq \mathbf{0} \end{aligned}$$

provide us with the best lower and upper bounds for $E[f(\xi_1, \ldots, \xi_s)]$ in Case (a). We call these bounding problems.

In Case (b) we define $\tilde{\mathbf{b}}$ as the vector obtained from $\mathbf{b}$ in such a way that we delete those moments $\mu_{\alpha_1\cdots\alpha_s}$ for which $\alpha_1 + \cdots + \alpha_s > m$. Deleting the corresponding rows from A, let $\tilde{A}$ designate the resulting matrix. Then, in Case (b), the bounding problems are:

$$\begin{aligned} &\min(\max) \quad \mathbf{f}^T \mathbf{p} \\ &\text{subject to} \\ (1.11) \qquad &\qquad \tilde{A}\mathbf{p} = \tilde{\mathbf{b}} \\ &\qquad\ \ \mathbf{p} \geqq \mathbf{0}. \end{aligned}$$

The matrix A has size $[(m_1+1)\cdots(m_s+1)]\times[(n_1+1)\cdots(n_s+1)]$ and is of full rank. The matrix $\tilde{A}$ has size $N\times[(n_1+1)\cdots(n_s+1)]$, where $N=\binom{s+m}{m}$ and is also of full rank.

It is well-known in linear programming theory that any dual feasible basis (i.e., that satisfies the optimality condition but is not necessarily primal feasible) has the property that the value of the objective function corresponding to the basic solution is smaller (greater) than or equal to the optimum value in case of a minimization (maximization) problem.

Let $V_{\min}$ (V_{max}) designate the minimum (maximum) value of any of the problems (1.10) and (1.11). Let further B_1 (B_2) designate a dual feasible basis in any of the minimization (maximization) problems (1.10) and (1.11). Then, in view of the above statement, we have the inequalities

$$\mathbf{f}_{B_1}^T\mathbf{p}_{B_1}\leqq V_{\min}\leqq E\left[f(\xi_1,\ldots,\xi_s)\right]\leqq V_{max}\leqq \mathbf{f}_{B_2}^T\mathbf{p}_{B_2}, \tag{1.12}$$

where $\mathbf{f}_B$ and $\mathbf{p}_B$ designate the vectors of basic components of $\mathbf{f}$ and $\mathbf{p}$, respectively.

We use some of the basic facts from linear programming and the dual algorithm of Lemke [15] for the solution of the linear programming problem. A simple and elegant presentation for both can be found in Prékopa [22]. For the reader's convenience the dual algorithm is briefly summarized in Section 7.

Note that, as it is customary in linear programming, the term "basis" and the symbol "B" mean a matrix and, at the same time, the collection of its column vectors.

We look for dual feasible bases allowing for inequalities (1.12) and providing us with bounding formulas. If such a bound is not sharp then, starting from the corresponding basis, as an initial dual feasible basis, the dual method of linear programming provides us with a sharp algorithmic bound. It is shown by Prékopa [18], [19], [20] that in case of $s=1$ the dual method can be executed in a very simple manner. We will show that some simplification is possible in the multidimensional case, too.

We can look at the moment problems from a more general point of view, by replacing Chebyshev systems for the matrices $A_1,\ldots,A_s$. Such a generality in handling the problem does not present any new theoretical challenges as compared to the power moment problem, however. On the other hand, the nice formulas that we obtain through Lagrange interpolation polynomials would not be immediately at hand. Therefore we keep the discussion on a more specialized level.

There is one case, however, to which we pay special attention, in addition to the multivariate power moment problem. This is the multivariate binomial moment problem.

We take $Z_j=\{0,\ldots,n_j\}$, $j=1,\ldots,s$, introduce the cross binomial moments of $\xi_1, \dots, \xi_s$ as

$$S_{\alpha_1 \cdots \alpha_s} = E\left[\binom{\xi_1}{\alpha_1} \cdots \binom{\xi_s}{\alpha_s}\right]$$

and formulate the problems

$$\begin{aligned}
&\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s} \\
&\text{subject to} \\
&\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} \binom{i_1}{\alpha_i} \cdots \binom{i_s}{\alpha_s} p_{i_1 \cdots i_s} = S_{\alpha_1 \cdots \alpha_s} \\
&\text{for} \quad 0 \leqq \alpha_j \leqq m_j, \quad j = 1, \dots, s \\
&p_{i_1 \cdots i_s} \geqq 0, \quad \text{all } i_1, \dots, i_s,
\end{aligned} \tag{1.13}$$

and

$$\begin{aligned}
&\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s} \\
&\text{subject to} \\
&\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} \binom{i_1}{\alpha_i} \cdots \binom{i_s}{\alpha_s} p_{i_1 \cdots i_s} = S_{\alpha_1 \cdots \alpha_s} \\
&\text{for} \quad \alpha_j \geqq 0, \quad j = 1, \dots, s, \quad \alpha_1 + \cdots + \alpha_s \leqq m \\
&p_{i_1 \cdots i_s} \geqq 0, \quad \text{all } i_1, \dots, i_s.
\end{aligned} \tag{1.14}$$

Problems (1.8) and (1.13) can be transformed into each other by the multivariate generalization of the transformation presented in Prékopa [20].

Problems (1.9) and (1.14) can also be transformed into each other by another but still simple rule.

It follows that a basis in problem (1.8) is primal (dual) feasible if and only if it is primal (dual) feasible in problem (1.13) and this simple correspondence carries over to problems (1.9) and (1.14), too.

2. Divided differences and Lagrange interpolation

First, let $s = 1$ and, for the sake of simplicity, designate the elements of Z_1 simply by $z_0, \dots, z_n$.

The divided difference of order 0, corresponding to z_i, is $f(z_i)$, by definition. The first order divided difference corresponding to z_{i_1}, z_{i_2} is designated and defined by

$$[z_{i_1}, z_{i_2}; f] = \frac{f(z_{i_2}) - f(z_{i_1})}{z_{i_2} - z_{i_1}},$$

where $z_{i_1} \neq z_{i_2}$. The k-th order divided difference is defined recursively by

$$[z_{i_1}, z_{i_2}, \dots, z_{i_k}, z_{i_{k+1}}; f] = \frac{[z_{i_2}, \dots, z_{i_{k+1}}; f] - [z_{i_1}, \dots, z_{i_k}; f]}{z_{i_{k+1}} - z_{i_1}},$$

where $z_{i_1}, \dots, z_{i_{k+1}}$ are pairwise different.

For the case of an arbitrary s, the divided difference corresponding to a subset

$$\mathbf{Z}_{I_1 \cdots I_s} = \{z_{1i}, i \in I_1\} \times \cdots \times \{z_{si}, i \in I_s\} = Z_{1I_1} \times \cdots \times = Z_{sI_s}$$

of the set $\mathbf{Z}$ can be defined in an iterative manner in such a way that first we take the k_1-th order divided difference of f with respect to z_1, where $k_1 = |I_1| - 1$, then the k_2-th order divided difference of that with respect to z_2, where $k_2 = |I_2| - 1$, etc. This can be executed in a mixed manner, the result will always be the same.

Let $[z_{1i}, i \in I_1; \dots; z_{si}, i \in I_s; f]$ designate this divided difference and call it of order $(k_1, \dots, k_s)$. The sum $k_1 + \cdots + k_s$ will be called the total order of the divided difference.

The set on which the above divided difference is defined is the Cartesian product of sets on the real line. Let us term such sets *rectangular*. Divided differences on non-rectangular sets have also been defined in the literature (see, e.g., Karlin, Micchelli and Rinott [11]). These require, however, smooth functions while ours are defined on discrete sets.

A Lagrange interpolation polynomial corresponding to the points in $\{z_{1i}, i \in I_1\} \times \cdots \times \{z_{si}, i \in I_s\}$ is defined by the equation

$$L_{I_1 \cdots I_s}(z_1, \dots, z_s) = \sum_{i_1 \in I_1} \cdots \sum_{i_s \in I_s} f(z_{1i_1}, \dots, z_{si_s}) L_{I_1 i_1}(z_1) \cdots L_{I_s i_s}(z_s), \tag{2.1}$$

where

$$L_{I_j i_j}(z_j) = \prod_{h \in I_j - \{i_j\}} \frac{z_j - z_{jh}}{z_{ji_j} - z_{jh}}, \qquad j = 1, \dots, s. \tag{2.2}$$

The polynomial (2.1) coincides with the function f at every point of the set $Z_{I_1 \cdots I_s}$ and is of degree $m_1 \cdots m_s$.

Newton's form of the Lagrange polynomial (2.1) can be given as follows. Let us order each set Z_{jI_j} and let $I_j^{(k_j)}$ designate the first $k_j + 1$ elements of I_j, $0 \leqq k_j \leqq m_j$, $j = 1, \dots, s$. Then the required form is

$$L_{I_1 \cdots I_s}(z_1, \dots, z_s) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_s=0}^{m_s} \prod_{j=1}^{s} \prod_{h \in I_j^{(k_j-1)}} (z_j - z_{jh}) \left[z_{1h}, h \in I_1^{(k_1)}; \dots; z_{sh}, h \in I_s^{(k_s)}; f\right]. \tag{2.3}$$

Let us introduce the notations $\mathbf{b}(z_1,\dots,z_s)$, $\widetilde{\mathbf{b}}(z_1,\dots,z_s)$, where

$$\mathbf{b}(z_1,\dots,z_s)=(1,z_1,\dots,z_1^{m_1})\otimes\cdots\otimes(1,z_s,\dots,z_s^{m_s})$$

and $\widetilde{\mathbf{b}}(z)$ is obtained from $\mathbf{b}(z)$ by deleting those components $z_1^{\alpha_1}\cdots z_s^{\alpha_s}$ for which $\alpha_1+\cdots+\alpha_s>m$. Then we have the equalities

$$\mathbf{b}=E[\mathbf{b}(\xi_1,\dots,\xi_s)],$$

$$\widetilde{\mathbf{b}}=E[\widetilde{\mathbf{b}}(\xi_1,\dots,\xi_s)].$$

Let $U=\{u_1,\dots,u_M\}$ be a set of points in $\mathbb{R}^s$ and $H=\{(\alpha_1,\dots,\alpha_s)\}$ a finite set of s-tuples of nonnegative integers $(\alpha_1,\dots,\alpha_s)$.

We say that the set U admits Lagrange interpolation of type H if for any real function $f(\mathbf{z})$, $\mathbf{z}\in U$, there exists a polynomial $p(\mathbf{z})$ of the form

$$p(\mathbf{z})=\sum_{(\alpha_1,\dots,\alpha_s)\in H}c(\alpha_1,\dots,\alpha_s)z_1^{\alpha_1}\cdots z_s^{\alpha_s}, \tag{2.4}$$

where all $c(\alpha_1,\dots,\alpha_s)$ are real, such that

$$p(u_i)=f(u_i),\qquad i=1,\dots,M. \tag{2.5}$$

Equations (2.5) form a system of linear equations for the coefficients $c(\alpha_1,\dots,\alpha_s)$. If $|H|=M$, then in (2.5) the number of equations is the same as the number of unknowns. Simple linear algebraic facts imply that if U admits Lagrange interpolation of type H, then it admits a unique Lagrange interpolation of type H.

Let B be a basis of the columns of the matrix A and H the collection of all power s-tuples of the components of the vector $\mathbf{b}(z_1,\dots,z_s)$. In this case $|H|=(m_1+1)\cdots(m_s+1)$. Let

$$I=\{(i_1,\dots,i_s)\mid a_{i_1\cdots i_s}\in B\}. \tag{2.6}$$

Then the unique H-type Lagrange polynomial corresponding to the set

$$U=\{(z_{1i_1},\dots,z_{si_s})\mid(i_1,\dots,i_s)\in I\} \tag{2.7}$$

is equal to

$$L_I(z_1,\dots,z_s)=\mathbf{f}_B^TB^{-1}\mathbf{b}(z_1,\dots,z_s). \tag{2.8}$$

Since $\mathbf{b}(z_{1i_1},\dots,z_{si_s})=a_{i_1\cdots i_s}$, it follows that the basis B is dual feasible in the minimization (maximization) problem (1.10) if and only if

$$\begin{aligned} &f(z_1,\dots,z_s)\geqq L_I(z_1,\dots,z_s), &&\text{all } (z_1,\dots,z_s)\in\mathbf{Z}\\ &(f(z_1,\dots,z_s)\leqq L_I(z_1,\dots,z_s), &&\text{all } (z_1,\dots,z_s)\in\mathbf{Z}). \end{aligned} \tag{2.9}$$

Note that in (2.9) equality holds for all $(z_1,\dots,z_s)\in U$.

Let $\widetilde{B}$ be a basis of the columns of $\widetilde{A}$ and H the collection of all power s-tuples of the components of $\widetilde{\mathbf{b}}(z_1,\dots,z_s)$. If we define I and U as

$$I=\{(i_1,\dots,i_s)\mid \widetilde{a}_{i_1\cdots i_s}\in\widetilde{B}\} \tag{2.10}$$

$$U=\{(z_{1i_1},\dots,z_{si_s})\mid (i_1,\dots,i_s)\in I\} \tag{2.11}$$

then

$$L_I(z_1,\dots,z_s)=\mathbf{f}_{\widetilde{B}}^T\widetilde{B}^{-1}\widetilde{\mathbf{b}}(z_1,\dots,z_s) \tag{2.12}$$

is the unique H-type Lagrange polynomial corresponding to the set U.

The dual feasibility of the basis $\widetilde{B}$ in the minimization (maximization) problem means that

$$\begin{aligned} &f(z_1,\dots,z_s)\geqq L_I(z_1,\dots,z_s), && \text{all } (z_1,\dots,z_s)\in\mathbf{Z}\\ &(f(z_1,\dots,z_s)\leqq L_I(z_1,\dots,z_s), && \text{all } (z_1,\dots,z_s)\in\mathbf{Z}),\end{aligned} \tag{2.13}$$

where equality holds in case of $(z_1,\dots,z_s)\in U$.

The inequalities (2.9) and (2.13) are the conditions of optimality of the minimization (maximization) problems (1.10) and (1.11), respectively.

Replacing $(\xi_1,\dots,\xi_s)$ for $(z_1,\dots,z_s)$ and taking expectations, relations (2.9) and (2.13) provide us with bounds for $E[f(\xi_1,\dots,\xi_s)]$ in Cases (a) and (b), respectively. If the basis is also primal feasible, then it is optimal and thus, the obtained inequality is sharp.

3. Inequalities based on rectangular dual feasible bases

In this section we assume that $f(z_1,\dots,z_s)=f_1(z_1)\cdots f_s(z_s)$ for $z_i\in Z_i$, $i=1,\dots,s$.

For each j, $1\leqq j\leqq s$, we consider the univariate moment problem

$$\begin{aligned} &\min(\max)\quad \sum_{i=0}^{n_j} f_j(z_{ji})p_i^{(j)}\\ &\text{subject to}\\ &\sum_{i=0}^{n_j} z_{ij}^{\alpha}p_i^{(j)}=\mu_\alpha^{(j)},\quad \alpha=0,\dots,m_j\\ &\qquad p_i^{(j)}\geqq 0,\quad i=0,\dots,n_j,\end{aligned} \tag{3.1}$$

where $\mu_\alpha^{(j)}=E(\xi_j^\alpha)$, $\alpha=0,\dots,m_j$, $j=1,\dots,s$ are known, together with the sets $Z_j=\{z_{ji},\ i=0,\dots,m_j\}$ and the unknown decision variables are the $p_i^{(j)}=P(\xi_j=z_{ji})$, $i=0,\dots,n_j$, $j=1,\dots,s$.

THEOREM 3.1. *Suppose that $f_j(z) \geqq 0$ for all $z \in Z_j$. If for each j, $1 \leqq j \leqq s$, we are given a B_j that is a dual feasible basis relative to the maximization problem* (3.1), *then $B = B_1 \otimes \cdots \otimes B_s$ is a dual feasible basis relative to the maximization problem* (1.10).

Moreover, if the set of subscripts of B_j is I_j and $L_{I_j}(z)$ is the corresponding univariate Lagrange polynomial, then we have the inequality

$$E\left[f(\xi_1, \ldots, \xi_s)\right] \leqq E\left[L_{I_1}(\xi_1) \cdots L_{I_s}(\xi_s)\right]. \tag{3.2}$$

PROOF. The dual feasibility of the bases $B_1, \ldots, B_s$ means that

$$\begin{aligned} L_{I_1}(z_1) &\geqq f_1(z_1), \quad z_1 \in Z_1 \\ &\vdots \\ L_{I_s}(z_s) &\geqq f_s(z_s), \quad z_s \in Z_s. \end{aligned} \tag{3.3}$$

On the other hand, the unique H-type Lagrange polynomial, with $H = \{(\alpha_1, \ldots, \alpha_s) \mid 0 \leqq \alpha_j \leqq m_j,\ \alpha_j \text{ integer},\ j = 1, \ldots, s\}$, is given by (2.8). Since $f(z_1, \ldots, z_s) = f_1(z_1) \cdots f_s(z_s)$, it follows that the polynomial (2.8) takes the form

$$L_{I_1 \ldots I_s}(z_1, \ldots, z_s) = L_{I_1}(z_1) \cdots L_{I_s}(z_s). \tag{3.4}$$

Since the dual feasibility of B relative to the maximization problem (1.10) is the same as the second inequality in (2.13), the theorem follows by (3.3) and (3.4). □

THEOREM 3.2. *Suppose that $L_{I_j}(z) \geqq 0$ for all $z \in Z_j$. If for each j, $1 \leqq j \leqq s$, we are given a B_j that is a dual feasible basis relative to the minimization problem* (3.1), *then $B = B_1 \otimes \cdots \otimes B_s$ is a dual feasible basis relative to the minimization problem* (1.10).

Moreover, if the set of subscripts of B_j is I_j and $L_{I_j}(z)$ is the corresponding Lagrange polynomial, then we have the inequality

$$E\left[f(\xi_1, \ldots, \xi_s)\right] \geqq E\left[L_{I_1}(\xi_1) \cdots L_{I_s}(\xi_s)\right]. \tag{3.5}$$

PROOF. The proof is the same as that of Theorem 3.1, with a slight modification. □

Theorem 3.1, combined with the one-dimensional dual feasible basis structure theorems of Prékopa [20] provides us with a variety of upper bounds for probabilities and expectations. Below we present a few examples. Define

$$f_j(z) = \begin{cases} 0, & \text{if } z < z_{r_j} \\ 1, & \text{if } z \geqq z_{r_j}. \end{cases}$$

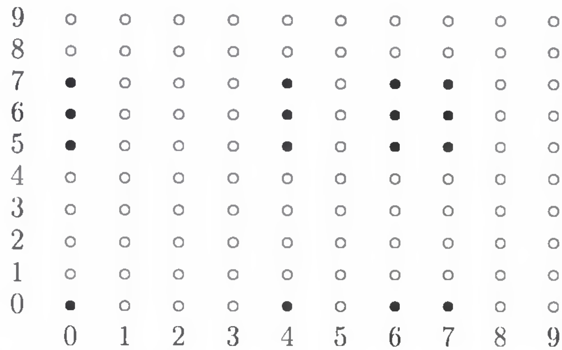

Figure 1. Illustration of a rectangular dual feasible basis through the planar points to which the basic columns of A correspond in the maximization problem (1.10). We chose $m_1+1=4, r_1=4, m_2+1=4, r_2=5$.

EXAMPLE 1. Let $s=2$, $Z_j=\{0,\dots,9\}$, $j=1,2$, $m_1=3$, $m_2=3$ and choose the dual feasible bases, relative to the maximization problem (3.1), as follows:

$$I_1=\{0,r_1,k,k+1\}, \qquad r_1\geqq 1$$
$$I_2=\{0,r_2,t,t+1\}, \qquad r_2\geqq 1.$$

Then we have

$$L_{I_1}(z)=\frac{(z-0)(z-k)(z-k-1)}{(r_1-0)(r_1-k)(r_1-k-1)}+\frac{(z-0)(z-r_1)(z-k-1)}{(k-0)(k-r_1)(k-k-1)}+\frac{(z-0)(z-r_1)(z-k)}{(k+1-0)(k+1-r_1)(k+1-k)}$$

$$L_{I_2}(z)=\frac{(z-0)(z-t)(z-t-1)}{(r_2-0)(r_2-t)(r_2-t-1)}+\frac{(z-0)(z-r_2)(z-t-1)}{(t-0)(t-r_2)(t-t-1)}+\frac{(z-0)(z-r_2)(z-t)}{(t+1-0)(t+1-r_2)(t+1-t)}.$$

To be more specific, let $r_1=4$, $r_2=5$, $k=6$, $t=6$ (see Figure 1). Then the above polynomials take the forms

$$L_{I_1}(z)=\frac{z}{168}\left(z^2-17z+94\right)$$
$$L_{I_2}(z)=\frac{z}{210}\left(z^2-18z+107\right).$$

The inequality (3.2) specializes to

$$P(\xi_1\geqq 4,\ \xi_2\geqq 5)\leqq\frac{1}{35280}(\mu_{33}-18\mu_{32}+107\mu_{31}$$
$$-17\mu_{23}+306\mu_{22}-1819\mu_{21}$$
$$+94\mu_{13}-1619\mu_{12}+10058\mu_{11}).$$

EXAMPLE 2. Let $m_j = 2$, $r_j = 1$, $j = 1, \dots, s$. Then the only dual feasible bases relative to the maximization problem, are those that correspond to the subscript sets $I_j = \{0, 1, n_j\}$, $j = 1, \dots, s$. The Lagrange polynomials take the form

$$L_{I_j}(z) = -\frac{(z - z_{j0})(z - z_{jn_j})}{z_{jn_j} - z_{j1}} + \frac{(z - z_{j0})(z - z_{j1})}{(z_{jn_j} - z_{j0})(z_{jn_j} - z_{j1})}, \quad j = 1, \dots, s.$$

In case of $s = 1$ the bound (3.2) is sharp because the unique dual feasible basis must be primal feasible, too (we have assumed that the right-hand side values in problem (3.1) are moments of some random variable which implies that the problem has feasible solution; it has finite optimum, too, because the set of feasible solutions is compact). The bound (3.2) is sharp in the multivariate case, too, in the sense that the basis in problem (1.10), corresponding to the subscript set $I = I_1 \times \cdots \times I_s$ is primal and dual feasible, hence optimal.

Of particular interest is the case where $Z_j = \{0, \dots, n_j\}$, $j = 1, \dots, s$ and the random variable ξ_j is equal to the number of events that occur among some events $E_{j1}, \dots, E_{jn_j}$, $j = 1, \dots, s$. We may write $L_{I_j}(z)$ in the form

$$L_{I_j}(z) = z - \frac{z(z-1)}{2}\frac{2}{n_j}, \quad j = 1, \dots, s$$

from which we derive

$$L_{I_1}(z_1) \cdots L_{I_s}(z_s) = \left[z_1 - \frac{z_1(z_1 - 1)}{2}\frac{2}{n_1}\right] \cdots \left[z_s - \frac{z_s(z_s - 1)}{2}\frac{2}{n_s}\right].$$

Using the cross binomial moments $S_{\alpha_1 \cdots \alpha_s}$, the inequality (3.2) can be obtained. It is also a sharp one. For example, if $s = 2$ then we obtain

$$\begin{aligned} &P((A_{11} \cup \cdots \cup A_{1n_1}) \cap (A_{21} \cup \cdots \cup A_{2n_2})) = P(\xi_1 \geqq 1, \xi_2 \geqq 1) \\ &\leqq S_{1,1} - \frac{2}{n_2}S_{1,2} - \frac{2}{n_1}S_{2,1} + \frac{4}{n_1 n_2}S_{2,2}. \end{aligned}$$

This result was first obtained by Galambos and Xu [6].

The general formula can be written in the form

$$\begin{aligned} &P\Big(\bigcap_{j=1}^{s} \bigcup_{i=1}^{n_j} A_{ji}\Big) = P(\xi_1 \geqq 1, \dots, \xi_s \geqq 1) \\ &\leqq \sum_{1 \leqq \alpha_j \leqq 2, j=1,\dots,s} (-1)^{\alpha_1 + \cdots + \alpha_s - s} S_{\alpha_1 \cdots \alpha_s} \frac{1}{n_1^{\alpha_1 - 1}} \cdots \frac{1}{n_s^{\alpha_s - 1}}. \end{aligned} \tag{3.6}$$

EXAMPLE 3. Let $r_j = 1$, $z_{j0} = 0$, $m_j = 2$, $I_j = \{0, z_{ji_j}, z_{ji_j+1}\}$, $j = 1, \dots, s$. Then the basis corresponding to the subscript set I_j is dual feasible in the minimization problem (3.1). The Lagrange polynomial $L_{I_j}(z)$ takes the form

$$\begin{aligned} L_{I_j}(z) &= \frac{z(z - z_{ji_j+1})}{z_{ji_j}(z_{ji_j} - z_{ji_j+1})} + \frac{z(z - z_{ji_j})}{z_{ji_j+1}(z_{ji_j+1} - z_{ji_j})} \\ &= \frac{z(z_{ji_j} + z_{ji_j+1} - z)}{z_{ji_j} - z_{ji_j+1}}. \end{aligned}$$

This polynomial is nonnegative for $0 \leqq z \leqq z_{n_j}$ iff $z_{ji_j} + z_{ji_j+1} \geqq z_{n_j}$.

In the special case where $\{z_{j0}, \dots, z_{jn_j}\} = \{0, \dots, n_j\}$, the nonnegativity condition for $L_{I_i}(z)$ is that $2i_j + 1 \geqq n_j$. Assuming this to be the case, for each j, $1 \leqq j \leqq s$, we may write

$$P(\xi_1 \geqq 1, \dots, \xi_s \geqq 1) \geqq E\left[\prod_{j=1}^{s} \frac{\xi_j(2i_j + 1 - \xi_j)}{i_j(i_j+1)}\right]. \tag{3.7}$$

If ξ_j, $j = 1, \dots, s$ designate the occurrences concerning the event sets $E_{j1}, \dots, E_{jn_j}$, $j = 1, \dots, s$, respectively, then it is desirable to give (3.7) another form, expressed in terms of the cross binomial moments. For the case of $s = 2$ the inequality (3.7) gives the following result

$$\begin{aligned} P(\xi_1 \geqq 1, \xi_2 \geqq 1) &\geqq E\left[\prod_{j=1}^{2}\left[\frac{2\xi_j}{i_j+1} - \frac{2}{i_j(i_j+1)}\binom{\xi_j}{2}\right]\right] \\ &= \frac{4}{(i_1+1)(i_2+1)} S_{1,1} - \frac{4}{(i_1+1)i_2(i_2+1)} S_{1,2} \\ &\quad - \frac{4}{(i_2+1)i_1(i_1+1)} S_{2,1} + \frac{4}{i_1(i_1+1)i_2(i_2+1)} S_{2,2}. \end{aligned} \tag{3.8}$$

For an arbitrary s the formula is:

$$\begin{aligned} &P\Big(\bigcap_{j=1}^{s} \bigcup_{i=1}^{n_j} A_{ji}\Big) \\ &= P(\xi_1 \geqq 1, \dots, \xi_s \geqq 1) \\ &\geqq \sum_{\substack{\alpha_1 + \cdots + \alpha_s \leqq 2s \\ 1 \leqq \alpha_j \leqq 2, j = 1, \dots, s}} (-1)^{\alpha_1 + \cdots + \alpha_s - s} S_{\alpha_1 \cdots \alpha_s} \frac{4}{i_1^{\alpha_1 - 1}(i_1+1) \cdots i_s^{\alpha_s - 1}(i_s+1)}, \end{aligned} \tag{3.9}$$

where it is assumed that $i_j \geqq (n_j - 1)/2$, $j = 1, \dots, s$.

It should be mentioned, in connection with problem (1.10), that the optimal basis is not necessarily a rectangular one as it has been shown by

Prékopa [21]. We can reach the optimal basis by starting from any dual feasible basis and carry out the dual method for solving problem (1.10).

In order to find a good rectangular dual feasible basis we can choose I_j, $j=1,\dots,s$ in such a way that I_j is optimal for problem (3.1), provided that it is a maximization problem. In case of the minimization problem we choose the best among those dual feasible bases for which the Lagrange polynomial is nonnegative. The term best means that any dual step that improves on the objective function does not preserve the nonnegativity of the Lagrange polynomial.

Note that having the best univariate bases $I_1,\dots,I_s$, the basis $I=I_1\times\dots\times I_s$ is not necessarily the best rectangular basis.

4. Bounds based on multivariate moments of total order m

We assume that the known moments are: $\mu_{\alpha_1\cdots\alpha_s}$, where $\alpha_j\geqq 0$, $j=1,\dots,s$, $\alpha_1+\dots+\alpha_s\leqq m$.

THEOREM 4.1. *Let* $I=\{(i_1,\dots,i_s)|i_j\geqq 0,$ *integers,* $j=1,\dots,s, i_1+\dots+i_s\leqq m\}$ *and assume that all divided differences of total order* $m+1$*, of the function* f*, are nonnegative. Then the following assertions hold.*

(a) *The set of columns* $\{a_{i_1\cdots i_s}|(i_1,\dots,i_s)\in I\}$ *is a basis* B *for the columns of* $\bar{A}$ *in problem* (1.11).

(b) *The Lagrange polynomial* $L_I(z_1,\dots,z_s)$*, corresponding to the points* $\{(z_{1i_1},\dots,z_{si_s})|=(i_1,\dots,i_s)\in I\}$ *is unique and is the following*

$$L_I(z_1,\dots,z_s)$$
$$=\sum_{\substack{i_1+\cdots i_s\leqq m\\ 0\leqq i_j\leqq n_j, j=1,\dots,s}}[z_{10},\dots,z_{1i_1};\dots;z_{s0},\dots,z_{si_s};f]\prod_{j=1}^{s}\prod_{h=0}^{i_j-1}(z_j-z_{jh}),$$

where, by definition, $\prod_{h=0}^{i_j-1}(z_j-z_{jh})=1$, *for* $i_j=0$.

(c) *We have the inequalities*

$$f(z_1,\dots,z_s)\geqq L_I(z_1,\dots,z_s),\qquad \text{for } (z_1,\dots,z_s)\in Z, \tag{4.2}$$

i.e., B *is a dual feasible basis in the minimization problem* (1.11),

$$E[f(\xi_1,\dots,\xi_s)]\geqq E[L_I(\xi_1,\dots,\xi_s)]. \tag{4.3}$$

If B *is also a primal feasible basis in problem* (1.11), *then the inequality* (4.3) *is sharp.*

(d) *If all divided differences of total order* $m+1$ *are nonpositive, then all assertions hold with the difference that* B *is a dual feasible basis in* *the maximization problem* (1.11) *and the inequalities* (4.2), (4.3) *are reversed.*

PROOF. We mention, without proof, that the determinant of B has a simple form

$$|B| = \prod_{j=1}^{s} \prod_{h=0}^{m-1} \prod_{i=0}^{m-(h+1)} (z_{jm-h} - z_{ji})^{h+1}.$$

This implies that $|B| \neq 0$, hence (a) holds. Assertion (b) follows from (2.13).

Let $Z_{hi} = \{z_{h0}, \ldots, z_{hi}\}$, $Z'_{hi} = \{z_{h0}, \ldots, z_{hi}, z_h\}$, $i = 0, \ldots, m$, $h = 1, \ldots, s$ and define the function $R_I(z_1, \ldots, z_s)$, $(z_1, \ldots, z_s) \in Z$ as follows:

$$\begin{aligned} &R_I(z_1, \ldots, z_s) \\ &= \sum_{h=1}^{s} \sum_{\substack{i_h + \cdots + i_s = m \\ 0 \le i_j \le n_j, j=h,\ldots,s}} [z_1; \ldots; z_{h-1}; Z'_{hi_h}; Z_{h+1 i_{h+1}}; \ldots; Z_{si_s}; f] \\ &\quad \times \prod_{l=0}^{i_h} (z_h - z_{hl}) \prod_{j=h+1}^{s} \prod_{k=0}^{i_j - 1} (z_j - z_{jk}). \end{aligned}$$

We show that $L_I(z_1, \ldots, z_s) + R_I(z_1, \ldots, z_s) = f(z_1, \ldots, z_s)$.

The proof can be carried out by induction. For $s = 1$ it reduces to

$$f(z) - L_I(z) = \prod_{z \in I} (z - z_j)[z_i, i \in I; f] \tag{4.5}$$

which is well-known in Lagrange interpolation theory. For the case of $s = 2$ we have

$$\begin{aligned} &L_I(z_1, z_2) \\ &= \sum_{\substack{i_1 + i_2 \le m \\ 0 \le i_j \le n_j, j=1,2}} [z_{10}, \ldots, z_{1i_1}; z_{20}, \ldots, z_{2i_2}; f] \prod_{j=1}^{2} \prod_{h=0}^{i_j - 1} (z_j - z_{jh}), \end{aligned} \tag{4.6}$$

where $\prod_{h=0}^{i_j-1} (z_j - z_{jh}) = 1$ for $i_j = 0$, by definition, and

$$\begin{aligned} R_I(z_1, z_2) &= [z_{10}, \ldots, z_{1m}, z_1; z_{20}; f](z_1 - z_{10}) \cdots (z_1 - z_{1m}) \\ &\quad + [z_{10}, \ldots, z_{1m-1}, z_1; z_{20}, z_{21}; f](z_1 - z_{10}) \cdots (z_1 - z_{1m-1})(z_2 - z_{20}) \\ &\quad \vdots \\ &\quad + [z_{10}, z_1; z_{20}, \ldots, z_{2m}; f](z_1 - z_{10})(z_2 - z_{20}) \cdots (z_2 - z_{2m-1}) \\ &\quad + [z_1; z_{20}, \ldots, z_{2m}, z_2; f](z_2 - z_{20}) \cdots (z_2 - z_{2m}). \end{aligned} \tag{4.7}$$

Combining terms from (4.6) with terms from (4.7), we may write

$$
\begin{aligned}
&L_I(z_1,z_2)+R_I(z_1,z_2)\\
&=\left\{\sum_{i=0}^{m}[z_{10},\dots,z_{1i};z_{20};f]\prod_{h=0}^{i-1}(z_1-z_{1h})\right.\\
&\qquad\left.+[z_{10},\dots,z_{1m},z_1;z_{20};f]\prod_{h=0}^{m}(z_1-z_{1h})\right\}\\
&\quad+\left\{\sum_{i=0}^{m-1}[z_{10},\dots,z_{1i};z_{20},z_{21};f]\prod_{h=0}^{i-1}(z_1-z_{1h})(z_2-z_{20})\right.\\
&\qquad\left.+[z_{10},\dots,z_{1m-1},z_1;z_{20},z_{21};f]\prod_{h=0}^{m-1}(z_1-z_{1h})(z_2-z_{20})\right\}\\
&\quad\vdots\\
&\quad+\left\{[z_{10};z_{20},\dots,z_{2m};f]\prod_{k=0}^{m-1}(z_2-z_{2k})\right.\\
&\qquad\left.+[z_{10},z_1;z_{20},\dots,z_{2m};f](z_1-z_{10})\prod_{k=0}^{m-1}(z_2-z_{2k})\right\}\\
&\quad+[z_1;z_{20},\dots,z_{2m},z_2;f]\prod_{k=0}^{m}(z_2-z_{2k})\\
&\quad=f(z_1,z_{20})+[z_1;z_{20},z_{21};f](z_2-z_{20})+\cdots\\
&\quad+[z_1;z_{20},\dots,z_{2m},z_2;f](z_2-z_{20})\cdots(z_2-z_{2m})\\
&\quad=f(z_1,z_2).
\end{aligned}
\tag{4.8}
$$

Assuming that the assertion holds for the case of $s-1$, for any function, we derive the equality

$$
\begin{aligned}
&L_I(z_1,\dots,z_s)+R_I(z_1,\dots,z_s)\\
&=\sum_{i_s=0}^{m}[z_1,\dots,z_{s-1};z_{s0},\dots,z_{si_s};f]\prod_{h=0}^{i_s-1}(z_s-z_{sh})\\
&\quad+[z_1,\dots,z_{s-1};z_{s0},\dots,z_{sm},z_s;f]\prod_{h=0}^{m}(z_s-z_{sh}).
\end{aligned}
\tag{4.9}
$$

By (4.5) we see that this is further equal to $[z_1;\dots;z_{s-1};z_s;f]$ which is the same as $f(z_1,\dots,z_s)$.

Since $R(z_1,\ldots,z_s) \geqq 0$ for every $(z_1,\ldots,z_s) \in Z$, we have the inequality (4.2) and its consequence (4.3). The rest of the theorem follows in a straightforward manner. $\square$

Figure 2 illustrates a dual feasible basis of Theorem 4.1 for the case of $Z_1 = Z_2 = \{0,\ldots,9\}$.

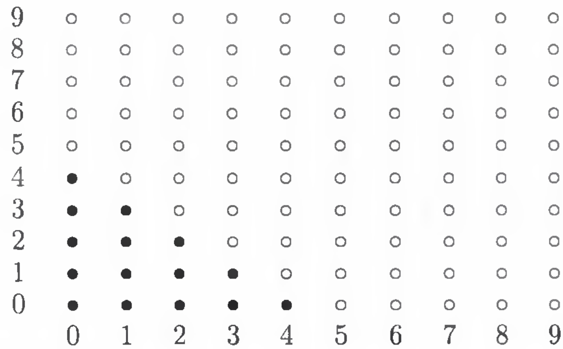

Figure 2. Illustration of a dual feasible basis through the planar points to which the basic columns of A correspond in the minimization problem (1.11). We chose $m+1=5$.

REMARK. If $Z_j = \{0,\ldots,n_j\}$, $j=1,\ldots,s$, then (4.3) can be written in the form

$$E[f(\xi_1,\ldots,\xi_s)] \geqq \sum_{\substack{i_1+\cdots+i_s \leqq m \\ 0 \leqq i_j \leqq n_j, j=1,\ldots,s}} i_1! \cdots i_s! [z_{10},\ldots,z_{1i_1};\ldots;z_{s0},\ldots,z_{si_s};f] S_{i_1\cdots i_s}. \tag{4.10}$$

THEOREM 4.2. *Let $I = \{(i_1,\ldots,i_s) | i_j \geqq 0$, integers, $j=1,\ldots,s, n_1-i_1+\cdots+n_s-i_s \leqq m\}$ and assume that all divided differences of total order $m+1$, of the function f, are nonnegative. Then the following assertions hold.*

(a) *The set of columns $\{a_{i_1\cdots i_s} | (i_1,\ldots,i_s) \in I\}$ is a basis B for the columns of $\tilde{A}$ in problem* (1.11).

(b) *The Lagrange polynomial $L_I(z_1,\ldots,z_s)$, corresponding to the points $\{(z_{1i_1},\ldots,z_{si_s}) | (i_1,\ldots,i_s) \in I\}$, is unique and is the following*

$$L_I(z_1,\ldots,z_s) = \sum_{\substack{i_1+\cdots+i_s \geqq n_1+\cdots+n_s-m \\ 0 \leqq i_j \leqq n_j, j=1,\ldots,s}} [z_{1n_1},\ldots,z_{1n_1-i_1};\ldots;z_{sn_s},\ldots,z_{sn_s-i_s};f] \times \prod_{j=1}^{s} \prod_{h=n_j-i_j+1}^{n_j} (z_j - z_{jh}). \tag{4.11}$$

(c) *If $m+1$ is odd, then*

$$f(z_1,\ldots,z_s) \leqq L_I(z_1,\ldots,z_s), \quad \text{for } (z_1,\ldots,z_s) \in Z, \tag{4.12}$$

i.e., B is a dual feasible basis in the maximization problem (1.11), *and*

$$E[f(\xi_1,\ldots,\xi_s)] \leqq E[L_I(\xi_1,\ldots,\xi_s)]. \tag{4.13}$$

If $m+1$ is even, then the inequalities (4.12) *and* (4.13) *are reversed, i.e., B is a dual feasible basis in the minimization problem* (1.11). *In either case the expectation inequality is sharp, if B is also a primal feasible basis in problem* (1.11).

(d) *If all divided differences of total order $m+1$ are nonpositive, then all assertions hold with the difference that* (4.12) *and* (4.13) *hold for $m+1$ odd, and the reversed inequalities hold for $m+1$ even.*

Proof. The polynomial (4.11) coincides with f at the points $\{(z_{1i_1},\ldots,z_{si_s}),\ (i_1,\ldots,i_s)\in I\}$, for every f. This proves assertions (a) and (b).

Assertion (c) can be proved in the same way as that of Theorem 4.1. If all divided differences that appear in the suitably defined $R_I(z_1,\ldots,z_s)$ are nonnegative (nonpositive), then still the sign of $R_I(z_1,\ldots,z_s)$ depends on the number of factors that multiply the divided differences in each term. Since all factors are nonpositive for all $(z_1,\ldots,z_s)\in Z$ and there are $m+1$ factors in each term, the assertion in (c) follows. Assertion (d) is a trivial modification of assertion (c). □

Figures 3.a and 3.b illustrate dual feasible bases of Theorem 4.2. We chose $Z_1 = Z_2 = \{0,\ldots,9\}$.

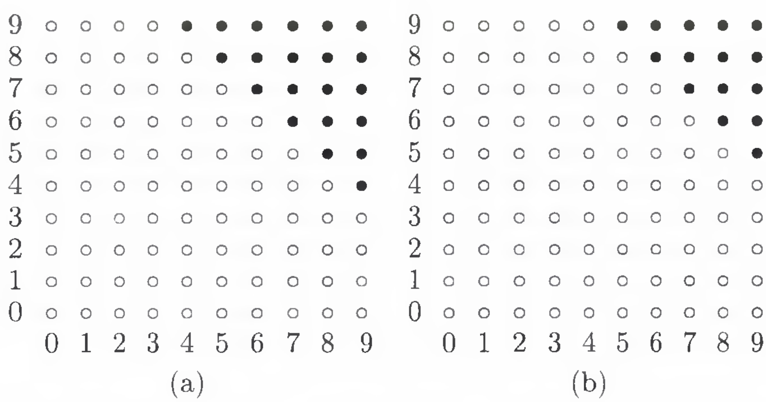

Figure 3. Illustration of dual feasible bases through the planar points to which the basic columns of $\tilde{A}$ correspond in the minimization problem (1.11). The basis in Figure 3.a (3.b) yields an upper (lower) bound because $m+1=5$ is odd ($m+1=4$ is even).

REMARK. If $Z_j = \{0, \dots, n_j\}$, $j = 1, \dots, s$, then the inequality (4.13) can be written in the form

$$
\begin{aligned}
&E[f(\xi_1, \dots, \xi_s)] \\
(4.14)\quad &\leqq \sum_{\substack{i_1+\cdots+i_s \geqq n_1+\cdots+n_s-s \\ 0 \leqq i_j \leqq n_j, j=1,\dots,s}} (n_1-i_1)! \cdots (n_s-i_s)! [n_1, \dots, n_1-i_1; \dots; n_s, \dots, n_s-i_s; f] \\
&\times S_{n_1-i_1 \cdots n_s-i_s}.
\end{aligned}
$$

5. Some bivariate inequalities

THEOREM 5.1. *Let* $I = \{(0,0), (1,0), (0,1), (n_1, 0), (0, n_2), (n_1, n_2)\}$ *and assume that all divided differences of total order* 3 *of the function* f *are nonnegative. Then the following assertions hold.*

(a) *The set of columns* $\{a_{i_1 i_2} | (i_1, i_2) \in I\}$ *is a basis* B *for the columns of* $\bar{A}$ *in problem* (1.11).

(b) *The Lagrange polynomial* $L_I(z_1, z_2)$ *corresponding to the points* $\{(z_{i_1}, z_{i_2}) \,|\, (i_1, i_2) \in I\}$ *is unique and is the following*

$$
\begin{aligned}
L_I(z_1, z_2) = {} & f(z_{10}, z_{20}) + [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\
& + [z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) \\
& + [z_{10}, z_{11}, z_{1n_1}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}) \qquad (5.1) \\
& + [z_{10}; z_{20}, z_{21}, z_{2n_2}; f](z_2 - z_{20})(z_2 - z_{21}) \\
& + [z_{10}, z_{1n_1}; z_{20}, z_{2n_2}; f](z_1 - z_{10})(z_2 - z_{20}).
\end{aligned}
$$

(c) *We have the inequalities*

$$(5.2) \qquad f(z_1, z_2) \leqq L_I(z_1, z_2) \quad \textit{for } (z_1, z_2) \in \mathbf{Z},$$

i.e., B *is dual feasible in the maximization problem* (1.11),

$$(5.3) \qquad E[f(\xi_1, \xi_2)] \leqq E[L_I(\xi_1, \xi_2)].$$

If B *is also a primal feasible basis in problem* (1.11), *then the inequality* (5.3) *is sharp.*

(d) *If all divided differences of total order* 3 *are nonpositive, then all assertions hold with the difference that* B *is dual feasible in the minimization problem* (1.11) *and the inequalities* (5.2) *and* (5.3) *are reversed.*

PROOF OF (a). It is a simple exercise to check that $|B| \neq 0$. Thus, B is in fact a basis.

PROOF OF (b). The uniqueness of the Lagrange polynomial follows from (a).

That the polynomial $L_I(z_1, z_2)$, given by (5.1), is the Lagrange polynomial corresponding to the points $\{(z_{1i_1}, z_{2i_2}), (i_1, i_2) \in I\}$, follows from the fact that $L_I(z_1, z_2)$ coincides with $f(z_1, z_2)$ on these points.

Now we show that (5.2) holds. First we assume that $z_1 > z_{10}$, $z_2 > z_{20}$.

In view of the assumption that the $(2,1), (1,2)$-order divided differences are nonnegative, we have the inequalities

$$[z_{10}, z_1; z_{20}, z_2; f] \leqq [z_{10}, z_{1n_1}; z_{20}, z_2; f] \leqq [z_{10}, z_{1n_1}; z_{20}, z_{2n_2}; f].$$

It follows from this that

$$\begin{aligned} f(z_1, z_2) \leqq\ & f(z_{10}, z_2) + f(z_1, z_{20}) - f(z_{10}, z_{20}) \\ & + [z_{10}, z_{1n_1}; z_{20}, z_{2n_2}; f](z_1 - z_{10})(z_2 - z_{20}). \end{aligned} \tag{5.4}$$

On the other hand, the nonnegativity of the $(3,0)$-order divided differences and the fact that $\{0, 1, n_1\}$ is a univariate dual feasible basis structure in the problem

$$\begin{aligned} & \max \quad \sum_{i=0}^{n_1} f(z_{1i}, z_{20}) p_i^{(1)} \\ & \text{subject to} \\ & \sum_{i=0}^{n_1} z_{1i}^{\alpha} p_i^{(1)} = \mu_{\alpha 0}, \quad \alpha = 0, 1, 2 \\ & p_i^{(1)} \geqq 0, \qquad i = 0, \ldots, n_1, \end{aligned} \tag{5.5}$$

(see [20]) imply that

$$\begin{aligned} f(z_1, z_{20}) \leqq\ & f(z_{10}, z_{20}) + [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\ & + [z_{10}, z_{1n_1}, z_{12}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}). \end{aligned} \tag{5.6}$$

In a similar way we obtain

$$\begin{aligned} f(z_{10}, z_2) \leqq\ & f(z_{10}, z_{20}) + [z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) \\ & + [z_{10}; z_{20}, z_{21}; z_{2n_2}; f](z_2 - z_{20})(z_2 - z_{21}). \end{aligned} \tag{5.7}$$

The inequalities (5.4), (5.6) and (5.7) imply (5.2).

The expectation inequality (5.3) follows from (5.2). If B is also primal feasible, then it is optimal in the maximization problem (1.11), hence the inequality is sharp. Assertion (d) follows from the fact that in this case the function $-f$ has nonnegative divided differences of total order 3, hence the inequalities (5.4), (5.6) and (5.7) hold if we replace $-f$ for f. These imply the reversed inequalities of (5.2) and (5.3).

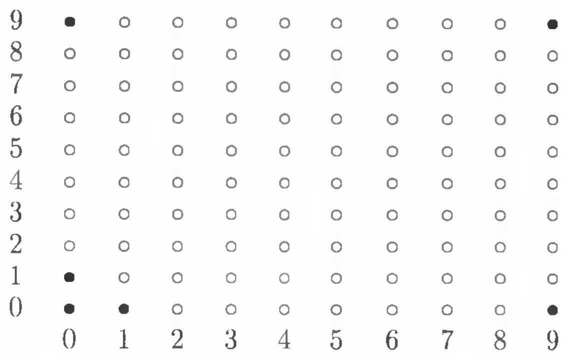

Figure 4. Illustration of a dual feasible basis through the planar points to which the basic columns of $\tilde{A}$ correspond in the maximization problem (1.11).

If $z_1 = z_{10}$ and/or $z_2 = z_{20}$, (5.2) reduces to (5.7) or (5.6). This completes the proof. □

A dual feasible basis of Theorem 5.1 is illustrated in Figure 4.

REMARK. If $Z_1 = \{0, \dots, n_1\}$, $Z_2 = \{0, \dots, n_2\}$, then the inequality in (5.3) takes the form

$$\begin{aligned} E[f(\xi_1, \xi_2)] &\leqq f(0,0) + [0,1;0;f]S_{10} + [0;0,1;f]S_{01} \\ &+ 2[0,1,n_1;0;f]S_{20} + 2[0;0,1,n_2;f]S_{02} + [0,n_1;0,n_2;f]S_{11}. \end{aligned} \tag{5.8}$$

THEOREM 5.2. *Let $I = \{(0,0),(0,1),(1,0),(1,1),(0,n_2),(n_1,0)\}$ and assume that all divided differences of orders $(2,1),(1,2)$ of the function f are nonnegative while the divided differences of orders $(3,0),(0,3)$ are nonpositive. Then the following assertions hold.*

(a) *The set of columns $\{a_{i_1 i_2} | (i_1, i_2) \in I\}$ is a basis B for the columns of $\tilde{A}$ in problem* (1.11).

(b) *The Lagrange polynomial $L_I(z_1, z_2)$, corresponding to the points $\{(z_{i_1}, z_{i_2}) | (i_1, i_2) \in I\}$ is unique and is the following*

$$\begin{aligned} L_I(z_1, z_2) = f(z_{10}, z_{20}) &+ [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\ &+ [z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) \\ &+ [z_{10}, z_{11}, z_{1n_1}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}) \\ &+ [z_{10}; z_{20}, z_{21}, z_{2n_2}; f](z_2 - z_{20})(z_2 - z_{21}) \\ &+ [z_{10}, z_{11}; z_{20}, z_{21}; f](z_1 - z_{10})(z_2 - z_{20}). \end{aligned} \tag{5.9}$$

(c) *We have the inequalities*

$$f(z_1, z_2) \geqq L_I(z_1, z_2) \quad \textit{for } (z_1, z_2) \in \mathbf{Z}, \tag{5.10}$$

i.e., B is dual feasible in the minimization problem (1.11),

$$E[f(\xi_1, \xi_2)] \geqq E[L_I(\xi_1, \xi_2)]. \tag{5.11}$$

If B is also a primal feasible basis in problem (1.11), *then the expectation inequality* (5.11) *is sharp.*

(d) *If all divided differences of orders* $(2,1),(1,2)$ *are nonpositive, and those of orders* $(3,0),(0,3)$ *are nonnegative, then all assertions hold with the difference that B is dual feasible in the maximization problem* (1.11) *and the inequalities* (5.10) *and* (5.11) *are reversed.*

PROOF. The proof is very similar to that of Theorem 5.1 and is omitted. □

Figure 5 illustrates a dual feasible basis of Theorem 5.2.

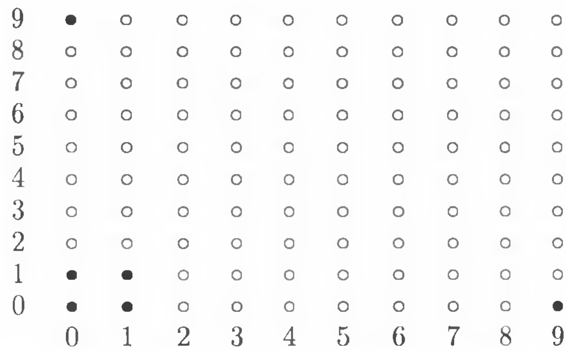

Figure 5. Illustration of a dual feasible basis through the planar points to which the basic columns of $\tilde{A}$ correspond in the minimization problem (1.11).

REMARK. If $Z_1=\{0,\dots,n_1\}$, $Z_2=\{0,\dots,n_2\}$, then the inequality in (5.11) can be written in the form

$$\begin{aligned} E[f(\xi_1,\xi_2)] \geqq f(0,0) &+ [0,1;0;f]S_{10} \\ &+ [0;0,1;f]S_{01} + 2[0,1,n_1;n_2 f]S_{20} \\ &+ 2[0;0,1,n_2;f]S_{02} + [0,1;0,1;f]S_{11}. \end{aligned} \tag{5.12}$$

The proof of Theorem 5.1 allows for the derivation of simular dual feasibility assertions for other lattices, using other assumptions. For example, if $m=2$ and the divided differences of orders $(2,0),(2,1),(1,2)$ and $(0,4)$ are nonnegative, then the set of points $\{(z_{1i_1},z_{2i_2})|(i_1,i_2)\in I\}$, with $I=\{(0,0),(0,1),(0,2),(0,3),(1,0),(1,1)\}$, determines a unique Lagrange interpolation and a dual feasible basis in the minimization problem (1.11). The Lagrange polynomial is

$$\begin{aligned} L_I(z_1,z_2) = [z_{10};z_{20};f] &+ [z_{10};z_{10},z_{21};f](z_2-z_{20}) \\ &+ [z_{10};z_{20},z_{21},z_{22};f](z_2-z_{20})(z_2-z_{21}) \\ &+ [z_{10};z_{20},z_{21},z_{22},z_{23};f](z_2-z_{20})(z_2-z_{21})(z_2-z_{22}) \\ &+ [z_{10},z_{11};z_{20};f](z_1-z_{10}) \\ &+ [z_{10},z_{11};z_{20},z_{21};f](z_1-z_{10})(z_2-z_{20}). \end{aligned}$$

From here inequalities of the type (5.2) and (5.3) can be derived for $s=2$.

6. Multivariate Bonferroni inequalities

In this section we assume that $Z_j = \{0, \ldots, n_j\}$, $j = 1, \ldots, s$. Defining

$$g(z) = \begin{cases} 0, & \text{if } z = 0 \\ 1, & \text{if } z \geqq 1, \end{cases}$$

we easily see that

$$[0, \ldots, i; g] = (-1)^{i-1} \frac{1}{i!}, \quad \text{for } i \geqq 1.$$

Let $f(z_1, \ldots, z_s) = g(z_1) \cdots g(z_s)$, $(z_1, \ldots, z_s) \in \mathbf{Z}$. Then we have

$$\begin{aligned} [0, \ldots, i_1; \ldots; 0, \ldots, i_s; f] &= \prod_{j=1}^{s} [0, \ldots, i_j; g] \\ &= \prod_{j=1}^{s} (-1)^{i_j - 1} \frac{1}{i_j!} \text{ for } i_j \geqq 1, j = 1, \ldots, s. \end{aligned} \tag{6.1}$$

If for at least one j we have $i_j = 0$, then the above divided difference is 0.

Let $A_{j1}, \ldots, A_{jn_j}$, $j = 1, \ldots, s$ be s finite sequences of arbitrary events and let ξ_j designate the number of those, in the jth sequence, that occur. Then $\xi_j \geqq 1$ is the same as $\cup_{i=1}^{n_j} A_{ji}$. Now, Theorem 4.1 and relation (6.1) imply

THEOREM 6.1. *If $m + 1 - s$ is even, then we have*

$$P\left(\bigcap_{j=1}^{s} \bigcup_{i=1}^{n_j} A_{ji}\right) \geqq \sum_{\substack{i_1 + \cdots + i_s \leq m \\ 1 \leqq i_j \leqq n_j, j = 1, \ldots, s}} (-1)^{i_1 + \cdots + i_s - s} S_{i_1 \cdots i_s} \tag{6.2}$$

and if $m + 1 - s$ is odd, then we have

$$P\left(\bigcap_{j=1}^{s} \bigcup_{i=1}^{n_j} A_{ji}\right) \leqq \sum_{\substack{i_1 + \cdots + i_s \leq m \\ 1 \leqq i_j \leqq n_j, j = 1, \ldots, s}} (-1)^{i_1 + \cdots + i_s - s} S_{i_1 \cdots i_s}. \tag{6.3}$$

For the case of $s = 1$ we obtain

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geqq \sum_{i=1}^{m} (-1)^{i-1} S_i,$$

if m is even, and

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leqq \sum_{i=1}^{m} (-1)^{i-1} S_i,$$

if m is odd. These are the original Bonferroni inequalities (see [1]*).*

We can also deduce inequalities for $P(\xi_1 = r_1, \dots, \xi_s = r_s)$ and $P(\xi_1 \geqq r_1, \dots, \xi_s \geqq r_s)$.

In the first case we define

$$f_j(z) = \begin{cases} 0, & \text{if } z \neq r_j \\ 1, & \text{if } z = r_j \end{cases}$$

and $f(z_1, \dots, z_s) = f_1(z_1) \cdots f_s(z_s)$ for $(z_1, \dots, z_s) \in Z$. By the determinantal form of the univariate divided differences (see, e.g., Jordan (1947)) we easily deduce that

$$[0, \dots, i_j; f_j] = (-1)^{i_j - r_j} \frac{1}{i_j!} \binom{i_j}{r_j}$$

which implies that

$$\begin{aligned} [0, \dots, i_1; \dots; 0, \dots, i_s; f] &= \prod_{j=1}^{s} [0, \dots, i_j; f_j] \\ &= \prod_{j=1}^{s} (-1)^{i_j - r_j} \frac{1}{i_j!} \binom{i_j}{r_j}. \end{aligned} \tag{6.4}$$

This is nonnegative if $i_1 + \cdots + i_s - (r_1 + \cdots + r_s)$ is even, otherwise it is nonpositive. Hence Theorem 4.1 implies

THEOREM 6.2. *If $m + 1 - (r_1 + \cdots + r_s)$ is even, then we have*

$$P(\xi_1 = r_1, \dots, \xi_s = r_s) \geqq \sum_{\substack{i_1 + \cdots + i_s \leqq m \\ r_j \leqq i_j \leqq n_j, j=1,\dots,s}} \prod_{j=1}^{s} (-1)^{i_j - r_j} \binom{i_j}{r_j} S_{i_1 \cdots i_s} \tag{6.5}$$

and if $m + 1 - (r_1 + \cdots + r_s)$ is odd, then we have

$$P(\xi_1 = r_1, \dots, \xi_s = r_s) \leqq \sum_{\substack{i_1 + \cdots + i_s \leqq m \\ r_j \leqq i_j \leqq n_j, j=1,\dots,s}} \prod_{j=1}^{s} (-1)^{i_j - r_j} \binom{i_j}{r_j} S_{i_1 \cdots i_s}. \tag{6.6}$$

Finally, in order to obtain inequalities for $P(\xi_1 \geqq r_1, \dots, \xi_s \geqq r_s)$ we define

$$f_j(z) = \begin{cases} 0, & \text{if } z < r_j \\ 1, & \text{if } z \geqq r_j \end{cases}$$

and $f(z_1,\ldots,z_s)=f_1(z_1)\cdots f_s(z_s)$ for $(z_1,\ldots,z_s)\in\mathbf{Z}$. Again, using the determinantal form of the univariate divided differences, we get

$$[0,\ldots,i_j;f]=\sum_{h=r_j}^{i_j}(-1)^{i_j-h}\frac{1}{i_j!}\binom{i_j}{h}.$$

On the other hand, we have the combinatorial identity

$$\sum_{h=r}^{i}(-1)^{i-h}\binom{i}{h}=(-1)^{i-r}\binom{i-1}{r-1}.$$

Thus, we have the following formula for the multivariate divided differences

$$\begin{aligned}[0,\ldots,i_1;\ldots;0,\ldots,i_s;f]&=\prod_{j=1}^{s}[0,\ldots,i_j;f_j]\\ &=\prod_{j=1}^{s}(-1)^{i_j-r_j}\frac{1}{i_j!}\binom{i_j-1}{r_j-1}.\end{aligned}\tag{6.7}$$

Theorem 4.1 and equation (6.7) imply

THEOREM 6.3. *If $m+1-(r_1+\cdots+r_s)$ is even, then we have the inequality*

$$\begin{aligned}P(\xi_1\geqq r_1,\ldots,\xi_s\geqq r_s)\geqq&\sum_{\substack{i_1+\cdots+i_s\leqq m\\ r_j\leqq i_j\leqq n_j,\,j=1,\ldots,s}}\prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j-1}{r_j-1}\\ &=S_{i_1\cdots i_s}\end{aligned}\tag{6.8}$$

and if $m+1-(r_1+\cdots+r_s)$ is odd, then we have the inequality

$$\begin{aligned}P(\xi_1\geqq r_1,\ldots,\xi_s\geqq r_s)\leqq&\sum_{\substack{i_1+\cdots+i_s\leqq m\\ r_j\leqq i_j\leqq n_j,\,j=1,\ldots,s}}\prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j-1}{r_j-1}\\ &=S_{i_1\cdots i_s}.\end{aligned}$$

The inequalities (6.2), (6.3), (6.4), (6.6), (6.7) and (6.8) have been derived first by Meyer [16].

7. Algorithmic bounds and numerical examples

The significance of the knowledge of a dual feasible basis is twofold. First, we can immediately present bound for the optimum value of the linear programming problem we are dealing with. Second, starting from this basis, we have an algorithmic tool with the aid of which we can improve on the bound or obtain the best possible bound. This tool is the dual method of linear programming, due to Lemke [15]. For a short and elegant description of it see [22].

Given a linear programming problem

$$\begin{aligned} &\min(\max)\ \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \\ &A\mathbf{x}=\mathbf{b} \\ &\mathbf{x}\geqq\mathbf{0}, \end{aligned} \tag{7.1}$$

where A is an $m\times n$ matrix ($m\leqq n$), assumed to be of full rank, any basis B is a nonsingular $m\times m$ part of A. We say that B is feasible or primal feasible if the solution of the equation $B\mathbf{x}_B=\mathbf{b}$ produces $\mathbf{x}_B\geqq\mathbf{0}$. Let I or I_B designate the set of subscripts of those columns of A which are in the basis. Further, let $\mathbf{c}_B$ designate the vector of components c_i, $i\in I$, arranged in the same order as they are in $\mathbf{c}$.

The basis B is said to be dual feasible if the solution of the equation $\mathbf{y}^TB=\mathbf{c}_B^T$ satisfies the constraints of the dual of problem (7.1):

$$\begin{aligned} &\max(\min)\ \mathbf{b}^T\mathbf{y} \\ &\text{subject to} \\ &A^T\mathbf{y}\leqq(\geqq)\mathbf{c}. \end{aligned} \tag{7.2}$$

If B is both primal and dual feasible, then it is optimal.

Let $A=(a_1,\dots,a_n)$, $\mathbf{c}^T=(c_1,\dots,c_n)$. With these notations the dual feasibility of B can be formulated as follows:

$$\mathbf{c}_B^TB^{-1}a_h\leqq(\geqq)c_h,\quad h=1,\dots,n. \tag{7.3}$$

For $h\in I$ equality holds in (7.3).

The dual method of linear programming starts from a dual feasible basis B. Then the following steps are performed. We assume the problem is a minimization problem.

STEP 1. Check if $B^{-1}\mathbf{b}\geqq\mathbf{0}$, i.e., the basis B is primal feasible. If yes, then stop, optimal basis has been found. Otherwise go to Step 2.

STEP 2. Pick any negative component of $B^{-1}\mathbf{b}$. If it is the ith one, then delete the ith vector from B. Go to Step 3.

STEP 3. Determine the incoming vector maintaining dual feasibility of the basis and making the objective function value nondecreasing. Go to Step 1.

Step 3 is usually costly. In case of the univariate discrete moment problems (see [20]), however, the structure of the dual feasible bases have been found and Step 3 can be carried out by performing simple combinatorial search.

In case of the multivariate discrete moment problems we have only a few dual feasible basis structures and we cannot spare Step 3 when solving the problem to obtain the best possible bound.

Still, the availability of an initial dual feasible basis is of great help. We can save the time needed to execute the first phase in a two-phase solution method that is roughly 50% of the time needed to solve the LP. In addition, since moment problems are numerically very sensitive, the knowledge of an initial dual feasible basis increases numerical stability.

The dual method, as applied to these problems, has many other features. For example, we may have more detailed information about the possible values of the random vector $(\xi_1, \dots, \xi_s)$, i.e., we may know that some of the values in the set $Z = Z_1 \times \cdots \times Z_s$ are not possible, in other words, have probability 0. Information of this type has not been exploited so far in former sections of the paper. The dual method, however, allows to take such information into account, in a trivial way. In fact, we simply have to delete those columns from the problem that are multiplied by the probabilities known to be 0. The basis remains dual feasible with respect to the new problem. This way we even improve on the bound.

Below we present one small numerical example for illustration.

Let $n_1 = n_2 = 9$, $m_1 + m_2 = 3$. The following power moments have been obtained from the uniform distribution: $p_{i_1 i_2} = 1/100$ for each $0 \leqq i_1, i_2 \leqq 9$:

$$\begin{array}{llll} \mu_{00} = 1, & \mu_{10} = 4.5, & \mu_{20} = 28.5, & \mu_{30} = 202.5 \\ \mu_{01} = 4.5, & \mu_{11} = 20.25, & \mu_{21} = 128.25, & \\ \mu_{02} = 28.5, & \mu_{12} = 128.25, & & \\ \mu_{03} = 202.5. & & & \end{array}$$

We want to obtain the sharp lower bound for $P(\xi_1 \geqq 1, \xi_2 \geqq 1)$. We start from the dual feasible basis with subscript set $I = \{(0,0), (1,0), (0,1), (2,0), (1,1), (0,2), (3,0), (2,1), (1,2), (0,3)\}$.

As optimal solution, for the minimization problem (1.11), we obtain

$$\begin{array}{lll} p_{40} = 0.075, & p_{50} = 0.125, & p_{90} = 0, \\ p_{04} = 0.175, & p_{94} = 0.125, & p_{45} = 0.225, \\ p_{95} = 0.075, & p_{09} = 0.025, & p_{59} = 0.175, \\ p_{99} = 0, & \text{and all other } p_{i_1 i_2} = 0. & \end{array}$$

The value of the objective function is the sum of those $p_{i_1 i_2}$ probabilities for which we have $i_1 \geqq 1$, $i_2 \geqq 1$. This sum equals 0.6. Thus, the result is

$$P(\xi_1 \geqq 1,\ \xi_2 \geqq 1) \geqq 0.6.$$

Note that the true probability is the sum of those $p_{i_1 i_2} = 1/100$, for which $i_1 \geqq 1, i_2 \geqq 1$. This number is 0.81.

Suppose now that we have the information concerning ξ_1 and ξ_2 that $\xi_1 + \xi_2 \leqq 12$. This means that the set of possible values of the random vector (ξ_1, ξ_2) is only a subset of the set $\{(i,j) | 0 \leqq i \leqq 9,\ 0 \leqq j \leqq 9\}$. Thus, we may delete those columns, variables and objective function coefficients from problem (1.11) which correspond to (i,j) with $i + j > 12$. Solving the restricted problem, the optimal solution is

$$\begin{aligned} &p_{30} = 0.11393, \quad p_{03} = 0.09749, \quad p_{04} = 0.08732,\\ &p_{45} = 0.23637, \quad p_{55} = 0.11857, \quad p_{56} = 0.00985,\\ &p_{97} = 0.12531, \quad p_{09} = 0.05191, \quad p_{49} = 0.10666,\\ &p_{59} = 0.05259, \qquad \text{and all other} \ \ p_{i_1 i_2} = 0. \end{aligned}$$

The value of the objective function is $p_{45} + p_{55} + p_{56} + p_{97} + p_{49} + p_{59} = 0.64935$. This improves on the former lower bound that is 0.6.

In case of $m = 3$ the Bonferroni inequality (6.2) produces the irrealistic result:

$$P\left(\bigcap_{j=1}^{2} \bigcup_{i=0}^{9} A_{ji}\right) \geqq S_{11} - S_{12} - S_{21} = 20.25 - 54 - 54 = -87.75.$$

This number is, at the same time, the value of the objective function in case of the initial dual feasible basis.

For further numerical examples see Prékopa [21].

REFERENCES

[1] Bonferroni, C. E., *Teoria statistica delle classi e calcolo delle probabilità*, Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze, Nr. 8, Libr. Internaz. Seeber, Firenze, 1936, 62 pp. *Zbl* **16**, 411

[2] Boros, E. and Prékopa, A., Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur, *Math. Oper. Res.* **14** (1989), 317–342. *MR* **90g**:60021

[3] Dawson, D. and Sankoff, A., An inequality for probabilities, *Proc. Amer. Math. Soc.* **18** (1967), 504–507. *MR* **35** #2305

[4] Dulá, J. H., Bounds on the expectations of convex functions, Ph. D. Thesis, Dept. of Operations and Industrial Engineering, Univ. of Michigan, Ann Arbor, MI, 1986.

[5] Galambos, J. and Simonelli, J., *Bonferroni-type inequalities with applications*, Probability and its Applications, Springer-Verlag, New York, 1996.

[6] Galambos, J. and Xu, Y., Bivariate extension of the method of polynomials for Bonferroni-type inequalities, *J. Multivariate Anal.* **52** (1995), 131–139. *MR* **96b**:60041

[7] Horn, R. A. and Johnson, C. R., *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991. *MR* **92e**:15003

[8] Jordan, C., *Calculus of finite differences*, Chelsea Publishing Co., New York, 1947. Third edition, 1965. *MR* **32** #1463

[9] KALL, P., An upper bound for SLP using first and total second moments, *Ann. Oper. Res.* **30** (1991), 267–276. *MR* **92f**:90045

[10] KARLIN, S. and STUDDEN, W. J., *Tchebycheff systems: With applications in analysis and statistics*, Pure and Applied Mathematics, Vol. 15, Interscience Publishers John Wiley & Sons, New York – London, 1966. *MR* **34** #4757

[11] KARLIN, S., MICCHELLI, C. A. and RINOTT, Y., Multivariate splines: A probabilistic perspective, *J. Multivariate Anal.* **20** (1986), 69–90. *MR* **88a**:41008

[12] KEMPERMAN, J. H. B., The general moment problem, a geometric approach, *Ann. Math. Statist.* **39** (1968), 93–122. *MR* **40** #909

[13] KEMPERMAN, J. H. B. and SKIBINSKY, M., Covariance spaces for measures on polyhedral sets, *Stochastic inequalities* (Seattle, WA, 1991), ed. by M. Shaked and Y. L. Tong, IMS Lecture Notes Monograph Ser., **22**, Inst. Math. Statist., Hayward, CA, 1992, 182–195. *MR* **94g**:60026

[14] KWEREL, S. M., Bounds on the probability of the union and intersection of m events, *Adv. Appl. Probability* **7** (1975), 431–448. *MR* **51** #4341

[15] LEMKE, C. E., The dual method of solving the linear programming problem, *Naval Res. Logist. Quart.* **1** (1954), 36–47. *MR* **16**, 752a

[16] MEYER, R. M., Note on a 'multivariate' form of Bonferroni's inequalities, *Ann. Math. Statist.* **40** (1969), 692–693. *MR* **38** #6643

[17] POPOVICIU, T., *Les fonctions convexes*, Actualités Sci. Ind., no. 992, Hermann et Cie., Paris, 1944. *MR* **8**, 319a

[18] PRÉKOPA, A., Boole–Bonferroni inequalities and linear programming, *Oper. Res.* **36** (1988), 145–162. *MR* **89g**:90163

[19] PRÉKOPA, A., Sharp bounds on probabilities using linear programming, *Oper. Res.* **38** (1990), 227–239. *MR* **91b**:90098

[20] PRÉKOPA, A., The discrete moment problem and linear programming, *Discrete Appl. Math.* **27** (1990), 235–254. *MR* **91j**:90050

[21] PRÉKOPA, A., Inequalities on expectations based on the knowledge of multivariate moments, *Stochastic inequalities* (Seattle, WA, 1991), ed. by M. Shaked and Y. L. Tong, IMS Lecture Notes Monograph Ser., **22**, Inst. Math. Statist., Hayward, CA, 1992, 309–331. *MR* **94h**:60024

[22] PRÉKOPA, A., A brief introduction to linear programming, *Math. Sci.* **21** (1996), 85–111.

[23] SAMUELS, S. M. and STUDDEN, W. J., Bonferroni-type probability bounds as an application of the theory of Tchebycheff systems, *Probability, statistics and mathematics*, Papers in honor of S. Karlin, Academic Press, Boston, MA, 1989, 271–289. *MR* **91c**: 60021

[24] WHEELER, J. C. and GORDON, R. G., Bounds for averages using moment constraints, *The Padé approximant in theoretical physics*, ed. by G. A. Baker, Jr. and J. L. Gammel, Mathematics in Science and Engineering, Vol. 71, Academic Press, New York – London, 1970, Chapter 3, 99–127. *MR* **56** #7587

(Received January 20, 1998)

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

prekopa@cs.elte.hu

and

RUTCOR
RUTGERS CENTER FOR OPERATIONS RESEARCH
P.O. BOX 5062
NEW BRUNSWICK, NJ 08903–5062
U.S.A.

prekopa@rutcor.rutgers.edu

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 379–389

THE RANGE OF A CRITICAL BRANCHING WIENER PROCESS

P. RÉVÉSZ

To the memory of A. Rényi

Abstract

Consider a critical branching Wiener process on $\mathbb{R}^1$. Let $R(n)$ be the range of the locations of the particles at time n. A limit distribution theorem is proved for $n^{-1/2}R(n)$.

1. Introduction

Consider the following

MODEL 1.

(i) a particle starts from the position $0 \in \mathbb{R}^1$ and executes a Wiener process $W(t) \in \mathbb{R}^1$,

(ii) arriving at time $t=1$ to the new location $W(1)$ it dies,

(iii) at death it is replaced by Y offspring where

$$\mathbf{P}\{Y=0\}=\mathbf{P}\{Y=2\}=\frac{1}{2},$$

(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, p. 91 of [1].

Let

(a) $B(n)$ be the number of particles living at time n, the particles born at time n to be counted as alive at time n but not at time $n+1$, i.e. $B(0)=1,\ \mathbf{P}\{B(1)=0\}=\mathbf{P}\{B(1)=2\}=1/2$,

(b) $X_{n1}, X_{n2}, \dots, X_{n,B(n)}$ be the locations of the particles at time n in $\mathbb{R}^1$,

(c) $M_n^+ = \max\{X_{n1}, X_{n2}, \dots, X_{n,B(n)}\}$,

(d) $\mathcal{M}^+(n,x) = \mathbf{P}\{M_n^+ < xn^{1/2} \mid B(n)>0\}$,

(e) $M_n^- = \min\{X_{n1}, X_{n2}, \dots, X_{n,B(n)}\}$,

1991 *Mathematics Subject Classification.* Primary 60G57; Secondary 60J65, 60J80.

Key words and phrases. Critical branching Wiener process, limit distribution, random trees, range.

(f) $\mathcal{M}^-(n,x)=\mathbf{P}\{M_n^- < xn^{1/2} \mid B(n)>0\}$,
(g) $R_n = M_n^+ - M_n^-$,
(h) $\mathcal{R}(n,x)=\mathbf{P}\{R_n < xn^{1/2} \mid B(n)>0\}$,
(i) $F_n(x,y)=\mathbf{P}\{M_n^+ < xn^{1/2},\ M_n^- > yn^{1/2} \mid B(n)>0\}$.

In [2] we studied the limit properties of M_n^+. We proved the following two theorems.

THEOREM A. *There exists a distribution function $\mathcal{M}^+(x)$ $(x\in\mathbb{R}^1)$ such that for any n big enough we have*

$$d(\mathcal{M}^+(n,x),\mathcal{M}^+(x)) \leqq n^{-1/2}(\log n)^4, \tag{1.1}$$

$$1-\mathcal{M}^+(x)+\mathcal{M}^+(-x) \leqq \exp\left(-\frac{x^2}{20}\right) \quad \text{if } x \geqq 200, \tag{1.2}$$

$\mathcal{M}^+(x)$ is a solution of the integral equation

$$F(x)=\int_0^1\int_{-\infty}^{+\infty}(F(\alpha^{-1/2}(x-y)))^2\varphi_\alpha(y)\,dy\,d\alpha, \tag{1.3}$$

where

$$\varphi_\alpha(y)=(2\pi(1-\alpha))^{-1/2}\exp\left(-\frac{y^2}{2(1-\alpha)}\right)$$

and $d(\cdot,\cdot)$ is the Lévy distance.

THEOREM B. *There is only one distribution function which satisfies* (1.2) *and* (1.3).

In the present paper we prove similar results for R_n. Our main result is:

THEOREM 1. *There exist distribution functions $\mathcal{R}(x)$ and $F(x,y)$ such that for any n big enough we have*

$$d(\mathcal{R}(n,x),\mathcal{R}(x)) \leqq n^{-1/2}(\log n)^4, \tag{1.4}$$

$$d(F_n(x,y).F(x,y)) \leqq n^{-1/2}(\log n)^4, \tag{1.5}$$

$$\begin{aligned} &F(\infty,x)+(1-F(\infty,-x))+(1-F(x,-\infty))+F(-x,-\infty) \\ &\qquad \leqq 2\exp\left(-\frac{x^2}{20}\right) \quad \text{if } x \geqq 200, \end{aligned} \tag{1.6}$$

$$1-\mathcal{R}(x)+\mathcal{R}(-x) \leqq 2\exp\left(-\frac{x^2}{20}\right) \quad \text{if } x \geqq 200. \tag{1.7}$$

$F(\cdot,\cdot)$ is a solution of the integral equation

$$F(u,v)=\int_0^1\int_{-\infty}^{+\infty}(F(\alpha^{-1/2}(u-y),\alpha^{-1/2}(v-y)))^2\varphi_\alpha(y)dyd\alpha. \tag{1.8}$$

$F(\cdot,\cdot)$ is the only distribution function which satisfies (1.6) *and* (1.8). *Further we have*

$$\mathcal{R}(z)=\int_{A_z}d(F(x,-\infty)-F(x,y)),$$

where

$$A_z=\{(u,v)\in\mathbb{R}^2,\ -\infty<v<u<\infty,\ 0<u-v<z\}.$$

2. A simplified model

Let

$$\{W_{ni}(t),\ t\geq 0,\ i=1,2,\dots,2^{n-1},\ n=1,2,\dots\}$$

be an array of independent Wiener processes. Let

$$\{U(n,i),\ i=1,2,\dots,2^{n-1},\ n=1,2,\dots\}$$

be an array of independent, uniform-[0,1] r.v.'s. We assume that the arrays $\{W_{ni}(t)\}$ and $\{U(n,i)\}$ are independent. Let

$$\begin{aligned}V(1,1)&=U(1,1),\\ V(2,1)&=U(1,1)U(2,1)=V(1,1)U(2,1),\\ V(2,2)&=U(1,1)U(2,2)=V(1,1)U(2,2),\\ V(n,2k-1)&=V(n-1,k)U(n,2k-1),\\ V(n,2k)&=V(n-1,k)U(n,2k),\end{aligned}$$

$(k=1,2,\dots,2^{n-2},\ n=2,3,\dots)$.

Model 2. Let

$$\begin{aligned}X_{11}^{(2)}&=W_{11}(1-V(1,1)),\\ X_{21}^{(2)}&=X_{11}^{(2)}+W_{21}(V(1,1)-V(2,1)),\\ X_{22}^{(2)}&=X_{11}^{(2)}+W_{22}(V(1,1)-V(2,2)),\\ X_{n,2k-1}^{(2)}&=X_{n-1,k}^{(2)}+W_{n,2k-1}(V(n-1,k)-V(n,2k-1)),\\ X_{n,2k}^{(2)}&=X_{n-1,k}^{(2)}+W_{n,2k}(V(n-1,k)-V(n,2k)),\end{aligned}$$

$(k=1,2,\dots,2^{n-2},\ n=2,3,\dots)$.

Introduce the following notations:

(a) $M_n^+(2) = \max(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^n-1}^{(2)})$,

(b) $\mathbf{P}\{M_n^+(2) < x\} = \mathcal{M}_2^+(n,x)$,

(c) $M_n^-(2) = \min(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^n-1}^{(2)})$,

(d) $\mathcal{M}_2^-(n,x) = \mathbf{P}\{M_n^-(2) < x\}$,

(e) $R_m(2) = M_n^+(2) - M_n^-(2)$,

(f) $\mathcal{R}_2(n,x) = \mathbf{P}\{R_n(2) < x\}$,

(g) $F_n^{(2)}(x,y) = \mathbf{P}\{M_n^+(2) < x,\ M_n^-(2) > y\}$.

In [2] we have proved:

THEOREM C. *Let*

$$M^+(2) = \lim_{n\to\infty} M_n^+(2) \quad a.s.$$

and

$$\mathcal{M}_2^+(x) = \mathbf{P}\{M^+(2) < x\} = \lim_{n\to\infty} \mathcal{M}_2^+(n,x),$$

where the second limit is in weak sense. Then we have

(i) $\mathbf{P}\{|M_n^+(2) - M^+(2)| \geqq \exp(-10^{-3}n)\} \leqq \exp(-0.2n)$,

(ii) $\mathcal{M}_2^+(x - \exp(-10^{-3}n))$
$$\leqq \mathcal{M}_2^+(n,x) \leqq \mathcal{M}_2^+(x + \exp(-10^{-3}n)) + \exp(-0.2n),$$

(iii) $1 - \mathcal{M}_2^+(x) + \mathcal{M}_2^+(-x) \leqq \exp\left(-\frac{x^2}{20}\right)$ *if* $x \geqq 200$,

(iv) $\mathcal{M}_2^+(x)$ *is a solution of the integral equation* (1.3),

(v) $\mathcal{M}_2^+(x)$ *is the only distribution function which satisfies* (iii) *and* (1.3).

Theorem C clearly implies

THEOREM 2. *The following limits exist:*

$$M_2^- = \lim_{n\to\infty} M_n^-(2) \quad a.s.,$$

$$R_2 = \lim_{n\to\infty} R_n(2) \quad a.s.,$$

$$\mathcal{M}_2^-(x) = \mathbf{P}\{M_2^- < x\} = \lim_{n\to\infty} \mathcal{M}_2^-(n,x),$$

$$\mathcal{R}_2(x) = \mathbf{P}\{R_2 < x\} = \lim_{n\to\infty} \mathcal{R}_2(n,x),$$

$$F_2(x,y) = \mathbf{P}\{M_2^+ < x,\ M_2^- > y\} = \lim_{n\to\infty} F_n^{(2)}(x,y),$$

where the last three limits are in weak sense.

Further we have

(i) $\mathbf{P}\{|M_n^-(2) - M_2^-| \geqq \exp(-10^{-3}n)\} \leqq \exp(-0.2n)$,

(ii) $\mathbf{P}\{|R_n(2) - R_2| \geqq 2\exp(-10^{-3}n)\} \leqq 2\exp(-0.2n)$,

(iii) $\mathcal{M}_2^-(x-\exp(-10^{-3}n))$
$$\leqq \mathcal{M}_2^-(n,x) \leqq \mathcal{M}_2^-(x+\exp(-10^{-3}n)) + \exp(-0.2n),$$

(iv) $1-\mathcal{M}_2^-(x)+\mathcal{M}_2^-(-x) \leqq \exp\left(-\frac{x^2}{20}\right)$ *if* $x \geqq 200$,

(v) $F_2(\infty,x)+(1-F_2(\infty,-x))+(1-F_2(x,-\infty))+F_2(-x,-\infty)$
$$\leqq 2\exp\left(-\frac{x^2}{20}\right) \text{ if } x \geqq 200,$$

(vi) $1-\mathcal{R}_2(x) \leqq \exp\left(-\frac{x^2}{80}\right)$ *if* $x \geqq 400$.

THEOREM 3. *$F_2(\cdot,\cdot)$ is a solution of the integral equation*

$$F_2(u,v) = \int_0^1 \int_{-\infty}^{+\infty} (F_2(\alpha^{-1/2}(u-y), \alpha^{-1/2}(v-y)))^2 \varphi_\alpha(y)\,dy\,d\alpha, \tag{2.1}$$

where $\varphi_\alpha(y)$ is defined in Theorem A.

PROOF. Observe that

$$M_n^+(2) = \max(\max(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^{n-2}}^{(2)}),\ \max(X_{n,2^{n-2}+1}^{(2)}, \dots, X_{n,2^{n-1}}^{(2)})),$$
$$M_n^-(2) = \min(\min(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^{n-2}}^{(2)}),\ \min(X_{n,2^{n-1}+1}^{(2)}, \dots, X_{n,2^{n-1}}^{(2)})),$$

$$\max(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^{n-2}}^{(2)}) \overset{\mathcal{D}}{=} (V(1,1))^{1/2} M_{n-1}^+(2) + W_{11}(1-V(1,1))$$

and

$$\min(X_{n1}^{(2)}, X_{n2}^{(2)}, \dots, X_{n,2^{n-2}}^{(2)}) \overset{\mathcal{D}}{=} (V(1,1))^{1/2} M_{n-1}^-(2) + W_{11}(1-V(1,1)).$$

Hence

$$\mathbf{P}\{M_n^+(2) < u,\ M_n^-(2) > v \mid V(1,1)=\alpha,\ W_{11}(1-\alpha)=y\} =$$
$$= (\mathbf{P}\{M_{n-1}^+(2) < \alpha^{-1/2}(u-y),\ M_{n-1}^- > \alpha^{-1/2}(v-y)\})^2.$$

Consequently,

$$\mathbf{P}\{M_2^+ < y,\ M_2^- > v \mid V(1,1)=\alpha,\ W_{11}(1-\alpha)=y\} =$$
$$= (\mathbf{P}\{M_2^+ < \alpha^{-1/2}(u-y),\ M_2^- > \alpha^{-1/2}(v-y)\})^2$$

and

$$F_2(u,v) = \mathbf{P}\{M_2^+ < u,\ M_2^- > v\}$$
$$= \mathbf{E}\{\mathbf{P}\{M_2^+ < u,\ M_2^- > v \mid V(1,1)=\alpha,\ W_{11}(1-\alpha)=y\}\} =$$
$$= \int_0^1 \int_{-\infty}^{+\infty} (\mathbf{P}\{M_2^+ < \alpha^{-1/2}(u-y),\ M_2^- > \alpha^{-1/2}(v-y)\})^2 \varphi_\alpha(y)\,dy\,d\alpha.$$

Hence we have (2.1).

THEOREM 4. *$F_2(\cdot,\cdot)$ is the only distribution function which satisfies* (v) *of Theorem 2 and the integral equation* (1.8).

PROOF. Let

$$\{Z_{ni}=(Z_{ni}(1),Z_{ni}(2)),\ i=1,2,\dots,2^{n-1},\ n=1,2,\dots\}$$

be an array of i.i.d. random vectors with

$$\mathbf{P}\{Z_{ni}(1)<x,\ Z_{ni}(2)>y\}=F(x,y),$$

where the distribution function $F(\cdot,\cdot)$ is a solution of the integral equation (1.8) satisfying (v) of Theorem 2.

The existence of such a distribution function follows from Theorems 2 and 3. Assume also that the arrays $\{Z_{nk}\}$ and $\{W_{nk}(\cdot)\}$ are independent. Let

$$\begin{aligned}
Y_{11}&=Z_{11},\\
Y_{21}&=\overline{X}_{11}^{(2)}+(V(1,1))^{1/2}Z_{11},\\
Y_{22}&=\overline{X}_{11}^{(2)}+(V(1,1))^{1/2}Z_{12},\\
Y_{n,2k-1}&=\overline{X}_{n-1,k}^{(2)}+(V(n-1,k))^{1/2}Z_{n,2k-1},\\
Y_{n,2k}&=\overline{X}_{n-1,k}^{(2)}+(V(n-1,k))^{1/2}Z_{n,2k},
\end{aligned}$$

where

$$\begin{aligned}
\overline{X}_{ni}^{(2)}&=(X_{ni}^{(2)},X_{ni}^{(2)}),\\
Y_{ni}&=(Y_{ni}(1),Y_{ni}(2)).
\end{aligned}$$

Let

$$\begin{aligned}
\mu_n&=\max(Y_{n1}(1),Y_{n2}(1),\dots,Y_{n,2^{n-1}}(1)),\\
\nu_n&=\min(Y_{n1}(2),Y_{n2}(2),\dots,Y_{n,2^{n-1}}(2)),\\
h(t)&=(V(1,1))^{-1/2}(t-W_{11}(1-V(1,1))).
\end{aligned}$$

Observe that

$$\begin{aligned}
\mu_2&=(V(1,1))^{1/2}\max(Z_{11}(1),Z_{12}(1))+W_{11}(1-V(1,1)),\\
\nu_2&=(V(1,1))^{1/2}\min(Z_{11}(2),Z_{12}(2))+W_{11}(1-V(1,1)).
\end{aligned}$$

Hence

$$\begin{aligned}
&\mathbf{P}\{\mu_2<u,\ \nu_2>v\}\\
&=\mathbf{P}\{\max(Z_{11}(1),Z_{12}(1))<h(u),\ \min(Z_{11}(2),Z_{12}(2))>h(v)\}\\
&=\int_0^1\int_{-\infty}^{+\infty}(F_2(\alpha^{-1/2}(u-y),\alpha^{-1/2}(v-y)))^2\varphi_\alpha(y)\,dy\,d\alpha=F_2(u,v).
\end{aligned}\tag{2.2}$$

By (1.5) we have

$$\begin{aligned}&\mathbf{P}\{\max(Y_{31}(1),Y_{32}(1))<u,\ \min(Y_{31}(2),Y_{32}(2))>v\mid\xi\}\\&=\mathbf{P}\{\max(Y_{33}(1),Y_{34}(1))<u,\ \min(Y_{33}(2),Y_{34}(2))>v\mid\xi\}\\&=F(h(u),h(v)),\end{aligned}$$

where

$$\xi=(V(1,1),W_{11}(1-V(1,1))).$$

Since given $V(1,1)$ and $W_{11}(1-V(1,1))$ the random vectors

$$\max(Y_{31}(1),Y_{32}(1)),\ \min(Y_{31}(2),Y_{32}(2))$$

and

$$\max(Y_{33}(1),Y_{34}(1)),\ \min(Y_{33}(2),Y_{34}(2))$$

are independent, we have

$$\mathbf{P}\{\mu_3<u,\ \nu_3>v\mid V(1,1),W_{11}(1-V(1,1))\}=(F_2(h(u),h(v)))^2.$$

Hence

$$\mathbf{P}\{\mu_3<u,\ \nu_3>v\}$$

$$=\int_0^1\int_{-\infty}^{+\infty}(F_2(\alpha^{-1/2}(u-y),\alpha^{-1/2}(v-y)))^2\varphi_\alpha(y)dyd\alpha=F_2(u,v).$$

Similarly we have

$$\mathbf{P}\{\mu_n<u,\ \nu_n>v\}=F_2(u,v).$$

Conditions (v) of Theorem 2 imply that

$$\lim_{n\to\infty}|\mu_n-M_n^+(2)|=0\quad\text{a.s.},$$

$$\lim_{n\to\infty}|\nu_n-M_n^-(2)|=0\quad\text{a.s.}.$$

Hence the limit distribution of μ_n,ν_n and $M_n^+(2),M_n^-(2)$ are equal to each other and we have Theorem 4.

THEOREM 5. *For any $z>0$ we have*

$$\mathcal{R}_2(z)=\int_{A_z}d(F_2(x,-\infty)-F_2(x,y)),$$

where

$$A_z=\{(u,v)\in\mathbb{R}^2,\ -\infty<v<u<\infty,\ 0<u-v<z\}$$

and $F_2(\cdot,\cdot)$ is defined in Theorem 2, determined in Theorems 3 and 4.

PROOF. Since

$$\mathbf{P}\{M_2^+<u,\ M_2^-<v\}=F_2(u,-\infty)-F_2(u,v),$$

Theorem 3 implies Theorem 5.

3. Proof of Theorem 1

In [2] we evaluated the limit distribution of the most right particle in case of Model 2 and we proved a kind of invariance principle saying that the limit properties of Model 2 were essentially the same as those of Model 1. Now we evaluated the limit distribution of the range in case of Model 2 and we show that Model 1 inherits the results of Model 2.

Consider Model 1. Then for any $0 \leq k < n$ let $Q(k,n)$ be the number of those particles which are living at time k and which have at least one offspring living at time n. Clearly

$$B(k) \geqq Q(k,n), \qquad B(n) \geqq Q(k,n),$$

$$\{Q(k,n)=0\} = \{B(n)=0\} \qquad (0 \leqq k \leqq n).$$

$Q(k,n)$ is a nondecreasing function of k $(0 \leqq k \leqq n)$ and $Q(0,n)=1$ provided that $B(n) \geqq 1$. Hence on the set $\{B(n)>0\}$ we can define a r.v. $\nu_{11} = \nu_{11}(n)$ as follows:

$$\nu_{11} = \inf\{k:\ 0<k \leqq n,\ Q(k,n)=2\}.$$

At time ν_{11} we have two particles and both of them have at least one offspring living at time n. ν_{11} will be called the first branching time of the process. These two particles can be considered as the roots of two independent branching processes living at least till time n (starting from ν_{11}). Let $\nu_{21} = \nu_{21}(n)$ resp. $\nu_{22} = \nu_{22}(n)$ be the first branching times of the branching processes starting from ν_{11}. Clearly $\nu_{11} < \nu_{2i} \leqq n$ $(i=1,2)$. In case $\nu_{11} = n$ define $\nu_{2i} = n$. Note that in case $\nu_{11} = n-1$ we have also $\nu_{2i} = n$.

We can say again that at times ν_{21} (resp. ν_{22}) we have two (resp. two) particles and they can be considered as the roots of four independent branching processes living at least till time n. Let $\nu_{31} = \nu_{31}(n)$ (resp. $\nu_{32} = \nu_{32}(n)$) be the first branching times of the branching processes starting from ν_{21}. Similarly let $\nu_{33} = \nu_{33}(n)$ (resp. $\nu_{34} = \nu_{34}(n)$) be the first branching times of the branching processes starting from ν_{22}. Note that in case $\nu_{21} \geqq n-1$ we have $\nu_{31} = \nu_{32} = n$ and in case $\nu_{22} \geqq n-1$ we have $\nu_{33} = \nu_{34} = n$.

In general at time ν_{kj} $(j=1,2,\ldots,2^{k-1})$ we have two particles and they can be considered as the roots of two independent branching processes living at least till time n (starting from ν_{kj}). Let $\nu_{k+1,2j-1} = \nu_{k+1,2j-1}(n)$ resp. $\nu_{k+1,2j} = \nu_{k+1,2j}(n)$ be the first branching times of the branching processes starting at ν_{kj}. Note that $\nu_{k+1,2j-1} = \nu_{k+1,2j} = n$ if $\nu_{kj} \geqq n-1$.

Now we build up our

MODEL 3.

Let

$$\{W_{ni}(t),\ t \geqq 0,\ i=1,2,\ldots,2^{n-1},\ n=1,2,\ldots\}$$

be an array of independent Wiener processes which is independent from the array

$$\{\nu_{kj}(n),\ j=1,2,\ldots,2^{k-1},\ k=1,2,\ldots,n\}.$$

Let

$$X_{11}^{(3)} = W_{11}(\nu_{11}),$$
$$X_{21}^{(3)} = X_{11}^{(3)} + W_{21}(\nu_{21} - \nu_{11}),$$
$$X_{22}^{(3)} = X_{11}^{(3)} + W_{22}(\nu_{22} - \nu_{11}),$$
$$X_{m,2k-1}^{(3)} = X_{m-1,k}^{(3)} + W_{m,2k-1}(\nu_{m,2k-1} - \nu_{m-1,k}),$$
$$X_{m,2k}^{(3)} = X_{m-1,k}^{(3)} + W_{m,2k}(\nu_{m,2k} - \nu_{m-1,k}),$$

$(k = 1, 2, \dots, 2^{m-2},\ m = 1, 2, \dots, n)$.

Note that the sequence

$$\{X_{n,k}^{(3)},\ k = 1, 2, \dots, 2^{n-1}\}$$

is equal to the sequence

$$\{X_{nk},\ k = 1, 2, \dots, B(n)\}$$

except that the elements of the second sequence might occur many times in the first sequence. Hence

$$\begin{aligned} M_n^+(3) &:= \max\{X_{n1}^{(3)}, X_{n2}^{(3)}, \dots, X_{n,2^{n-1}}^{(3)}\} = M_n^+ \\ M_n^- &:= \min\{X_{n1}^{(3)}, X_{n2}^{(3)}, \dots, X_{n,2^{n-1}}^{(3)}\} = M_n^- \\ R_n(3) &:= M_n^+(3) - M_n^-(3) = R_n. \end{aligned} \tag{3.1}$$

Let

$$\overline{X}_{11}^{(2)} = W_{11}(n(1 - V(1,1))),$$
$$\overline{X}_{21}^{(2)} = \overline{X}_{11}^{(2)} + W_{21}(n(V(1,1) - V(2,1))),$$
$$\overline{X}_{22}^{(2)} = \overline{X}_{11}^{(2)} + W_{22}(n(V(1,1) - V(2,2))),$$
$$\overline{X}_{m,2k-1}^{(2)} = \overline{X}_{m-1,k}^{(2)} + W_{m,2k-1}(n(V(n-1,k) - V(n,2k-1))),$$
$$\overline{X}_{m,2k}^{(2)} = \overline{X}_{m-1,k}^{(2)} + W_{m,2k}(n(V(n-1,k) - V(n,2k))),$$

$(k = 1, 2, \dots, 2^{m-2},\ m = 1, 2, \dots, n)$.

Note that

$$\{n^{1/2} X_{mk}^{(2)},\ k = 1, 2, \dots, 2^{m-1},\ m = 1, 2, \dots, n\} \stackrel{\mathcal{D}}{=}$$
$$\stackrel{\mathcal{D}}{=} \{\overline{X}_{mk}^{(2)},\ k = 1, 2, \dots, 2^{m-1},\ m = 1, 2, \dots, n\}.$$

Now we recall

THEOREM D. *For any $n=1,2,\dots$ there exist:*

(i) *a probability space $\{\Omega_n, S_n, \mathbf{P}_n\}$,*

(ii) *a branching process* (cf. Section 3) *on Ω_n with $B(n)>0$,*

(iii) *an array of independent, uniform-*[0,1] *r.v.'s*

$$\{U(k,i),\ \ i=1,2,\dots,2^{k-1};\ \ k=1,2,\dots,n\},$$

(iv) *an array of independent Wiener processes $\{W_{kj}(\cdot),\ k=1,2,\dots;\ j=1,2,\dots,2^{k-1}\}$ which is independent from $\{\nu_{kj}\}$ and $\{U_{kj}\}$*

such that

$$\mathbf{P}\{\max_{1\leqq k\leqq \kappa}\max_{1\leqq j\leqq 2^{k-1}}\max_{l=2j-1,2j}|A(k,j,l)|\geqq C(\log n)^2\}\leqq \kappa 2^{\kappa-1}\exp\left(-\frac{(\log n)^2}{2}\right), \tag{3.2}$$

where

$$A(k,j,l)=W_{k+1,l}(\nu_{k+1,l}-\nu_{k,j})-W_{k+1,l}(nV(k,j)(1-U(k+1,l)))$$

and

$$\mathbf{P}\{|W_{11}(\nu_{11})-W_{11}(n(1-V(1,1)))|\geqq C(\log n)^{3/2}\}\leqq \exp\left(-\frac{\log^2 n}{2}\right). \tag{3.3}$$

Further

$$\mathbf{P}\left\{\max_{1\leqq l\leqq 2^{n-1}} V(\kappa,l)\geqq\left(\frac{99}{100}\right)^{\kappa}\right\}\leqq\left(\frac{6}{10}\right)^{\kappa} \tag{3.4}$$

and

$$\mathbf{P}\{n=\nu_{\kappa l},\ \ l=1,2,\dots,2^{\kappa-1}\}\geqq 1-\exp\left(-\left(C\log\frac{10}{6}\right)\log n\right), \tag{3.5}$$

where

$$\kappa=[C\log n]$$

and

$$C>\frac{2}{\log\frac{100}{99}}.$$

Consequently,

$$\mathbf{P}\{\max_{1\leqq k\leqq \kappa}\max_{1\leqq j\leqq 2^{k-1}}|X_{kj}^{(3)}-n^{1/2}X_{kj}^{(2)}|\geqq C(\log n)^2\}\leqq \kappa 2^{\kappa-1}\exp\left(-\frac{(\log n)^2}{2}\right). \tag{3.6}$$

Note that by (3.4) we have

$$\mathbf{P}\{X_{nj}^{(3)} = X_{\kappa j}^{(3)},\ j = 1, 2, \ldots, 2^{\kappa-1}\} \geqq 1 - \exp\left(-\left(C \log \frac{10}{6}\right) \log n\right) \tag{3.7}$$

taking into account only the different elements of the sequences $\{X_{nj}^{(3)}\}$ and $\{X_{\kappa j}^{(3)}\}$. Hence

$$\mathbf{P}\{R_\kappa(3) \neq R_n(3)\} \leqq \exp\left(-\left(C \log \frac{10}{6}\right) \log n\right) \tag{3.8}$$

and by (3.6)

$$\begin{aligned} &\mathbf{P}\{|R_\kappa(3) - n^{1/2} R_\kappa(2)| \geqq C(\log n)^2\} \\ &\leqq \kappa 2^{\kappa-1} \exp\left(-\frac{(\log n)^2}{2}\right). \end{aligned} \tag{3.9}$$

(3.1), (3.8) and (3.9) combined imply

$$\begin{aligned} &\mathbf{P}\{|R_n - n^{1/2} R_\kappa(2)| \geqq C(\log n)^2\} \\ &\leqq \exp\left(-\left(C \log \frac{10}{6}\right) \log n\right) + \kappa 2^{\kappa-1} \exp\left(-\frac{(\log n)^2}{2}\right). \end{aligned} \tag{3.10}$$

Remember that (ii) of Theorem 2 claims that

$$\begin{aligned} &\mathbf{P}\{|n^{1/2} R_\kappa(2) - n^{1/2} R_2| \geqq 2n^{1/2} \exp(-10^{-3}\kappa)\} \\ &\leqq 2 \exp(-0.2\kappa). \end{aligned} \tag{3.11}$$

Hence (3.10) and (3.11) imply that for any $K > 0$ there exists a $C = C(K) > 0$ such that

$$\mathbf{P}\{|n^{-1/2} R_n - R_2| \geqq C n^{-1/2} (\log n)^2\} \leqq n^{-K}. \tag{3.12}$$

Consequently, we have Theorem 1 by Theorems 2 and 3.

REFERENCES

[1] RÉVÉSZ, P., *Random walks of infinitely many particles*, World Scientific, Singapore, 1994.

[2] RÉVÉSZ, P., The maximum of a critical branching Wiener process, *J. Theoret. Probability* (to appear).

(Received November 20, 1997)

INSTITUT FÜR STATISTIK UND
WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
WIEDNER HAUPTSTRASSE 8–10/107
A–1040 WIEN
AUSTRIA

revesz@ci.tuwien.ac.at

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 391–400

ASYMPTOTICS OF MULTIVARIATE EXTREMES WITH RANDOM SAMPLE SIZE

D. S. SILVESTROV and J. L. TEUGELS

Dedicated to the memory of Alfréd Rényi

Abstract

We investigate the asymptotics of multivariate extremes with random sample size under general dependence-independence conditions for samples and random sample size indexes.

1. Introduction

Asymptotics of extremes with random sample size indexes have been thoroughly studied for two types of models with asymptotically independent sample and sample size indexes. In papers by Berman [3], Thomas [21], Galambos [7], [9], [10], Gnedenko, B. and Gnedenko, D. [11], Beirlant and Teugels [2] and Korolev [12] the model where the sample and the sample size indexes are independent has been investigated. The papers by Berman [3], Barndorff-Nielsen [1], Mogyoródi [14], Sen [17], Galambos [6], [8], [9] deal with the model where sample size indexes depend on the sample but converge in probability. This type of convergence is stronger than weak convergence of random sample size indexes. It implies the asymptotic independence of random sample size indexes and the corresponding extremes with the non-random sample size due to a well-known result by Rényi [16] concerning mixing sequences of random events.

In a recent paper, Silvestrov and Teugels [20] derived limit theorems for extremes with random sample size for the model where the sample and the sample size are dependent in an arbitrary way. The general limit theorems for superpositions of random processes developed in Silvestrov [19] have been used as a basic tool. In the present paper we generalize some of the main results of Silvestrov and Teugels [20] from the univariate to the multivariate model of extremes with random sample size. We also show how the results, related to the model where sample and sample size indexes are asymptotically independent, can be obtained from these general theorems. Finally we present general triangular array versions of the results related to the case where the sample size indexes converge in probability. As was already mentioned above, the remarkable result by Rényi concerning mixing sequences of random events plays here an essential role.

1991 *Mathematics Subject Classification.* Primary 26A12; Secondary 60F05.
Key words and phrases. Extreme, random sample size, weak convergence.

2. Weak convergence of multivariate extremal processes with random sample size indexes

For every $\varepsilon > 0$, let $\{\bar{\xi}_{\varepsilon,n} = (\xi_{\varepsilon,n,i},\ i = 1,\dots,m),\ n = 1,2,\dots\}$ be a sequence of real-valued i.i.d. random vectors. Further we need random vectors $\bar{\mu}_\varepsilon = (\mu_{\varepsilon,i},\ i = 1,\dots,m)$ with non-negative components, and non-random functions $n_\varepsilon > 0$ for which $n_\varepsilon \to \infty$ as $\varepsilon \to 0$.

If we are interested in multivariate extremal processes with non-random sample size indexes then we deal with the vector process $\bar{\xi}_\varepsilon(t) = (\xi_{\varepsilon,i}(t),\ i = 1,\dots,m)$, $t \geqq 0$ where

$$\xi_{\varepsilon,i}(t) = \max_{k \leq tn_\varepsilon} \xi_{\varepsilon,k,i}, \qquad t > 0. \tag{1}$$

In formula (1) and below, we assign the value zero to a maximum over an empty set.

Our interest lies in the relevant analogues of these processes when the sample size indexes are random as well. So, define $\bar{\zeta}_\varepsilon(t) = (\zeta_{\varepsilon,i}(t),\ i = 1,\dots,m)$, $t \geqq 0$ where

$$\zeta_{\varepsilon,i}(t) = \max_{k \leq t\mu_{\varepsilon,i}} \xi_{\varepsilon,k,i}, \qquad t > 0. \tag{2}$$

Let us denote by $\nu_{\varepsilon,i} = \mu_{\varepsilon,i}/n_\varepsilon$ normalized random sample size indexes and $\bar{\nu}_\varepsilon = (\nu_{\varepsilon,i},\ i = 1,\dots,m)$. Then the process $\{\bar{\zeta}_\varepsilon(t), t > 0\}$ can be represented in the form of the vector composition of the two processes $\{\bar{\xi}_\varepsilon(t), t > 0\}$ and $\{\bar{\nu}_\varepsilon(t) = t\bar{\nu}_\varepsilon, t > 0\}$, i.e. $\bar{\zeta}_\varepsilon(t) = (\zeta_{\varepsilon,i}(t),\ i = 1,\dots,m)$, $t \geqq 0$ where

$$\zeta_{\varepsilon,i}(t) = \xi_{\varepsilon,i}(t\nu_{\varepsilon,i}), \qquad t > 0. \tag{3}$$

This representation points to the use of limit theorems for compositions of random processes as a tool in obtaining limit theorems for extremal processes with random sample size.

Let G be the class of non-increasing continuous functions $g(\bar{u})$ acting from R_m into $[0,\infty]$ such that $e^{-g(\bar{u})}$ is an m-dimensional distribution function. (If $g(\bar{u}) = \infty$ we understand continuity in such a point in the sense that $g(\bar{v}) \to \infty$ as $\bar{v} \geqq \bar{u}$, $\bar{v} \to \bar{u}$.)

Define $G_\varepsilon(\bar{u}) = \mathbf{P}\{\bar{\xi}_{\varepsilon,1} \leqq \bar{u}\}$. The following condition is standard in papers dealing with limit theorems for extremes:

A_0: $n_\varepsilon(1 - G_\varepsilon(\bar{u})) \to g(\bar{u})$ as $\varepsilon \to 0$, $\bar{u} \in R_m$.

Denote by D_m the space of step functions on $(0,\infty)$, taking values in R_m, continuous from the right and with a finite number of jumps in every finite interval (a,b), $0 < a < b < \infty$; these jumps have to be non-negative and strictly positive in at least one component.

It is known (see for example Resnick [15], Leadbetter, Lindgren and Rootzén [13]) that Condition A_0 is equivalent to the condition

A: $\{\bar{\xi}_\varepsilon(t), t>0\} \Rightarrow \{\bar{\xi}_0(t), t>0\}$ as $\varepsilon \to 0$.

The limiting process $\{\bar{\xi}_0(t), t>0\}$ in A is called an *extremal process*. It is a stochastically continuous Markov jump process whose trajectories belong with probability 1 to the space D_m; its transition probabilities are given by

$$\mathbf{P}\{\bar{\xi}_0(s+t) \leqq \bar{u} \mid \bar{\xi}_0(s) = \bar{v}\} = \chi(\bar{v} \leqq \bar{u}) e^{-tg(\bar{u})}, \tag{4}$$

where $\chi(A)$ is the indicator of the event A.

We also have to assume a condition concerning the asymptotic behaviour of the random sample size which is consistent with A. A minimal such condition is

B: $\bar{\nu}_\varepsilon = \bar{\mu}_\varepsilon / n_\varepsilon \Rightarrow \bar{\nu}_0$ as $\varepsilon \to 0$, where $\bar{\nu}_0 = (\nu_{0,i},\ i=1,\dots,m)$ is a random vector with a.s. positive components.

It can be expected that Conditions A and B are sufficient to provide the weak convergence of extremal processes when the extremal process $\{\bar{\xi}_\varepsilon(t), t>0\}$ and the random sample size index $\bar{\nu}_\varepsilon$ are independent. However, in the case of dependence, Conditions A and B need to be replaced by a stronger condition in terms of the joint distribution of $\{\bar{\xi}_\varepsilon(t), t>0\}$ and $\bar{\nu}_\varepsilon$, i.e.

C: $\{(\bar{\xi}_\varepsilon(t), \bar{\nu}_\varepsilon), t>0\} \Rightarrow \{(\bar{\xi}_0(t), \bar{\nu}_0), t>0\}$ as $\varepsilon \to 0$ where $\{\bar{\xi}_0(t), t>0\}$ and $\bar{\nu}_0$ were defined in Conditions A and B, respectively.

Let $w>0$. Denote by $w < \tau_{1,w} < \tau_{2,w} < \dots$ the successive moments of jumps of the process $\bar{\xi}_0(t)$ in the interval $[w,\infty)$. For convenience we put $\tau_{0,w} = \tau_{-1,w} = w$. Denote by S the set of points $t>0$ for which $\mathbf{P}\{\tau_{k,w_n} = t\nu_{0,i}\} = 0$ for all $i=1,\dots,m$ and $k,n=1,2\dots$ where $w_n = n^{-1}$. The set $\bar{S}$ contains not more than a countable number of points since it coincides with the set of atoms for the distributions of random variables $\tau_{k,w_n}/\nu_{0,i}$, $i=1,\dots,m$; $k,n=1,2,\dots$. Therefore the set S is $(0,\infty)$ up to at most a countable set of points.

The set S coincides with $(0,\infty)$ if the random variables $\tau_{k,w_n}/\nu_{0,i}$, $i=1,\dots,m$; $k,n=1,2,\dots$ have continuous distributions. Since the random variables τ_{k,w_n} have continuous distribution functions, $\tau_{k,w_n}/\nu_{0,i}$ also have continuous distribution functions and so $S=(0,\infty)$ if the process $\{\bar{\xi}_0(t),\ t>0\}$ and the random variable $\bar{\nu}_0$ are independent or at any rate, the random variables τ_{k,w_n} and $\nu_{0,i}$ are independent for every $i=1,\dots,m$ and $k,n=1,2,\dots$. In the latter case the process $\{\bar{\xi}_0(t), t>0\}$ and the random vector $\bar{\nu}_0$ can be dependent.

The main result of this paper is the following theorem which generalizes Theorem 1 in [20] to the case of multivariate extremes with random sample size indexes.

THEOREM 1. *Let Condition* C *hold. Then*

$$\begin{aligned}&\{\bar{\zeta}_\varepsilon(t) = (\xi_{\varepsilon,i}(t\nu_{\varepsilon,i}),\ i=1,\dots,m),\ t\in S\} \Rightarrow \\ &\{\bar{\zeta}_0(t) = (\xi_{0,i}(t\nu_{0,i}),\ i=1,\dots,m),\ t\in S\} \quad \text{as } \varepsilon \to 0.\end{aligned} \tag{5}$$

PROOF. Theorem 1 can be obtained by using general results concerning weak convergence of randomly stopped processes as given in [18], [19]. However, thanks to the monotonicity of extremal processes we are able to give a simplified version of the proof by following the procedure in [20] for the univariate case.

Note first of all that the definition of the extremal process, as given in (1) causes some side effect at zero. The process $\bar{\zeta}_\varepsilon(t)$ has step trajectories, is continuous from the right and has possibly jumps only at points k/n_ε, $k \geqq 1$. All jumps with $k \geqq 2$ have non-negative components and at least one of them is positive; hence all components of the resulting process are monotonically non-decreasing on the interval $[1/n_\varepsilon, \infty)$. However, on the interval $(0, 1/n_\varepsilon)$ the process takes the value zero and the first jump can possess negative components if the random vector $\bar{\xi}_{\varepsilon,1}$ has negative components. To avoid this situation we consider a slight modification of the extremal processes as defined by (1) and (2). We replace the respective processes by

$$\tilde{\xi}_{\varepsilon,i}(t) = \max_{k \leqq (tn_\varepsilon \vee 1)} \xi_{\varepsilon,k,i}, \quad t > 0 \tag{6}$$

and

$$\tilde{\zeta}_{\varepsilon,i}(t) = \max_{k \leqq (t\mu_{\varepsilon,i} \vee 1)} \xi_{\varepsilon,k,i}, \quad t > 0. \tag{7}$$

By the definition of these processes, $\tilde{\zeta}_{\varepsilon,i}(t) - \zeta_{\varepsilon,i}(t) = \xi_{\varepsilon,1,i}\chi(t\nu_{\varepsilon,i} \leqq 1/n_\varepsilon)$ and under Condition B for any $t > 0$

$$\mathbf{P}\left\{\sup_{s \geqq t} |\tilde{\zeta}_{\varepsilon,i}(s) - \zeta_{\varepsilon,i}(s)| > 0\right\} \leqq \mathbf{P}\{t\nu_{\varepsilon,i} < 1/n_\varepsilon\} \to 0 \ \text{ as } \varepsilon \to 0. \tag{8}$$

For this reason both versions of the extremal processes with random sample size indexes will have the same asymptotic behaviour in the sense of weak convergence.

For every $n = 1, 2, \dots,$ let $\dots < z_{-1,n} < z_{0,n} < z_{1,n} < \dots$ be a partition of the interval $(0, \infty)$ such that: (a) $z_{-k,n} \to 0$ as $k \to \infty$; (b) $z_{k,n} \to \infty$ as $k \to \infty$; (c) $h_n = \max_k (z_{k+1,n} - z_{k,n}) \to 0$ as $n \to \infty$.

For every $\varepsilon > 0$ and $i = 1, \dots, m$ we define the approximative extremal processes with non-random sample size

$$\xi^{\pm}_{\varepsilon,n,i}(t) = \tilde{\xi}_{\varepsilon,i}(z_{k+\frac{1\pm 1}{2},n}) \ \text{ for } \ z_{k,n} \leqq t < z_{k+1,n} \qquad -\infty < k < \infty \tag{9}$$

and in the limiting case for every $i = 1, \dots, m$

$$\xi^{\pm}_{0,n,i}(t) = \xi_{0,i}(z_{k+\frac{1\pm 1}{2},n}) \ \text{ for } \ z_{k,n} \leqq t < z_{k+1,n} \qquad -\infty < k < \infty. \tag{10}$$

Similarly, we define for every $\varepsilon \geqq 0$ and $i = 1, \dots, m$ and the corresponding approximative extremal processes with random sample size indexes

$$\zeta^{\pm}_{\varepsilon,n,i}(t) = \xi^{\pm}_{\varepsilon,n,i}(t\nu_{\varepsilon,i}), \quad t > 0. \tag{11}$$

By definition of the approximative processes $\{\zeta^{\pm}_{\varepsilon,n,i}(t), t>0\}$ for all $i = 1,\dots,m$, $n=1,2,\dots$ and $\varepsilon>0$

$$\zeta^{-}_{\varepsilon,n,i}(t) \leqq \bar{\zeta}_{\varepsilon,i}(t) \leqq \zeta^{+}_{\varepsilon,n,i}(t) \quad \text{for } t>0. \tag{12}$$

Similar inequalities are valid for all $i=1,\dots,m$, $n=1,2,\dots$ in the limiting case $\varepsilon=0$

$$\zeta^{-}_{0,n,i}(t) \leqq \zeta_{0,i}(t) \leqq \zeta^{+}_{0,n,i}(t) \quad \text{for } t>0. \tag{13}$$

Also by the definition of the partitions $\{z_{kn}\}$

$$|\zeta_{0,i}(t) - \zeta^{\pm}_{0,n,i}(t)| \leqq \sup_{|s|\leqq h_n} |\xi_{0,i}(t\nu_{0,i}) - \xi_{0,i}(t\nu_{0,i}+s)|, \tag{14}$$

where $\xi_{0,i}(s)=0$ for $s\leqq 0$.

By the definition of S, the random point $t\nu_{0,i}$ is with probability 1 a point of continuity of the random process $\{\xi_{0,i}(t), t>0\}$ for every $i=1,\dots,m$ and $t\in S$. Using this fact and taking into account that $h_n\to 0$ as $n\to\infty$, the relation (14) implies that for $i=1,\dots,m$ and $t\in S$

$$\zeta^{\pm}_{0,n,i}(t) \xrightarrow{P} \zeta_{0,i}(t) \quad \text{as } n\to\infty. \tag{15}$$

From (13) and (15) follows that

$$\begin{gathered}\{(\zeta^{\pm}_{0,n,i}(t),\ i=1,\dots,m),\ t\in S\} \Rightarrow \\ \{(\zeta_{0,i}(t),\ i=1,\dots,m),\ t\in S\} \quad \text{as } n\to\infty.\end{gathered} \tag{16}$$

Let us take arbitrary points $t_1,\dots,t_r\in S$. By definition for every $\varepsilon>0$

$$\begin{aligned}&\mathbf{P}\{\zeta^{\pm}_{\varepsilon,n,i}(t_j)\leqq u_{ij},\ i=1,\dots,m,\ j=1,\dots,r\}\\ &=\sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k_j=-\infty}^{\infty}\mathbf{P}\{\bar{\xi}_{\varepsilon,i}(z_{k_j+\frac{1\pm1}{2},n})\leqq u_{ij},\\ &\quad t_j\nu_{\varepsilon,i}\in[z_{k_j,n}, z_{k_j+1,n}),\ i=1,\dots,m,\ j=1,\dots,r\}.\end{aligned} \tag{17}$$

We can always choose the partitions $z_{k,n}$ in such a way that $z_{k,n}\in S$ and $\mathbf{P}\{t_j\nu_{0,i}=z_{k,n}\}=0$ for all i,j,k and n. In this case using (17) and Condition C we get for all $n=1,2\dots$

$$\begin{gathered}(\zeta^{\pm}_{\varepsilon,n,i}(t_j),\ i=1,\dots,m, j=1,\dots r)\Rightarrow \\ (\zeta^{\pm}_{0,n}(t_i),\ i=1,\dots,m,\ j=1,\dots,r) \quad \text{as } \varepsilon\to 0.\end{gathered} \tag{18}$$

It is always possible to find a set U of points $(u_{ij}, i=1,\dots,m,\ j=1,\dots,r)$ dense in R_{mr}, which are points of continuity for the distribution functions of random vectors $(\zeta_{0,i}(t_j), i=1,\dots,m, j=1,\dots,r)$ and $(\zeta^{\pm}_{0,n,i}(t_j), i=1,\dots,m, j=1,\dots,r)$ for all $n \geqq 1$. Using the inequality (13), and relations (15) and (18) we get for points $(u_{ij}, i=1,\dots,m, j=1,\dots,r) \in U$

$$\begin{aligned}\varliminf_{\varepsilon\to 0} &\mathbf{P}\{\bar{\zeta}_{\varepsilon,i}(t_j) \leqq u_{ij}, i=1,\dots,m, j=1,\dots,r\}\\ &\geqq \lim_{n\to 0}\varliminf_{\varepsilon\to 0}\mathbf{P}\{\zeta^{+}_{\varepsilon,n,i}(t_j) \leqq u_{ij}, i=1,\dots,m, j=1,\dots,r\}\\ &= \lim_{n\to 0}\mathbf{P}\{\zeta^{+}_{0,n,i}(t_j) \leqq u_{ij}, i=1,\dots,m, j=1,\dots,r\}\\ &= \mathbf{P}\{\zeta_{0,i}(t_j) \leqq u_{ij}, i=1,\dots,m, j=1,\dots,r\}\end{aligned} \tag{19}$$

and in an analogous fashion

$$\begin{aligned}\varlimsup_{\varepsilon\to 0}&\mathbf{P}\{\bar{\zeta}_{\varepsilon,i}(t_j) \leqq u_{ij}, i=1,\dots,m, j=1,\dots,r\}\\ &\leqq \mathbf{P}\{\zeta_{0,i}(t_j) \leqq u_{ij}, i=1,\dots m, j=1,\dots,r\}.\end{aligned} \tag{20}$$

Relations (19) and (20) are equivalent to

$$\begin{aligned}&(\bar{\zeta}_{\varepsilon,i}(t_j), i=1,\dots,m, j=1,\dots,r) \Rightarrow\\ &(\zeta_{0,i}(t_j), i=1,\dots,m, j=1,\dots,r) \quad \text{as } \varepsilon\to 0.\end{aligned} \tag{21}$$

Now we can use the relation (8) and come back to the extremal processes $\{\bar{\zeta}_\varepsilon(t), t>0\}$. The relations (21) and (8) imply that

$$\begin{aligned}&(\zeta_{\varepsilon,i}(t_j), i=1,\dots,m, j=1,\dots,r) \Rightarrow\\ &(\zeta_{0,i}(t_j), i=1,\dots,m, j=1,\dots,r) \quad \text{as } \varepsilon\to 0.\end{aligned} \tag{22}$$

Since t_i, $i=1,\dots,m$ are arbitrary points in S the last relation completes the proof. □

3. Consequences of the main theorem

Let us first apply Theorem 1 to the important case where the sample and the sample size indexes are asymptotically independent. In addition to the basic Condition C we assume the following condition:

D: the process $\{\bar{\xi}_0(t), t>0\}$ and the random vector $\bar{\nu}_0$ are independent.

As was mentioned in Section 1 Condition D implies that the set $S=(0,\infty)$. So, as a corollary to Theorem 1, we can formulate the following theorem.

THEOREM 2. *Let Conditions* C *and* D *hold. Then*

$$\begin{aligned}(23)\qquad &\{\bar\zeta_\varepsilon(t)=(\xi_{\varepsilon,i}(t\nu_{\varepsilon,i}),i=1,\dots,m),t>0\}\Rightarrow\\ &\{\bar\zeta_0(t)=(\xi_{0,i}(t\nu_{0,i}),i=1,\dots,m),t>0\}\quad as\ \varepsilon\to 0.\end{aligned}$$

Theorem 2 obviously covers the case where the sample values $\{\bar\xi_{\varepsilon,n}, n=0,1\dots\}$ and the vector sample size index $\bar\mu_\varepsilon$ are independent for every $\varepsilon>0$. In this case Condition D automatically holds and Condition C is equivalent to Conditions A and B.

This reduction of Theorem 1 to the case of independent sample and sample size indexes is a triangular array vector version of the results given in a variety of different forms by Berman [3], Thomas [21], Galambos [7], [8], Gnedenko, B. and Gnedenko, D. [11], and Beirlant and Teugels [2].

An even more interesting application of Theorem 1 deals with the model where the random sample size indexes converge in probability.

It is natural to assume in this case that the random vectors $\{\bar\xi_{\varepsilon,n}, n=1,2,\dots\}$ and $\bar\mu_\varepsilon$ for all $\varepsilon>0$ are defined on the same probability space. We assume also that the independence condition for the random variables $\bar\xi_{\varepsilon,n}$ is satisfied in the following stronger sense.

E: The sets of random vectors $\{\bar\xi_{\varepsilon,n},\varepsilon>0\}$ are mutually independent for $n=1,2,\dots$.

Obviously, Condition E holds for the scale-location model. In this case the random vectors $\bar\xi_{\varepsilon,n}$ are represented in the form $\bar\xi_{\varepsilon,n}=(\xi_{n,i}-a_{\varepsilon,i})/b_{\varepsilon,i}$, $i=1,\dots,m$, where $\bar\xi_n=(\xi_{n,i}, i=1,\dots,m)$, $n=1,2,\dots$ are i.i.d. random vectors, and $a_{\varepsilon,i},b_{\varepsilon,i}$, $i=1,\dots,m$ are some non-random centralization and normalization constants. It also holds for the more general model with random vectors $\bar\xi_{\varepsilon,n}=h_\varepsilon(\bar\xi_n)$, $n=1,2,\dots$ where $h_\varepsilon(\cdot)$ are non-random measurable functions acting from R_m into R_m.

The condition for weak convergence B is replaced by the following condition:

F: $\bar\nu_\varepsilon=\bar\mu_\varepsilon/n_\varepsilon\xrightarrow{P}\bar\nu_0$ as $\varepsilon\to 0$, where $\bar\nu_0$ is a random vector with all components a.s. positive.

Note that the independence of the sample and sample size indexes is not assumed. However, as we see, Conditions A, E and F imply Conditions C and D, i.e. asymptotic independence of the extremal processes with non-random sample size and the random sample size indexes.

The following theorem is a triangular array vector version of the results given in different variants for the case of a scale-location model by Berman [3], Barndorff-Nielsen [1], Mogyoródi [14], Sen [17], Galambos [7], [8], [9] and Eriksson [5].

THEOREM 3. *Let Conditions* A, E *and* F *hold. Then*

$$\begin{aligned}(24)\qquad &\{\bar\zeta_\varepsilon(t)=(\xi_{\varepsilon,i}(t\nu_{\varepsilon,i}),\ i=1,\dots,m),\ t>0\}\Rightarrow\\ &\{\bar\zeta_0(t)=(\xi_{0,i}(t\nu_{0,i}),\ i=1,\dots,m),\ t>0\}\quad as\ \varepsilon\to 0,\end{aligned}$$

where the process $\{\bar{\xi}_0(t), t > 0\}$ *and the random vector* $\bar{\nu}_0$ *are independent.*

PROOF. The proof of this theorem follows from the following statement which generalizes to a triangular array scheme the analogous statements obtained for the scale-location model in the above named papers. The proof of this result, which seems to be of some independent interest, is based on the use of a key result by Rényi [16] concerning mixing sequences of random events.

We are going to prove that Conditions A, E and F imply that

$$\{(\bar{\xi}_\varepsilon(t), \bar{\nu}_\varepsilon), t > 0\} \Rightarrow \{(\bar{\xi}_0(t), \bar{\nu}_0), t > 0\} \quad \text{as } \varepsilon \to 0, \tag{25}$$

where the process $\{\bar{\xi}_0(t), t > 0\}$ and the random vector $\bar{\nu}_0$ are independent.

Let us take some subsequence $\varepsilon_n \to 0$ as $n \to \infty$ and choose some $0 < t_1 < \dots < t_r < \infty$ and $s_{i1} \leqq \dots \leqq s_{ir} < \infty$ for $i = 1, \dots, m$. Define

$$A_n = \{\max_{k \leqq t_j n_{\varepsilon_n}} \xi_{\varepsilon_n,k,i} \leqq s_{ij},\ i = 1, \dots, m,\ j = 1, \dots, r\}$$

and

$$A = \{\xi_{0,i}(t_j) \leqq s_{ij},\ i = 1, \dots, m,\ j = 1, \dots, r\}.$$

We are going to prove first that the sequence of events $\{A_n\}$ is a mixing sequence in the sense of [16], i.e. for any $l \geqq 1$

$$\lim_{n \to \infty} \mathbf{P}(A_n \cap A_l) = \mathbf{P}(A)\mathbf{P}(A_l). \tag{26}$$

The latter result is only non-trivial in the case when $\mathbf{P}(A) > 0$. Obviously, the event A_n can be written in the form $A_n = A_{nl}^+ \cap A_{nl}$ where

$$A_{nl}^- = \{\max_{k \leqq t_r n_{\varepsilon_l}} \xi_{\varepsilon_n,k,i} \leqq s_{i1},\ i = 1, \dots, m\}$$

and

$$A_{nl}^+ = \{\max_{t_r n_{\varepsilon_l} < k \leqq t_j n_{\varepsilon_n}} \xi_{\varepsilon_n,k,i} \leqq s_{ij},\ i = 1, \dots, m, j = 1, \dots, r\}.$$

From Condition A and supposition $\mathbf{P}(A) > 0$ it follows that for chosen $\bar{s}_1 = (s_{11}, \dots, s_{m1})$

$$\lim_{n \to \infty} \mathbf{P}(A_{nl}^-) = \lim_{n \to \infty} \{G_{\varepsilon_n}(\bar{s}_1)\}^{[t_r n_{\varepsilon_l}]} = 1. \tag{27}$$

Now, taking into account that for n large enough $t_m n_{\varepsilon_l} < t_1 n_{\varepsilon_n}$ and using A and E, we get by (27)

$$\begin{aligned} &\lim_{n \to \infty} \mathbf{P}(A_n \cap A_l) = \lim_{n \to \infty} \mathbf{P}(A_{nl}^+ \cap A_{nl}^- \cap A_l) \\ &= \lim_{n \to \infty} \mathbf{P}(A_{nl}^+)\mathbf{P}(A_{nl}^- \cap A_l) = \lim_{n \to \infty} \mathbf{P}(A_{nl}^+)\mathbf{P}(A_l) \\ &= \lim_{n \to \infty} \mathbf{P}(A_{nl}^+ \cap A_{nl}^-)\mathbf{P}(A_l) = \lim_{n \to \infty} \mathbf{P}(A_n)\mathbf{P}(A_l) = \mathbf{P}(A)\mathbf{P}(A_l). \end{aligned} \tag{28}$$

Since the sequence A_n is mixing in the sense of [16], $\mathbf{P}(A_n \cap B) \to \mathbf{P}(A)\mathbf{P}(B)$ as $n \to \infty$ for an arbitrary random event B. We can choose this set as $B_{\bar z} = \{\bar\nu_0 \leqq \bar z\}$. Let also $B_{\bar z,n} = \{\bar\nu_{\varepsilon_n} \leqq \bar z\}$. From Condition F follows that $\mathbf{P}(B_{\bar z} \Delta B_{\bar z,n}) \to 0$ as $n \to \infty$ for any $\bar z$ which is a point of continuity for the distribution function of the random vector $\bar\nu_0$. Using these asymptotic relations we finally get

$$\lim_{n\to\infty} \mathbf{P}(A_n \cap B_{\bar z,n}) = \lim_{n\to\infty} \mathbf{P}(A_n \cap B_{\bar z}) = \mathbf{P}(A)\mathbf{P}(B_{\bar z}). \tag{29}$$

As the choices of a subsequence ε_n, the points $0 < t_1 < \cdots < t_m < \infty$ and $s_1 \leqq \ldots \leqq s_m < \infty$ are all arbitrary, relation (29) leads to (25).

From (25) it follows that Condition C holds with the independence between the limiting process $\{\xi_0(t), t > 0\}$ and the random vector $\bar\nu_0$. Thus, Theorem 3 follows directly from Theorem 2. □

REFERENCES

[1] Barndorff-Nielsen, O., On the limit distribution of the maximum of a random number of independent random variables, *Acta Math. Acad. Sci. Hungar.* **15** (1964), 399–403. *MR* **30** #598

[2] Beirlant, J. and Teugels, J. L., Limit distributions for compounded sums of extreme order statistics, *J. Appl. Probab.* **29** (1992), 557–574. *MR* **93k**:60051

[3] Berman, S. M., Limiting distribution of the maximum term in sequences of dependent random variables, *Ann. Math. Statist.* **33** (1962), 894–908. *MR* **25** #5535

[4] Berman, S. M., *Sojourns and extremes of stochastic processes*, The Wadsworth & Brooks/Cole Statistics/Probability Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992. *MR* **93a**:60055

[5] Eriksson, E., Limit theorems for extreme statistics with random sample size, Master's Thesis, Dept. of Math. Stat., Umea Univ., 1995.

[6] Galambos, J., The distribution of the maximum of a random number of random variables with applications, *J. Appl. Probab.* **10** (1973), 122–129. *MR* **50** #5903

[7] Galambos, J., Limit laws for mixtures with applications to asymptotic theory of extremes, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **32** (1975), 197–207. *MR* **52** #1838

[8] Galambos, J., *The asymptotic theory of extreme order statistics*, Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1978. *MR* **80b**:60040

[9] Galambos, J., Random sample sizes: limit theorems and characterizations, *Probability Theory and Applications* (Ed. by J. Galambos and I. Kátai), Math. Appl. 80, Kluwer Acad. Publ., Dordrecht, 1992, 107–123. *MR* **94a**:60029

[10] Galambos, J., The development of the mathematical theory of extremes in the past half-century, *Teor. Veroyatnost. i Primenen.* **39** (1994), 272–293 (in Russian). Translation: *Theory Probab. Appl.* **39** (1994), 234–248. *MR* **97e**:60085

[11] Gnedenko, B. V. and Gnedenko, D. B., Laplace distribution and the logistic distribution as limit distributions in probability theory, *Serdica* **8** (1982), 229–234 (in Russian). *MR* **84m**:60029

[12] Korolev, V. Yu., Convergence of random sequences with independent random indices. II, *Theory Probab. Appl.* **40** (1995), 770–772. *MR* **97e**:60045

[13] Leadbetter, M. R., Lindgren, G. and Rootzén, H., *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983. *MR* **84h**:60050

[14] Mogyoródi, J., On the limiting distribution of the maximal term of a stratified sample with a random number of individuals, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **17** (1967), 75–83. *MR* **36** #2205

[15] Resnick, S. I., *Extreme values, regular variation, and point processes*, Applied Probability. A series of the Applied Probability Trust, 4, Springer-Verlag, New York, 1987. *MR* **89b**:60241

[16] Rényi, A., On mixing sequences of sets, *Acta Math. Acad. Sci. Hungar.* **9** (1958), 215–228. *MR* **20** #4623

[17] Sen, P. K., On weak convergence of extremal processes for random sample sizes, *Ann. Math. Statist.* **43** (1972), 1355–1362. *MR* **47** #1108

[18] Silvestrov, D. S., Remarks on the limit of a composite random function, *Teor. Veroyatnost. i Primenen.* **17** (1972), 707–715 (in Russian). *MR* **47** #5932

[19] Silvestrov, D. S., *Limit theorems for composite random functions*, Kiev University Press, Kiev, 1974 (in Russian). *MR* **54** #3811

[20] Silvestrov, D. S. and Teugels, J. L., Limit theorems for extremes with random sample size. Research Report No. 1, Department of Mathematical Statistics, Umeå University, 1997 (To appear in *Adv. Appl. Probab.*).

[21] Thomas, D. L., On limiting distributions of a random number of dependent random variables, *Ann. Math. Statist.* **43** (1972), 1719–1726. *MR* **50** #14897

(Received June 5, 1997)

DEPARTMENT OF MATHEMATICAL STATISTICS
UMEÅ UNIVERSITY
S–90187 UMEÅ
SWEDEN

dmitrii@matstat.umu.se

DEPARTEMENT WISKUNDE
FAKULTEIT WETENSCHAPPEN
KATHOLIEKE UNIVERSITEIT LEUVEN
CELESTIJNENLAAN 200B
B–3001 LEUVEN (HEVERLEE)
BELGIUM

jef.teugels@wis.kuleuven.ac.be

RECENTLY ACCEPTED PAPERS

BOGNÁR MÁTHÉ, K., Über ein halbreguläres Polyeder

BOGNÁR MÁTHÉ, K. and BÖRÖCZKY, K., Regular polyhedra and Hajós polyhedra

BÖLCSKEI, A., On weakly-neighbourly polyhedra

ČOMIĆ, I., One version of Miron's geometry in Osc M (II)

DIETHELM, K., Existence and construction of definite estimation functionals

DZIEDZIUL, K., Quasi-projectors

EWERT, J., Semicontinuity of topological limits of multivalued maps

IMHOF, L., On a problem of Dickmeis and Nessel

IVANČO, J. and TRENKLER, M., 3-polytopes with constant weight of faces

LUU, D. Q., On convergence in probability of martingale-like sequences

MALKOWSKY, E. and RAKOCEVIC, V., The measure of noncompactness of linear operators between spaces of Mth-order difference sequences

RENDER, H., Generalized uniform spaces and applications to function spaces

SHI, Z., A problem of Erdős–Révész on one-dimensional random walks

TANIGAWA, T., Asymptotic behaviour of positive solutions to nonlinear singular differential equations of second order

WU, J. and LI, R., Basic properties of locally convex A-spaces

ZHANG, W., On the primitive roots and the quadratic residues modulo p

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.

PRINTED IN HUNGARY
Akadémiai Nyomda, Martonvásár

CONTENTS

HU ISSN 0081–6906

Studia Scientiarum Mathematicarum Hungarica

VOLUME 34
NUMBER 4
1998

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 1 318 2875 Fax: +36 1 317 7166
e-mail: merza @ math-inst.hu

Subscription information

Orders should be addressed to

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

For 1999 volume 35 is scheduled for publication. The subscription price is $ 164.00, air delivery plus $ 20.00.

Coden: SSMHAX October, 1998
Vol: 34 **Pages**: 401-486
Number: 4 **Whole**: 67

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 401–412

PERFECTNESS OF REVERSIBLE MATRIX MAPS

E. JÜRIMÄE

Dedicated to Professor Károly Tandori on the occasion of his 70th birthday

1. Introduction

In the present paper we shall consider the matrix maps $y = Ax$, i.e.

$$y_n = \sum_k a_{nk} x_k, \qquad n \in \mathbb{N},$$

where $A = (a_{nk})$ is an infinite matrix of complex scalars and $x = (x_k), y = (y_n)$. Let X and Y be sets of sequences of complex numbers. We shall write $A \in (X, Y)$ if $Ax \in Y$ for every $x \in X$. We shall consider the case $Y = c_\pi$ (see Section 2). For these matrices domains $c_{\pi A}$ are defined. We shall investigate some properties of these domains under the hypothesis $A \in (c_\rho, c_\pi)$ or $A \in (c^\lambda, c_\pi)$, where c_ρ is a rate-space with rate ρ and c^λ a space with speed λ. Similar properties have been studied by several authors in case $A \in (c, c)$, i.e. for conservative matrices (see [5, 8]). The purpose of this paper is to show that many concepts (coregularity, conullity, perfectness etc.) and methods applied in summability could be profitable for more general cases of matrix maps. These possibilities are demonstrated by the study of conditions for the perfectness. By the way, our general definition and the investigations in Sections 3 and 4 originate from the ideas of [6] and [3, 4].

If Y is an FK-space, then the set (the domain of A)

$$Y_A := \{x \in \omega \mid Ax \in Y\}$$

is also an FK-space.

Definition 1.1. If $A \in (X, Y)$ and

$$\mathrm{cl}_{Y_A} X = Y_A,$$

then A is said to be (X, Y)-perfect.

1991 *Mathematics Subject Classification*. Primary 40A05, 40D09, 40H05; Secondary 46A45, 46B15.

Key words and phrases. Matrix map, rate-space, space with speed, perfectness, coregularity, conullity.

This research was supported in part by Estonian Scientific Foundation Grant No. 276.

In Section 3 we shall consider the case $A \in (c_\rho, c_\pi)$ and give some conditions for (c_ρ, c_π)-perfectness for coregular and also for conull matrices. It is shown that a matrix A can be perfect by one pair of spaces (c_ρ, c_π) but not by another pair.

In Section 4 the same questions are studied if $A \in (c^\lambda, c_\pi)$.

In the investigations in Sections 3 and 4 we shall apply the concepts of the spaces $m_\pi, c_\pi, c_{0\pi}$ and l_π (see Section 2) the rate-spaces with rate π.

2. Notations and preliminaries

Let $\pi = (\pi_n)$ be a sequence of positive numbers and ω be the set of all sequences of complex numbers. We shall consider the sets:

$$m_\pi := \{x = (x_k) \in \omega \mid (\frac{x_k}{\pi_k}) \in m\},$$

$$c_\pi := \{x \in m_\pi \mid \exists \lim_\pi x := \lim_n \frac{x_n}{\pi_n}\},$$

$$c_{0\pi} := \{x \in c_\pi \mid \lim_\pi x = 0\},$$

$$l_\pi := \{x \in c_{0\pi} \mid \sum_k \mid \frac{x_k}{\pi_k} \mid < \infty\}.$$

The sets m_π, c_π and $c_{0\pi}$ are BK-spaces with norm

$$\| x \| = \sup_k \mid \frac{x_k}{\pi_k} \mid .$$

Some properties of the spaces m_π, c_π and $c_{0\pi}$ are investigated in [1]. It is also shown there that these spaces are closely connected with spaces

$$c^\lambda := \{x \in c \mid \exists \lim_k \lambda_k (x_k - \lim x)\},$$

where $\lambda = (\lambda_k)$ is a sequence of positive numbers with $\lim \lambda = \infty$. The connection between the rate-spaces and spaces with speed is given by the relation

$$c^\lambda = c_{\frac{1}{\lambda}} \oplus < e >,$$

where $e = (1, 1, ...)$ (see [1]).

Here we shall mention some properties of the spaces l_π which are BK-spaces with norms

$$\| x \| = \sum_k \mid \frac{x_k}{\pi_k} \mid .$$

Let $e_k = (0, ..., 0, 1, 0, ...)$ (non-zero term 1 in the k-th place) and ϕ be the set of all finitely non-zero sequences. Then we have the next properties.

PROPOSITION 2.1. $e_k \in l_\pi, \quad \forall k \in \mathbb{N}$.

PROPOSITION 2.2. $\phi \subset l_\pi$.

PROPOSITION 2.3. $0 < \delta \leqq \pi_n \quad \forall n \in \mathbb{N} \Rightarrow l_\pi \subset l$.

PROPOSITION 2.4. $0 < \delta \leqq \frac{\pi_n}{\rho_n} \leqq M \Leftrightarrow l_\pi = l_\rho$.

PROPOSITION 2.5. $\rho_n \neq O(\pi_n) \Leftrightarrow l_\rho \supsetneqq l_\pi$.

PROPOSITION 2.6. *l_π has AK (sectional convergence), i.e.*

$$x = \sum_k x_k e_k$$

for every $x \in l_\pi$.

The proofs of these propositions are trivial. We mention only that 2.5 follows from the theorem of Knopp-Lorentz on the map $A \in (l, l)$.

If $A = (a_{nk})$ is an infinite matrix, then we call the set

$$c_{\pi A} := \{x \in \omega \mid Ax \in c_\pi\}$$

the domain of the matrix A. It is an FK-space and every $f \in (c_{\pi A})'$ has a representation

$$f(x) = \sum_k t_k x_k + \sum_n \tau_n \sum_k a_{nk} x_k + \mu \lim_{\pi A} x, \tag{1}$$

where

$$(t_k) \in (c_{\pi A})^\beta, \ \mu \in \mathbb{C}, \ \lim_{\pi A} x := \lim_n \frac{1}{\pi_n} \sum_k a_{nk} x_k,$$

$$\tau = (\tau_n) \in l_{\frac{1}{\pi}} \quad \text{and} \quad \frac{1}{\pi} := (\pi_k^{-1}).$$

If A is c_π-reversible, i.e. for each $y \in c_\pi$ there exists a unique x such that $Ax = y$, then $c_{\pi A}$ is a BK-space and every $f \in (c_{\pi A})'$ has a representation (1) with $t_k = 0 \ \forall k \in \mathbb{N}$.

The next statements are true (see [1]).

THEOREM 2.7. *A matrix $A = (a_{nk}) \in (c_\rho, c_\pi)$ if and only if it satisfies the following conditions:*

(i) $\quad \exists \lim_{\pi A} e_k = \lim_n \frac{a_{nk}}{\pi_n} =: a_k^\pi, \qquad k \in \mathbb{N}$,

(ii) $\quad \exists \lim_{\pi A} \rho = \lim_n \frac{1}{\pi_n} \sum_k a_{nk} \rho_k =: a^{\rho\pi}$,

(iii) $\quad \sum_k |a_{nk}| \rho_k = O(\pi_n)$.

If $x \in c_\rho$, *then*

$$\lim_{\pi A} x = \Big(a^{\rho\pi} - \sum_k a_k^\pi \rho_k\Big) \lim_\rho x + \sum_k a_k^\pi x_k.$$

THEOREM 2.8. *A matrix* $A = (a_{nk}) \in (c^\lambda, c_\pi)$ *if and only if the following statements are true:*

(i) $\exists \lim_{\pi A} e_k = a_k^\pi, \quad k \in \mathbb{N},$

(ii) $\exists \lim_{\pi A} \dfrac{1}{\lambda} = \lim_n \dfrac{1}{\pi_n} \sum_k \dfrac{a_{nk}}{\lambda_k} =: a^{\lambda^{-1}\pi},$

(iii) $\exists \lim_{\pi A} e = \lim_n \dfrac{1}{\pi_n} \sum_k a_{nk} =: a^{1\pi},$

(iv) $\sum_k \dfrac{|a_{nk}|}{\lambda_k} = O(\pi_n).$

If $x \in c^\lambda$, *then*

$$\lim_{\pi A} x = a^{1\pi} \lim x + \Big(a^{\lambda^{-1}\pi} - \sum_k \frac{a_k^\pi}{\lambda_k}\Big) \lambda(x) + \sum_k \frac{a_k^\pi}{\lambda_k} \lambda^k(x),$$

where

$$\lambda^k(x) = \lambda_k(x_k - \lim x) \quad \text{and} \quad \lambda(x) = \lim_k \lambda^k(x).$$

Let $A \in (c_\rho, c_\pi)$. Then the matrix A is called (c_ρ, c_π)-coregular (see [2]) if

$$\chi_{c\rho}^{c\pi}(A) := a^{\rho\pi} - \sum_k a_k^\pi \rho_k \neq 0$$

and (c_ρ, c_π)-conull if the characteristic $\chi_{c\rho}^{c\pi}(A) = 0$. Similar definitions are also used for matrices $A \in (c^\lambda, c_\pi)$. In this case

$$\chi_{c^\lambda}^{c\pi}(A) := a^{\lambda^{-1}\pi} - \sum_k \frac{a_k^\pi}{\lambda_k}.$$

3. (c_ρ, c_π)-perfectness

In this and in the next section we shall only consider c_π-reversible matrices whereas the representation of linear continuous functionals which have an important role in our investigations is simpler in this case.

By 1.1 a matrix A is said to be (c_ρ, c_π)-perfect if

$$\mathrm{cl}_{c_{\pi A}} c_\rho = c_{\pi A}.$$

The set $E=\{\rho,\ e_k \mid k\in\mathbb{N}\}$ is fundamental (see [1]) for c_ρ. Therefore we get by the Hahn–Banach theorem that A is (c_ρ,c_π)-perfect if and only if $f(x)=0\ \forall x\in E$ implies $f=0$. By (1) it means that the system

$$\text{(2)}\qquad \begin{cases} \sum_n a_{nk}\tau_n+\mu\lim_{\pi A} e_k=0, & k\in\mathbb{N},\\ \sum_n\sum_k a_{nk}\rho_k\tau_n+\mu\lim_{\pi A}\rho=0 \end{cases}$$

has in $l_{\frac{1}{\pi}}$ only trivial solution $\tau_n=\mu=0, \forall n\in\mathbb{N}$.

If A is (c_ρ,c_π)-coregular, then the system (2) is equivalent to the system

$$\text{(3)}\qquad \sum_n a_{nk}\tau_n=0,\quad k\in\mathbb{N}.$$

So we get the next theorem.

THEOREM 3.1. *Let A be c_π-reversible and (c_ρ,c_π)-coregular. The next conditions are equivalent:*

(i) *A is (c_ρ,c_π)-perfect;*

(ii) *(3) has only trivial solution in $l_{\frac{1}{\pi}}$;*

(iii) *the matrix $(\pi_n^{-1}a_{nk})$ is of type M.*

COROLLARY 3.2. *If a c_π-reversible and (c_ρ,c_π)-coregular matrix A is (c_ρ,c_π)-perfect, then it is (c_κ,c_π)-perfect for any κ whenever A is (c_κ,c_π)-coregular.*

EXAMPLE 3.3. Let

$$a_{nk}=\begin{cases} 1 & \text{if}\quad k=n-1,\\ \frac{1}{n} & \text{if}\quad k=n,\\ 0 & \text{if}\quad k\neq n-1,n. \end{cases}$$

This matrix $A=(a_{nk})$ is triangular and (c,c)-coregular. For A all solutions of the system (3) can be represented in the form

$$\overline{\tau}=(((-1)^n\big/(n-1)!)\tau_1),$$

where $\tau_1\in\mathbb{C}$. It means that this matrix is not (c,c)-perfect.

The matrix A is also (c_ρ,c_π)-coregular if $\rho_k=\frac{1}{2}k!$ and $\pi_n=(n-1)!$. Whereas $\overline{\tau}\in l_{\frac{1}{\pi}}$ only for $\overline{\tau}_1=0$, then A is (c_ρ,c_π)-perfect.

COROLLARY 3.4. *Let $M \geqq \pi_n \geqq \varepsilon > 0$ $\forall n \in \mathbb{N}$. Then a c_π-reversible and (c_ρ, c_π)-coregular matrix A is (c_ρ, c_π)-perfect if and only if A is of type M.*

PROOF. This assertion follows from 2.4 and 3.1. □

The class "matrix of type M" can be defined only for coregular matrices. A similar class of conservative matrices was introduced by the author in 1970 and denoted by P.

DEFINITION 3.5. A conservative matrix $B = (b_{nk})$ is said to be of type P if the system

$$\begin{cases} \sum\limits_n b_{nk}\tau_n + \mu b_k = 0, & k \in \mathbb{N}, \\ \chi(B)\mu = 0 \end{cases}$$

has only trivial solution $\tau_n = \mu = 0$, $n \in \mathbb{N}$, in l.

In [4] G. Kangro proved the next theorem.

THEOREM 3.6. *Let $B = (b_{nk})$ be a conservative matrix with right inverse $B' = (b'_{nk})$. Then B is of type P if*

(i) *B is coregular and $b^n := (b'_{nk})_{k=1}^\infty \in m$*

or

(ii) *B is conull, $b^n = (b'_{nk})_{k=1}^\infty \in c$ and $b_k = \lim\limits_n b_{nk} \neq 0$ for some k.*

The right inverse of a matrix A is connected with the existence and form of the solution of the equation $y = Ax$. This becomes obvious in the next theorem.

THEOREM 3.7. *Let $A = (a_{nk})$ be a c_π-reversible matrix. Then A has a unique right inverse $A' = (a'_{kv})$. The rows of A' belong to $l_{\frac{1}{\pi}}$. There is a sequence b such that the equation $y = Ax$ has, for $y \in c_\pi$, the unique solution*

$$x = b \lim_\pi y + A'y.$$

PROOF. Applying the representation of $f \in (c_{\pi A})'$ to the coordinates, we have for $x \in c_{\pi A}$ and $y = Ax$, $x_k = \mu_k \lim\limits_{\pi A} x + \sum\limits_n \tau_{kn} \sum\limits_\nu a_{n\nu} x_k = \mu_k \lim\limits_\pi y + \sum\limits_n \tau_{kn} y_n$, where $(\tau_{kn})_{n=1}^\infty \in l_{\frac{1}{\pi}}$. Now setting $b = (\mu_k)$ and $a'_{k\nu} = \tau_{k\nu}$ we have all the theorem except that $A' = (a'_{k\nu})$ is a right inverse. We see this by taking $y = e_\nu = (\delta_{\nu n})$. Then $\lim\limits_\pi e_k = 0$ and $x_k = \sum\limits_n \tau_{nk}\delta_{\nu n} = \tau_{k\nu} = a'_{k\nu}$. The equation $y = Ax$ becomes

$$\delta_{\nu n} = \sum_k a_{nk} x_k = \sum_k a_{nk} a'_{k\nu},$$

i.e. $AA' = I$. □

The proof for the case $\pi = e$ is given in [8] (Theorem 5.4.5).

LEMMA 3.8. (i) *If* $A=(a_{nk})\in(c_\rho,c_\pi)$, *then* $B=\left(\frac{a_{nk}\rho_k}{\pi_n}\right)\in(c,c)$, *i.e. B is conservative.*

(ii) *If A is* (c_ρ,c_π)*-coregular, then B is* (c,c)*-coregular.*

(iii) *If A has a right inverse* $A'=(a'_{k\nu})$, *then B has a right inverse* $B'=\left(\frac{a'_{k\nu}\pi_\nu}{\rho_k}\right)$.

PROOF. Theorem 2.7 implies (i) and (ii). For (iii) we have

$$\sum_k \frac{a_{nk}\rho_k}{\pi_n}\frac{a'_{k\nu}\pi_\nu}{\rho_k}=\frac{\pi_\nu}{\pi_n}\sum_k a_{nk}a'_{k\nu}=\frac{\pi_\nu}{\pi_n}\delta_{n\nu}=\delta_{n\nu}. \qquad \square$$

The assertion "system (2) has only trivial solution in $l_{\frac{1}{\pi}}$" is equivalent to the statement "the matrix $B=(\frac{a_{nk}\rho_k}{\pi_n})$ is of type P". Therefore we get the next theorem.

THEOREM 3.9. *Let* $A\in(c_\rho,c_\pi)$ *be a* c_π*-reversible matrix. Then the next statements are equivalent:*

(i) $A=(a_{nk})$ *is* (c_ρ,c_π)*-perfect;*

(ii) $B=(\frac{a_{nk}\rho_k}{\pi_n})$ *is* (c,c)*-perfect;*

(iii) B *is of type P.*

Applying Theorems 3.6 and 3.9 and Lemma 3.8 we get the next assertion.

THEOREM 3.10. *Let* $A=(a_{nk})\in(c_\rho,c_\pi)$ *be a* c_π*-reversible matrix with right inverse* $A'=(a'_{nk})$. *Then A is* (c_ρ,c_π)*-perfect if*

(i) $\chi_{c\rho}^{c\pi}(A)\neq 0$ *and* $(a'_{nk})_{k=1}^{\infty}\in m_{\frac{1}{\pi}}$

or

(ii) $\chi_{c\rho}^{c\pi}(A)=0$, $(a'_{nk})_{k=1}^{\infty}\in c_{\frac{1}{\pi}}$ *and* $a_k^\pi\neq 0$ *for some k.*

EXAMPLE 3.11. Let $A=(R_\kappa,p_k)$ with $p_k>0$ be Riesz matrix of order κ. Then, for inverse matrix, $a_{k\nu}^{-1}=0$ for $k>\nu+\kappa$ (see [7]). It means that $a^\nu=(a_{k\nu}^{-1})_{k=1}^{\infty}\in m_{\frac{1}{\pi}}$. Therefore, every (c_ρ,c_π)-coregular Riesz matrix is (c_ρ,c_π)-perfect.

4. (c^λ,c_π)-perfectness

In this section we shall study the conditions for (c^λ,c_π)-perfectness which are a little more complicated because of the structure of c^λ. By Definition 1.1 a matrix $A=(a_{nk})$ is (c^λ,c_π)-perfect if

$$\mathrm{cl}_{c_{\pi A}}c^\lambda=c_{\pi A}.$$

The set $G=\{e,\lambda^{-1},e_k \mid k\in\mathbb{N}\}$ is fundamental in c^λ (see [1]). Therefore a c_π-reversible A is (c^λ,c_π)-perfect if and only if the next system ($f\in(c_{\pi A})'$)

$$
\begin{aligned}
&f(e_k)=\sum_n a_{nk}\tau_n+\mu a_k^\pi=0, \quad k\in\mathbb{N},\\
(4)\qquad &f(e)=\sum_n\sum_k a_{nk}\tau_n+\mu a^{1\pi}=0,\\
&f(\lambda^{-1})=\sum_n\sum_k \frac{a_{nk}}{\lambda_k}\tau_n+\mu a^{\lambda^{-1}\pi}=0
\end{aligned}
$$

has in $l_{\frac{1}{\pi}}$ only trivial solution $\mu=\tau_n=0 \quad \forall n\in\mathbb{N}$.

If A is a (c^λ,c_π)-coregular matrix, then the next theorem is true.

THEOREM 4.1. *Let $A=(a_{nk})$ be a c_π-reversible and (c^λ,c_π)-coregular matrix. Then A is (c^λ,c_π)-perfect if and only if the system*

$$
\begin{aligned}
&\sum_n a_{nk}\tau_n=0, \quad k\in\mathbb{N},\\
(5)\qquad &\sum_n\sum_k a_{nk}\tau_n=0
\end{aligned}
$$

has in $l_{\frac{1}{\pi}}$ only trivial solution $\tau=(\tau_n)=0$.

In the general case the last equation of (5) does not follow from the others.

EXAMPLE 4.2. Let

$$
b_{nk}=\begin{cases} 1+2^{n-1}, & k=n-1,\\ -2^{n-1}, & k=n,\\ 0, & k\neq n-1,n.\end{cases}
$$

If $\lambda=(2^n)$ then $B=(b_{nk})\in(c^\lambda,c)$. For this matrix the system

$$\sum_n b_{nk}\tau_n=0, \quad k\in\mathbb{N},$$

has in l solutions $\tau^\circ=(\tau_n^\circ)$, where

$$\tau_n^\circ=\frac{2^2\cdot 2^3\cdot\ldots\cdot 2^{n-2}}{(1+2^2)\ldots(1+2^{n-1})}\tau_1<\frac{\tau_1}{2^{n-1}}, \quad n>1.$$

We have

$$\sum_n\sum_k b_{nk}\tau_n^\circ=\sum_n\tau_n^\circ\neq 0 \quad \text{if} \quad \tau_1\neq 0.$$

The matrix B is (c^λ, c)-coregular and a triangle ($b_{nn} \neq 0$ and $b_{nk} = 0,\ k > n$), therefore it is also c-reversible.

Our investigation implies that

$$\mathrm{cl}_{c_B}\{\lambda^{-1}, e_k \mid k \in \mathbb{N}\} \neq c_B$$

but

$$\mathrm{cl}_{c_B}\{\lambda^{-1}, e, e_k, \mid k \in \mathbb{N}\} = c_B.$$

This means that B is not $(c_{\lambda^{-1}}, c)$-perfect but it is (c^λ, c)-perfect. □

If the rate $\pi = (\pi_n)$ is sufficiently great for the matrix A, then we can omit the last equation of (5). This follows from the next theorem.

THEOREM 4.3. *Let $A = (a_{nk})$ be a c_π-reversible, (c^λ, c_π)-coregular matrix and $\sup\limits_{n,m} \frac{1}{\pi_n} \mid \sum\limits_{k=1}^{m} a_{nk} \mid < \infty$. Then the following statements are equivalent:*

(i) *A is (c^λ, c_π)-perfect;*

(ii) *the system*

$$\sum_n a_{nk}\tau_n = 0, \quad \forall k \in \mathbb{N},$$

has only trivial solution in $l_{\frac{1}{\pi}}$;

(iii) *the matrix $(\pi_n^{-1} a_{nk})$ is of type M.*

PROOF. We must ascertain that for $\tau = (\tau_n) \in l_{\frac{1}{\pi}}$ the equality

$$\sum_n \sum_k a_{nk}\tau_n = 0$$

follows from

$$\sum_n a_{nk}\tau_n = 0, \qquad \forall k \in \mathbb{N}.$$

It is so if the limit

$$\lim_m \sum_{k=1}^{m} \sum_n a_{nk}\tau_n = \lim_m \sum_n \sum_{k=1}^{m} a_{nk}\tau_n$$

exists for any $\tau = (\tau_n) \in l_{\frac{1}{\pi}}$. This is equivalent to the condition:

$$\lim_m \sum_n \sum_{k=1}^{m} \frac{a_{nk}}{\pi_n}(\pi_n\tau_n)$$

exists for every $(\pi_n\tau_n) \in l$. By Hahn's theorem (see [8], Ch. 8) this is true if

$$\sup_{n,m} \Big| \sum_{k=1}^{m} \frac{a_{nk}}{\pi_n} \Big| < \infty. \qquad \square$$

We can represent the conditions for perfectness of a matrix by the right inverse of A (cf. [8], Ch. 3 and [4]). Similar assertions are also true in our case.

THEOREM 4.4. *Let A be a c_π-reversible, (c^λ, c_π)-coregular matrix and suppose that A has a right inverse $A' = (a'_{nk})$ with columns $(a'_{nk})_{n=1}^\infty \in m_{\frac{1}{\lambda}}$. Then A is (c^λ, c_π)-perfect.*

PROOF. If

$$\sum_n a_{nk}\tau_n = 0, \quad \forall k \in \mathbb{N},$$

then by 2.8 for any $\nu \in \mathbb{N}$

$$0 = \sum_k a'_{k\nu} \sum_n a_{nk}\tau_n = \sum_k \sum_n (\tau_n \pi_n) \frac{a_{nk}}{\pi_n \lambda_k} (\lambda_k a'_{k\nu}) =$$

$$= \sum_n \tau_n \sum_k a_{nk} a'_{k\nu} = \sum_n \tau_n \delta_{n\nu} = \tau_\nu. \qquad \square$$

For (c^λ, c_π)-conull matrices the third equation in (4) is an implication from the first one. So, we can say that a c_π-reversible and (c^λ, c_π)-conull matrix $A = (a_{nk})$ is (c^λ, c_π)-perfect if and only if the system

$$\text{(6)} \qquad \begin{aligned} f(e_k) &= \sum_n a_{nk}\tau_n + \mu a_k^\pi = 0, \quad \forall k \in \mathbb{N}, \\ f(e) &= \sum_n \sum_k a_{nk}\tau_n + \mu a^{1\pi} = 0 \end{aligned}$$

has in $l_{\frac{1}{\pi}}$ only trivial solution $\mu = \tau_n = 0 \quad \forall n \in \mathbb{N}$.

We shall prove

THEOREM 4.5. *Let $A = (a_{nk})$ be a c_π-reversible, (c^λ, c_π)-conull matrix with $a_k^\pi \neq 0$ for some k and suppose that A has right inverse $A' = (a'_{k\nu})$ with columns $a^\nu = (a'_{k\nu})_{k=1}^\infty \in c^\lambda \bigcap c_0$. Then A is (c^λ, c_π)-perfect.*

PROOF. Since $Aa^\nu = e_\nu = (\delta_{n\nu})$ we have by 2.8

$$0 = \lim_n \frac{\delta_{n\nu}}{\pi_n} = \lim_{\pi A} a^\nu = a^{1\pi} \lim_k a'_{k\nu} + \sum_k \frac{a_k^\pi}{\lambda_k} \lambda^k(a^\nu) =$$

$$= a^{1\pi} \lim_k a'_{k\nu} + \sum_k \frac{a_k^\pi}{\lambda_k} \lambda_k (a'_{k\nu} - \lim_k a'_{k\nu}).$$

The hypothesis $a^\nu = (a'_{k\nu})_{k=1}^\infty \in c^\lambda \bigcap c_0$ gives that

$$\sum_k a_k^\pi a'_{k\nu} = 0, \ \nu \in \mathbb{N}.$$

Applying this and Theorem 2.8, the first equation of the last system (6) implies that

$$0=\mu\sum_k a'_{k\nu}a^{\pi}_k+\sum_k a'_{k\nu}\sum_n a_{nk}\tau_n=$$

$$=\sum_k\sum_n(\pi_n\tau_n)\frac{a_{nk}}{\pi_n\lambda_k}(\lambda_k a'_{k\nu})=\sum_n\tau_n\sum_k a_{nk}a'_{k\nu}=$$

$$=\sum_n\tau_n\delta_{n\nu}=\tau_\nu.$$

The hypothesis $a^{\pi}_k\neq 0$ for some k implies now that $\mu=0$ in (6). This completes the proof of the theorem. □

5. Appendix

In this paper we have considered for reversible matrix maps some conditions for the perfectness which are described by several infinite systems of equations. These conditions are quite perspicuous and by them it is also quite easy to settle whether a given reversible matrix map is perfect or not.

For non-reversible matrix maps the (c,c)-perfectness has been studied by several authors using the test functions and the corresponding distinguished subsets of the domain (see [5, 8]). It could be assumed that this way can also be used in more general cases which have been discussed in the present paper. In connection with this, it should be emphasized that the nature of the perfect part of a matrix $A\in(c^{\lambda},c_{\pi})$ (and also $A\in(c_{\rho},c_{\pi})$) is an interesting question in itself.

REFERENCES

[1] Jürimäe, E., Matrix mappings between rate-spaces and spaces with speed, *Tartu Ülik. Toimetised*, **970** (1994), 29–52. *MR* **96j**:40006

[2] Jürimäe, E., Properties of domains of matrix mappings on rate-spaces and spaces with speed, *Tartu Ülik. Toimetised*, **970** (1994), 53–64. *MR* **96j**:40007

[3] Kangro, G., On λ-perfectness of summability methods, and applications of it. I, *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.* **20** (1971), 111–120 (in Russian). *MR* **45** #2372

[4] Kangro, G., On λ-perfectness of summability methods and applications of it. II, *Eesti NSV Tead. Akad. Toimetised Füüs.-Mat.* **20** (1971), 375–385 (in Russian). *MR* **45** #4004

[5] Leiger T., *Methods of functional analysis in summability*, Tartu, 1992 (in Estonian).

[6] Meyer-König, W. und Zeller, K., Tauber-Sätze und M-Perfektheit, *Math. Z.* **177** (1981), 257–266. *MR* **83d**:40006

[7] Türnpu, H., Summability factors for the Riesz method, *Tartu Riikl. Ülik. Toimetised* **206** (1967), 90–105 (in Russian). *MR* **41** #7335

[8] WILANSKY, A., *Summability through functional analysis*, Notas de matemática, 91; North-Holland Mathematics Studies, 85, North-Holland Publ. Co., Amsterdam–New York–Oxford, 1984. *MR* **85d**:40006

(Received February 26, 1996)

INSTITUTE OF PURE MATHEMATICS
UNIVERSITY OF TARTU
VANEMUISE 46–201
EE–2400 TARTU
ESTONIA

endel@math.ut.ee

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 413–428

THE METHOD OF LINES FOR FIRST ORDER PARTIAL DIFFERENTIAL-FUNCTIONAL EQUATIONS

BARBARA ZUBIK-KOWAL

Abstract

The Cauchy problem for nonlinear first order partial differential-functional equations in unbounded domains is treated with a general class of the method of lines. Existence and convergence properties of the method are investigated under the assumption that the right-hand side of the equation satisfies the Lipschitz condition with respect to the functional argument. The theorems are proved by means of the differential-difference inequalities technique. Examples of differential-functional problems and corresponding methods of lines are given.

1. Introduction

For any two metric spaces X and Y we denote by $C(X,Y)$ the class of all continuous functions from X into Y. We will use vectorial inequalities with the understanding that the same inequalities hold between their corresponding components. Let $D=[-\tau_0,0]\times[-\tau,\tau]$, where $\tau_0\in\mathbb{R}_+$, $\mathbb{R}_+=[0,+\infty)$ and $\tau=(\tau_1,\dots,\tau_n)\in\mathbb{R}^n_+$. Write

$$E=[0,a]\times\mathbb{R}^n,\quad E_0=[-\tau_0,0]\times\mathbb{R}^n,\quad \Theta=E\times C(D,\mathbb{R})\times\mathbb{R}^n$$

and suppose that

$$f\colon\Theta\to\mathbb{R},\quad \varphi\colon E_0\to\mathbb{R}$$

are given functions. For any function $z\colon E_0\cup E\to\mathbb{R}$ and for $(x,y)\in E$, $y=(y_1,\dots,y_n)$, we define a function $z_{(x,y)}\colon D\to\mathbb{R}$ by $z_{(x,y)}(t,s)=z(x+t,y+s)$, $(t,s)\in D$. The function $z_{(x,y)}$ is the restriction of z to the set $[x-\tau_0,x]\times[y-\tau,y+\tau]$ and this restriction is shifted to the set D.

The paper deals with the nonlinear Cauchy problem

$$D_x z(x,y)=f(x,y,z_{(x,y)},D_y z(x,y)),\quad (x,y)\in E, \tag{1}$$

$$z(x,y)=\varphi(x,y),\quad (x,y)\in E_0, \tag{2}$$

1991 *Mathematics Subject Classification.* Primary 65N40; Secondary 35A40.

Key words and phrases. Cauchy problem, unbounded solutions, differential-difference inequalities, comparison technique.

0081–6906/98/$ 5.00

where $D_y z = (D_{y_1} z, \dots, D_{y_n} z)$. We will consider classical solutions of problem (1), (2). We will assume that there exists a classical solution of (1), (2). Sufficient conditions for the existence can be found in [3], [4], [12].

In general, first order partial differential equations are used to describe the growth of a population of cells which constantly differentiate in time, see [6], [7]. The paper [10] discusses, using differential-integral equations, optimal harvesting policies for age-structured populations harvested with effort independent of age. In the theory of the distribution of wealth [8] a differential equation with a deviated argument is used. Differential-integral equations describing the dynamic of muscle contraction was studied in [9]. There are various problems in nonlinear optics [1] which lead to non-linear hyperbolic differential-integral equations.

Differential and differential-functional equations considered in [1], [6], [7], [8], [9], [10] are particular cases of (1). We give next examples in Section 6.

2. Discretization

For $y, \overline{y} \in \mathbb{R}^n$ we write $y * \overline{y} = (y_1 \overline{y}_1, \dots, y_n \overline{y}_n)$. We will use the letters $\mathbb{Z}$ and $\mathbb{N}$ to denote the set of integers and the set of natural numbers. Now we define a mesh in $\mathbb{R}^n$. Suppose that for $h = (h_1, \dots, h_n)$, where $h_i > 0$, there exists $N = (N_1, \dots, N_n) \in \mathbb{N}^n$ such that $N * h = \tau$. We denote by I_0 the set of all h having the above property. Let $m = (m_1, \dots, m_n) \in \mathbb{Z}^n$ and $y^m = m * h$, $y^m = (y_1^{m_1}, \dots, y_n^{m_n})$. Write $\mathbb{R}^n_h = \{y^m : m \in \mathbb{Z}^n\}$ and

$$E_h = [0, a] \times \mathbb{R}^n_h, \qquad E_{0.h} = [-\tau_0, 0] \times \mathbb{R}^n_h, \qquad D_h = D \cap E_{0.h}.$$

Let $\delta = (\delta_1, \dots, \delta_n)$ be a difference operator defined in the following way. Put $S = \{-1, 0, 1\}^n$ and

$$\delta_i z(x, y^m) = \frac{1}{h_i} \sum_{r \in S} c^{(i)}_{r,m} z(x, y^{m+r}), \tag{3}$$

where $z \colon E_{0,h} \cup E_h \to \mathbb{R}$ and $c^{(i)}_{r,m}$ are given numbers, $\delta z = (\delta z_1, \dots, \delta z_n)$. We will approximate $D_y z(x, y^m)$ by means of $\delta z(x, y^m)$. In the next part of the paper we adopt additional assumptions on $c^{(i)}_{r,m}$. Since the coefficients $c^{(i)}_{r,m}$ depend on $m \in \mathbb{Z}^n$, the approximation of the spatial derivative may be different in different points of the spatial mesh.

For $z \colon E_{0.h} \cup E_h \to \mathbb{R}$, $x \in [0, a]$, $m \in \mathbb{Z}^n$ we define a function $z_{(x,m)} \colon D_h \to \mathbb{R}$ by

$$z_{(x,m)}(t, y^s) = z(x + t, y^{m+s}), \quad (t, y^s) \in D_h.$$

Let $\Theta_h = E_h \times C(D_h, \mathbb{R}) \times \mathbb{R}^n$ and suppose that

$$\Phi_h \colon \Theta_h \to \mathbb{R} \qquad \phi_h \colon E_{0.h} \to \mathbb{R}$$

are given functions. We will approximate the solutions of (1), (2) by means of the solutions of the problem

$$D_x z(x, y^m) = \Phi_h(x, y^m, z_{(x,m)}, \delta z(x, y^m)), \quad (x, y^m) \in E_h, \tag{4}$$

$$z(x, y^m) = \phi_h(x, y^m), \quad (x, y^m) \in E_{0.h}. \tag{5}$$

The initial-value problem (4), (5) will be called the method of lines for (1), (2). The function Φ_h and the difference operator δ characterize a general class of differential-difference schemes which can be applied to (1), (2). We give sufficient conditions for the solvability of (4), (5) and for the convergence of the sequence $\{u_n\}$ of solutions of (4), (5) to a solution of (1), (2). We will consider unbounded solutions of (1), (2) and (4), (5).

The main theorems concerning (1), (2) and (4), (5) will be based on a comparison theorem where a function satisfying some differential-difference inequalities in an unbounded domain is estimated by a solution of an adequate ordinary differential-functional problem. The first result of this type was given by Łojasiewicz [13], (see Lemma 1, p. 96) for a function of two variables satisfying linear differential-difference inequalities with constant coefficients. It is an essential fact in [13] that the differential-difference inequalities are periodic with respect to the spatial variable.

The comparison result from [13] is extended in [20] (see also [5]) on differential-difference inequalities with a functional argument and on nonlinear comparison problems. It is an essential fact in [20] that the finite systems of differential-difference equations have been considered.

In the paper we consider infinite systems of equations (4) or infinite systems of differential-difference inequalities. The comparison theorem given in Section 3 is new also in the case when f does not depend on the functional argument.

There is a great amount of literature on the method of lines. The monograph by Walter [18] contains a large bibliography. The existence, uniqueness, monotonicity and convergence properties of the method of lines for the Cauchy problem for nonlinear parabolic differential equations in unbounded sets are given in [16]. The convergence of the method of lines for a parabolic differential equation is shown in [17]. The error estimations for the method applied to the first boundary value problem and for the Cauchy problem for parabolic equations are given in [18, 19]. The method is also treated as a tool for proving the existence theorems. Such existence theorems based on the method of lines are given in [18, 19]. The method of lines for a first order differential system with a functional right side as well as for an equation of the parabolic type is studied in [11]. The method for equations of higher orders is studied in [14].

The method of lines is treated as a tool of numerical solving of differential problems. The book [15] demonstrates lots of examples of the use of the numerical method of lines.

3. Comparison theorem

For a function $\omega\colon[-\tau_0,a]\to\mathbb{R}$ and for $x\in[0,a]$ we define a function $\omega_{(x)}\colon[-\tau_0,0]\to\mathbb{R}$ by $\omega_{(x)}(t)=\omega(x+t)$, $t\in[-\tau_0,0]$. We denote by $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^n$ and by $\|\cdot\|_0$ the supremum norm in $C([-\tau_0,0],\mathbb{R})$. We use the symbols $\|\cdot\|_D$ and $\|\cdot\|_{D_h}$ to denote the supremum norms in $C(D,\mathbb{R})$ and $C(D_h,\mathbb{R})$, respectively.

We will prove that problem (4), (5) has exactly one solution in the following class of functions.

DEFINITION. Let $\alpha,\beta\in\mathbb{R}$. A function $z\colon[\alpha,\beta]\times\mathbb{R}^n_h\to\mathbb{R}$ will be called the function of class $\sum$ if the following conditions are satisfied

(i) $z(\cdot,y^m)\in C([-\tau_0,a],\mathbb{R})$ for all $m\in\mathbb{Z}^n$;

(ii) there exists a function $\Psi\colon\mathbb{R}_+\to\mathbb{R}_+$ such that

(a) Ψ has a continuous derivative on $\mathbb{R}_+$ which is bounded and nonnegative,

(b) $\lim\limits_{t\to+\infty}\Psi(t)=+\infty$,

(c) for $t\in\mathbb{R}_+$ we have

$$\max\{|z(x,y^m)|: x\in[-\tau_0,a],\ \|y^m\|\leqq t\}\leqq\Psi(t).$$

In a comparison theorem we will estimate a function of several variables by means of a function of one variable. Therefore we will need the following operator $\Gamma_h\colon C(D_h\mathbb{R})\to C([-\tau_0,0]\mathbb{R}_+)$. If $w\in C(D_h,\mathbb{R})$ then

$$[\Gamma_h w](t)=\max\{|w(t,y^m)|: -N\leqq m\leqq N\}.$$

Suppose that $\sigma\colon[0,a]\times C([-\tau_0,0],\mathbb{R}_+)\to\mathbb{R}_+$ and $\lambda=(\lambda_1,\ldots,\lambda_n)\colon E_h\times C(D_h,\mathbb{R})\to\mathbb{R}^n$ are given functions. In the section we consider the differential-difference inequalities

$$\Big|D_z z(x,y^m)-\sum_{i=1}^{n}\lambda_i(x,y^m,z_{(x,m)})\delta_i z(x,y^m)\Big|\leqq\sigma(x,\Gamma_h z_{(x,m)}),\qquad(x,y^m)\in(0,a]\times\mathbb{R}^n_h. \tag{6}$$

THEOREM 1. *Suppose that*

1° *the function* $\sigma\colon[0,a]\times C([-\tau_0,0],\mathbb{R}_+)\to\mathbb{R}_+$ *is continuous, nondecreasing with respect to the functional argument and there is* $K\in\mathbb{R}_+$ *such that*

$$|\sigma(x,w)-\sigma(x,\bar w)|\leqq K\|w-\bar w\|_0\quad\text{on}\quad[0,a]\times C([-\tau_0,0],\mathbb{R}_+), \tag{7}$$

2° $u\colon E_{0.h}\cup E_h\to\mathbb{R}$ *is of class* Σ, *there exists the derivative* $D_x u(x,y)$ *for* $(x,y)\in(0,a]\times\mathbb{R}^n_h$, u *satisfies differential-difference inequalities* (6), 3° *there are* $L=(L_1,\dots,L_n)\in\mathbb{R}^n_+$ *and* $\eta\in C([-\tau_0,0],\mathbb{R}_+)$ *such that*

$$|\lambda_i(x,y,w)|\leqq L_i \quad \text{on} \quad E_h\times C(D_h,\mathbb{R}), \quad i=1,\dots,n$$

and

$$|u(x,y)|\leqq\eta(x) \quad \text{on} \quad E_{0.h}, \tag{8}$$

4° *the operator* δ *and the function* λ *satisfy the condition*

$$c^{(i)}_{r,m}\lambda_i(x,y^m,w)\geqq 0 \quad \text{for} \quad r\in S-\{0\},$$
$$(x,y^m)\in(0,a]\times\mathbb{R}^n_h, \quad i=1,\dots,n,$$

5° *there exists* C_1 *such that* $\sum_{r\in S}|c^{(i)}_{r,m}|\leqq C_1$ *for* $i=1,\dots,n$, $m\in\mathbb{Z}^n$.

Under these assumptions we have

$$|u(x,y)|\leqq\omega(x;\eta), \quad (x,y)\in E_h,$$

where $\omega(\cdot;\eta)$ *is the solution of the problem*

$$D_x\omega(x)=\sigma(x,\omega_{(x)}), \quad x\in(0,a],$$
$$\omega(x)=\eta(x), \quad x\in[-\tau_0,0].$$

PROOF. We first show that $u(x,y^m)\leqq\omega(x;\eta)$ for $(x,y^m)\in[0,a]\times\mathbb{R}^n_h$. Let us define

$$v(x,y)=u(x,y)-\omega(x;\eta), \quad (x,y)\in E_{0.h}\cup E_h$$

and

$$\Psi_0(t)=\max\{v(x,y):(x,y)\in E_h,\|y\|\leqq t\}, \quad t\in\mathbb{R}_+.$$

It follows from assumption 2° that there exists a function $\Psi\colon\mathbb{R}_+\to\mathbb{R}_+$ such that conditions (a), (b) of Definition are satisfied and $\Psi_0(t)\leqq\Psi(t)$ for $t\in\mathbb{R}_+$. Let $C>0$ be the constant such that $\Psi'(t)\leqq C$ for $t\in\mathbb{R}_+$. We define the function $H\colon E_0\cup E\to\mathbb{R}$ in the following way

$$H(x,y)=\exp\Big(\Psi\Big(\sqrt{1+\|y\|^2}\Big)+x\Big(1+K\exp(C\|\tau\|)+C_1\exp(C\|h\|)\sum_{i=1}^n\frac{L_i}{h_i}\Big)\Big)$$
$$\text{for} \quad (x,y)\in E,$$
$$H(x,y)=\exp\Big(\Psi\Big(\sqrt{1+\|y\|^2}\Big)\Big) \quad \text{for} \quad (x,y)\in E_0.$$

It can be shown that H is a solution of the differential-difference inequality

$$D_x z(x,y^m) \geqq z(x,y^m) + K\|z_{(x,m)}\|_{D_h} + \sum_{i=1}^{n} L_i|\delta_i z(x,y^m)|, \quad (x,y^m)\in(0,a]\times\mathbb{R}^n_h. \tag{9}$$

Let $\varepsilon > 0$ be fixed. It follows that there exists $A > 0$ such that

$$\frac{v(x,y)}{H(x,y)} < \varepsilon \quad \text{for} \quad (x,y)\in(0,a]\times\mathbb{R}^n_h, \quad \|y\|\geqq A. \tag{10}$$

Write

$$U(x,y) = \frac{u(x,y)}{H(x,y)}, \qquad V(x,y) = \frac{v(x,y)}{H(x,y)},$$
$$\Omega(x,y) = \frac{\omega(x;\eta)}{H(x,y)},$$

where $(x,y)\in E_{0.h}\cup E_h$. We prove that

$$V(x,y) < \varepsilon \quad \text{for} \quad (x,y)\in E_{0.h}\cup E_h, \quad \|y\| < A. \tag{11}$$

The initial inequality (8) implies $V(x,y)\leqq 0$ for $(x,y)\in E_{0.h}$. Suppose that assertion (11) is false. Then there exists $\tilde{x}\in(0,a]$ and $\tilde{m}\in\mathbb{Z}^n$ such that $\|y^{\tilde{m}}\| < A$ and

$$V(x,y^m) < \varepsilon \quad \text{for } x\in[0,\tilde{x}),\ m\in\mathbb{Z}^n, \tag{12}$$

$$V(\tilde{x},y^{\tilde{m}}) = \max\{V(\tilde{x},y^m) : m\in\mathbb{Z}^n\} = \varepsilon. \tag{13}$$

Using the definition of function v and (6) we get

$$D_x v(\tilde{x},y^{\tilde{m}}) + \sigma(\tilde{x},\omega_{(\tilde{x})}) \leqq \sigma(\tilde{x},\Gamma_h u_{(\tilde{x},\tilde{m})}) + \sum_{i=1}^{n}\lambda_i(\tilde{x},y^{\tilde{m}},u_{(\tilde{x},\tilde{m})})\delta_i u(\tilde{x},y^{\tilde{m}}).$$

We have $\delta v(\tilde{x},y^{\tilde{m}})$, therefore from (3), (13) and from assumptions 3°, 4° it follows that

$$\begin{aligned}
&\sum_{i=1}^{n}\lambda_i(\tilde{x},y^{\tilde{m}},u_{(\tilde{x},\tilde{m})})\delta_i u(\tilde{x},y^{\tilde{m}}) \\
&= \sum_{i=1}^{n}\lambda_i(\tilde{x},y^{\tilde{m}},u_{(\tilde{x},\tilde{m})})h_i^{-1}\sum_{r\in S}c^{(i)}_{r,\tilde{m}}v^{(\tilde{m}+r)}(\tilde{x}) \\
&= \sum_{i=1}^{n}\sum_{r\in S}h_i^{-1}c^{(i)}_{r,\tilde{m}}\lambda_i(\tilde{x},y^{\tilde{m}},u_{(\tilde{x},\tilde{m})})V(\tilde{x},y^{(\tilde{m}+r)})H(\tilde{x},y^{\tilde{m}+r})
\end{aligned}$$

$$\leqq V(\tilde{x}, y^{\tilde{m}}) \sum_{i=1}^{n} |\lambda_i(\tilde{x}, y^{\tilde{m}}, u_{(\tilde{x},\tilde{m})})| \, |\delta_i H(\tilde{x}, y^{\tilde{m}})|$$

$$\leqq V(\tilde{x}, y^{\tilde{m}}) \sum_{i=1}^{n} L_i |\delta_i H(\tilde{x}, y^{\tilde{m}})|.$$

Hence

$$D_x v(\tilde{x}, y^{\tilde{m}}) + \sigma(\tilde{x}, \omega_{(\tilde{x})}) \leqq \sigma(\tilde{x}, \Gamma_h u_{(\tilde{x},\tilde{m})}) + V(\tilde{x}, y^{\tilde{m}}) \sum_{i=1}^{n} L_i |\delta_i H(\tilde{x}, y^{\tilde{m}})|. \tag{14}$$

Let $M_{\{\tilde{x},\tilde{m}\}}\colon D_h \to \mathbb{R}$ defined by

$$M_{\{\tilde{x},\tilde{m}\}}(t, s) = \max\{\Gamma_h u_{(\tilde{x},\tilde{y})}(t, s), \omega_{(\tilde{x})}(t)\}, \quad (t, s) \in D_h.$$

It is easy to verify that

$$\|M_{\{\tilde{x},\tilde{m}\}} - \omega_{(\tilde{x})}\|_0 \leqq V(\tilde{x}, y^{\tilde{m}}) \|H_{(\tilde{x},y^{\tilde{m}})}\|_D. \tag{15}$$

Since

$$\begin{aligned} V(\tilde{x}, y^{\tilde{m}}) > 0, &\qquad D_x V(\tilde{x}, y^{\tilde{m}}) \geqq 0, \\ H(\tilde{x}, y^{\tilde{m}}) > 0, &\qquad D_x H(\tilde{x}, y^{\tilde{m}}) > 0, \end{aligned}$$

and the function σ is nondecreasing with respect to the functional argument, it follows from (14), (15) and (7) that

$$\begin{aligned} 0 \leqq D_x v(\tilde{x}, y^{\tilde{m}}) &= D_x V(\tilde{x}, y^{\tilde{m}}) H(\tilde{x}, y^{\tilde{m}}) + V(\tilde{x}, y^{\tilde{m}}) D_x H(\tilde{x}, y^{\tilde{m}}) \\ &\leqq \sigma(\tilde{x}, M_{\{\tilde{x},\tilde{m}\}}) - \sigma(\tilde{x}, \omega_{(\tilde{x})}) + V(\tilde{x}, y^{\tilde{m}}) \sum_{i=1}^{n} L_i |\delta_i H(\tilde{x}, y^{\tilde{m}})| \\ &\leqq K V(\tilde{x}, y^{\tilde{m}}) \|H_{(\tilde{x},y^{\tilde{m}})}\|_D + V(\tilde{x}, y^{\tilde{m}}) \sum_{i=1}^{n} L_i |\delta_i H(\tilde{x}, y^{\tilde{m}})|. \end{aligned}$$

Therefore

$$\begin{aligned} &0 < D_x V(\tilde{x}, y^{\tilde{m}}) H(\tilde{x}, y^{\tilde{m}}) \\ &\leqq V(\tilde{x}, y^{\tilde{m}}) \Big[K \|H_{(\tilde{x},y^{\tilde{m}})}\|_D + \sum_{i=1}^{n} L_i |\delta_i H(\tilde{x}, y^{\tilde{m}})| - D_x H(\tilde{x}, y^{\tilde{m}}) \Big]. \end{aligned} \tag{16}$$

(16) and (9) give a contradiction. Then (11) is proved. It follows from (10), (11) that $V(x, y^m) < \varepsilon$ for $(x, y^m) \in E_h$ where $\varepsilon > 0$ is arbitrary. Then we have $u(x, y^m) \leqq \omega(x; \eta)$ on E_h. The proof of the inequality $-u(x, y^m) \leqq \omega(x; \eta)$, $(x, y^m) \in E_h$, is analogous. The proof of the theorem is complete.

4. Stability and existence theorem

In this section we prove that there exists a solution of problem (4), (5) and that the method of lines is stable. Denote by $u_h\colon E_{0.h}\cup E_h\to\mathbb{R}$ a solution of problem (4), (5) and suppose that $v_h\colon E_{0.h}\cup E_h\to\mathbb{R}$. Assume that u_h and v_h are of class Σ and that the derivatives $D_x v_h(x,y^m)$ exist for $(x,y^m)\in(0,a]\times\mathbb{R}^n_h$. Suppose that there are $\gamma_0,\gamma_1:\mathbb{R}_+\to\mathbb{R}_+$ such that $\lim_{t\to 0}\gamma_0(t)=\lim_{t\to 0}\gamma_1(t)=0$ and

$$|v_h(x,y^m)-\phi_h(x,y^m)|\leqq\gamma_0(\|h\|),\quad (x,y^m)\in E_{0.h},$$
$$|D_x v_h(x,y^m)-\Phi_h(x,y^m,(v_h)_{(x,m)},\delta v_h(x,y^m))|\leqq\gamma_1(\|h\|),$$
$$(x,y^m)\in(0,a]\times\mathbb{R}^n_h.$$

The method of lines (4), (5) is called stable if there exists $\omega\colon\mathbb{R}_+\to\mathbb{R}_+$ such that $\lim_{t\to 0}\omega(t)=0$ and

$$|u_h(x,y^m)-v_h(x,y^m)|\leqq\omega(\|h\|),\quad (x,y^m)\in E_h.$$

THEOREM 2. *Suppose that*

1° *the function Φ_h characterizing the method of lines satisfies the Lipschitz condition*

$$|\Phi_h(x,y,w,q)-\Phi_h(x,y,\bar{w},q)|\leqq K\|w-\bar{w}\|_{D_h}\tag{17}$$

for every $(x,y)\in E_h$, $w,\bar{w}\in C(D_h,\mathbb{R})$, $q\in\mathbb{R}^n$, where $K>0$,

2° *there exist partial derivatives $D_q\Phi_h=(D_{q_1}\Phi_h,\dots,D_{q_n}\Phi_h)$ on Θ_h and for every $(x,y^m,w,q)\in\Theta_h$, $r\in S-\{0\}$, $i=1,\dots,n$, the inequalities*

$$D_{q_i}\Phi_h(x,y^m,w,q)c^{(i)}_{r,m}\geqq 0,\tag{18}$$

$$|D_{q_i}\Phi_h(x,y^m,w,q)|\leqq L_i,\tag{19}$$

are satisfied, where $L_i\geqq 0$,

3° *there exists C_1 such that $\sum_{r\in S}|c^{(i)}_{r,m}|\leqq C_1$ for $i=1,\dots,n$, $m\in\mathbb{Z}^n$,*

4° *$\phi_h\colon E_{0.h}\to\mathbb{R}$, $h\in I_0$, is of class Σ.*

Then there exists a solution $u_h\colon E_{0.h}\cup E_h\to\mathbb{R}$ of the problem (4), (5) *belonging to class Σ and the method of lines is stable.*

PROOF. Let $h\in I_0$ be fixed. We shall show the existence of the solution of the problems (4), (5). Let X be the set of all real sequences $\xi=\{\xi^m\}_{m\in\mathbb{Z}^n}$, $\xi^m\in\mathbb{R}$, such that $\sup\{|\xi^m|\colon m\in\mathbb{Z}^n\}<\infty$. X is a Banach space if we define the norm

$$\|\xi\|_X=\sup\{|\xi^m|:m\in\mathbb{Z}^n\}.$$

Then $C([-\tau_0,0],X)$ is also a Banach space with the norm

$$\|w\|_{CX}=\max\{\|w(t)\|_X : t\in[-\tau_0,0]\}$$

for $w\in C([-\tau_0,0],X)$. Define $A_m\colon C([-\tau_0,0],X)\to C(D_h,\mathbb{R})$ for $m\in\mathbb{Z}^n$ in the following way:

$$[A_m w](t,y^s)=w^{m+s}(t),$$

where $w\in C([-\tau_0,0],X)$, $(t,y^s)\in D_h$. Let $\tilde{\phi}_h\colon E_{0.h}\cup E_h\to\mathbb{R}$ be defined by $\tilde{\phi}(x,s)=\phi_h(x,s)$ for $(x,s)\in E_{0.h}$ and $\tilde{\phi}_h(x,s)=\phi_h(0,s)$ for $(x,s)\in E_h$. We also define the function $g_h\colon[0,a]\times C([-\tau_0,0],X)\to X$ by

$$g_h^m(x,w)=\Phi_h(x,y^m,A_m w+(\tilde{\phi}_h)_{(x,m)},\delta\omega(0,y^m)+\delta\tilde{\phi}_h(x,y^m)),$$

where $x\in[0,a]$, $w\in C([-\tau_0,0],X)$, $m\in\mathbb{Z}^n$.

Consider the following problem

$$\zeta'(x)=g_h(x,\zeta_{(x)}),\quad x\in[0,a], \tag{20}$$

$$\zeta(x)=0,\qquad x\in[-\tau_0,0], \tag{21}$$

in the Banach space $C([-\tau_0,0],X)$. Let us show that the function g_h satisfies a Lipschitz condition with respect to the functional argument. We have

$$\begin{aligned}
&|g_h^m(x,w)-g_h^m(x,\bar{w})|\\
&\leqq|\Phi_h(x,y^m,A_m w+(\tilde{\phi}_h)_{(x,m)},\delta w(0,y^m)+\delta\tilde{\phi}_h(x,y^m))\\
&\quad-\Phi_h(x,y^m,A_m\bar{w}+(\tilde{\phi}_h)_{(x,m)},\delta\bar{w}(0,y^m)+\delta\tilde{\phi}_h(x,y^m))|\\
&\leqq K\|A_m w-A_m\bar{w}\|_{D_h}+\sum_{i=1}^n L_i|\delta_i(w-\bar{w})(0,y^m)|\\
&\leqq K\max\{|w(t,y^{m+s})-\bar{w}(t,y^{m+s})| : t\in[-\tau_0,0],-N\leqq s\leqq N\}\\
&\quad+C_1\sum_{i=1}^n\frac{L_i}{h_i}\max\{|w(0,y^{m+r})-\bar{w}(0,y^{m+r})| : r\in S\}\\
&\leqq\left[K+C_1\sum_{i=1}^n\frac{L_i}{h_i}\right]\|w-\bar{w}\|_{CX},
\end{aligned}$$

for $x\in[0,a]$, $w,\bar{w}\in C([-\tau_0,0],X)$ and arbitrary $m\in\mathbb{Z}^n$. Therefore

$$\|g_h(x,w)-g_h(x,\bar{w})\|_X\leqq L[h]\|w-\bar{w}\|_{CX},$$

where $L[h]$ is a constant depending on h.

We consider differential-functional problem (20), (21) in the Banach space X. The right-hand side of the equation satisfies the global Lipschitz condition with respect to the functional argument. Hence problem (21), (22) has exactly one solution $\zeta\colon [-\tau_0, a] \to X$. The function $u_h\colon E_{0.h} \cup E_h \to \mathbb{R}$ defined by

$$u_h(x, y^m) = \zeta^m(x) - \tilde{\phi}_h(x, y^m),$$

$x \in [-\tau_0, a]$, $m \in \mathbb{Z}^n$, is a solution of (4), (5) belonging to class Σ.

Now we shall prove the stability of (4), (5). Let us define $\omega\colon [-\tau_0, a] \times \mathbb{R}_+ \to \mathbb{R}_+$ in the following way:

$$\omega(x,t) = \exp(Kx)\left[\frac{1}{K}\gamma_1(t) + \gamma_0(t)\right] - \frac{1}{K}\gamma_1(t), \quad x \in [0, a],$$

$$\omega(x,t) = \gamma_0(t), \quad x \in [-\tau_0, 0].$$

It is easy to verify that the function $\omega(\cdot, \|h\|)$ is a solution of the problem

$$\begin{aligned} \omega'(x) &= K\|\omega_{(x)}\|_0 + \gamma_1(\|h\|), && x \in (0, a], \\ \omega(x) &= \gamma_0(\|h\|), && x \in [-\tau_0, 0] \end{aligned}$$

and $\lim_{h \to 0} \omega(x, \|h\|) = 0$ uniformly with respect to $x \in [-\tau_0, a]$. Since

$$\begin{aligned} &D_x[u_h(x,y^m) - v_h(x,y^m)] \\ &= \Phi_h(x, y^m, (u_h)_{(x,m)}, \delta u_h(x,y^m)) - \Phi_h(x, y^m (u_h)_{(x,m)}, \delta v_h(x,y^m)) \\ &\quad + \Phi_h(x, y^m (u_h)_{(x,m)}, \delta v_h(x,y^m)) - \Phi_h(x, y^m, (v_h)_{(x,m)}, \delta v_h(x,y^m)) \\ &\quad + \Phi_h(x, y^m, (v_h)_{(x,m)}, \delta v_h(x,y^m)) - D_x v_h(x,y^m) \end{aligned}$$

for $x \in (0, a]$, $m \in \mathbb{Z}^n$, it follows that

$$\begin{aligned} &|D_x[u_h(x,y^m) - v_h(x,y^m)] \\ &\quad - \sum_{i=1}^{n} \int_0^1 D_{q_i}\Phi_h(Q_m(x,s))ds\delta_i[u_h(x,y^m) - v_h(x,y^m)]| \\ &\leqq K\|(u_h)_{(x,m)} - (v_h)_{(x,m)}\|_h + \gamma_1(\|h\|) \\ &= K\|\Gamma_h(u_h - v_h)_{(x,m)}\|_0 + \gamma_1(\|h\|) \end{aligned}$$

for $x \in (0, a]$, $m \in \mathbb{Z}^n$, where

$$Q_m(x,s) = (x, y^m, (u_h)_{(x,m)}, \delta v_h(x,y^m) + s(\delta u_h(x,y^m) - \delta v_h(x,y^m))).$$

The function $u_h - v_h$ is of class Σ. Thus in force of Theorem 1 we have

$$|u_h(x,y^m) - v_h(x,y^m)| \leqq \omega(x, \|h\|), \quad x \in [0, a],\ m \in \mathbb{Z}^n,$$

which completes the proof.

5. Convergence theorem

We will need the following

ASSUMPTION H. Suppose that

1° the operator δ satisfies the conditions

$$\sum_{r\in S} c_{r,m}^{(i)} r_j = \delta_{i,j}$$

for $i,j=1,\dots,n$, $m\in\mathbb{Z}^n$, where $\delta_{i,j}$ is the Kronecker symbol,

$$\sum_{r\in S} c_{r,m}^{(i)} = 0$$

for $i=1,\dots,n$, $m\in\mathbb{Z}^n$,

2° there exists $C_2>0$ such that $h_i h_j^{-1}\leqq C_2$, $i,j=1,\dots,n$.

REMARK. Suppose that Assumption H and Condition 3° of Theorem 2 are satisfied. If $\tilde{z}\colon E\to\mathbb{R}$ is of class C^2 and $|D_{y_iy_j}\tilde{z}(x,y)|\leqq C_2$, $(x,y)\in E$, $1\leqq i,j\leqq n$, then there is $\bar{C}\geqq 0$ such that

$$\|\delta\tilde{z}(x,y^m)-D_y\tilde{z}(x,y^m)\|\leqq\bar{C}\|h\|,\quad (x,y^m)\in E_h. \tag{22}$$

We omit a simple proof of (22).

THEOREM 3. *Suppose that assumptions 1°–4° of Theorem 2 are satisfied and*

1° $v\in C^2(E,\mathbb{R})$ *is a solution of the problem* (1), (2) *satisfying the inequalities* $|D_{y_iy_j}v(x,y)|\leqq C_3$ *on* E *for every* $i,j=1,\dots,n$ *and a certain constant* $C_3>0$,

2° *the function* $v_h\colon E_{0.h}\cup E_h\to\mathbb{R}$, *defined by* $v_h=v|_{E_{0.h}\cup E_h}$ *is of class* Σ,

3° $f\in C(\Omega,\mathbb{R})$ *and* $|D_{q_i}f(Q)|\leqq C_4$ *for every* $Q\in\Omega$,

4° *there exist functions* $\chi_i\colon\mathbb{R}_+\to\mathbb{R}_+$, $i=0,1$, *such that* $\lim\limits_{\alpha\to 0+}\chi_i(\alpha)$ *and*

$$|\phi(x,y^m)-\phi_h(x,y^m)|\leqq\chi_0(\|h\|),\quad (x,y^m)\in E_{0.h}, \tag{23}$$

$$\begin{aligned}&|f(x,y^m,v_{(x,y^m)},\delta v(x,y^m))-\Phi_h(x,y^m,(v_h)_{(x,m)},\delta v_h(x,y^m))|\\&\leqq\chi_1(\|h\|),\quad (x,y^m)\in E_h.\end{aligned} \tag{24}$$

Then there exist $\bar{a}>0$ *and* $\vartheta\colon[0,\bar{a}]\to\mathbb{R}_+$ *such that*

$$\lim_{\alpha\to 0+}\vartheta(\alpha)=0 \quad\textit{and}\quad |u_h(x,y^m)-v_h(x,y^m)|\leqq\vartheta(\|h\|)$$

for $\|h\|\in(0,\bar{a}]$, $x\in[0,a]$, $m\in\mathbb{Z}^n$, *where* $u_h\in\Sigma$ *is a solution of the problem* (4), (5).

PROOF. We will apply Theorem 2. From (23) we have

$$|v_h(x,y^m)-\phi_h(x,y^m)|\leqq\chi_0(\|h\|)\quad\text{for } (x,y^m)\in E_{0.h}.$$

There exists $\bar{C}\in\mathbb{R}_+$ such that (22) holds. Thus from assumption 6° and (23) there exists a function $\chi\colon \mathbb{R}_+\to\mathbb{R}_+$ such that $\lim_{\alpha\to 0+}\chi(\alpha)=0$ and

$$\begin{aligned}
&|D_x v_h(x,y^m)-\Phi_h(x,y^m,(v_h)_{(x,m)},\delta v_h(x,y^m))|\\
&\leqq |f(x,y^m,v_{(x,y^m)},D_y v(x,y^m))-f(x,y^m,v_{(x,y^m)},\delta v(x,y^m))|\\
&\qquad+|f(x,y^m,v_{(x,y^m)},\delta v(x,y^m))-\Phi_h(x,y^m,(v_h)_{(x,m)},\delta v_h(x,y^m))|\\
&\leqq \chi(\|h\|)+\chi_1(\|h\|)\quad \text{for } x\in(0,a],\ m\in\mathbb{Z}^n.
\end{aligned}$$

The assumptions of the Stability Theorem are satisfied by

$$\gamma_0=\chi_0,\qquad \gamma_1=\chi+\chi_1.$$

Taking $\vartheta(\|h\|)=\omega(a,\|h\|)$ we have the assertion of the Convergence Theorem.

6. Examples of the method of lines

Let $F(X,Y)$ denote the set of all functions mapping X into Y, where X,Y are arbitrary sets. We will denote by $T_h\colon F(E_{0.h}\cup E_h,\mathbb{R})\to F(E_0\cup E,\mathbb{R})$ the approximating operator defined in the following way. Put $S_+=\{0,1\}^n$. If $(x,y)\in E_0\cup E$ then there exists $m\in\mathbb{Z}^n$ such that $y^m\leqq y\leqq y^{m+1}$, where $m+1=(m_1+1,\dots,m_n+1)$. For $w\in F(E_{0.h}\cup E_h,\mathbb{R})$ we set

$$[T_h w](x,y)=\sum_{r\in S_+} w(x,y^{m+r})\Big[\frac{y-y^m}{h}\Big]^r\Big[1-\frac{y-y^m}{h}\Big]^{1-r},$$

where

$$\begin{aligned}
\Big[\frac{y-y^m}{h}\Big]^r&=\prod_{i=1}^n\Big[\frac{y_i-y_i^{m_i}}{h_i}\Big]^{r_i},\\
\Big[1-\frac{y-y^m}{h}\Big]^{1-r}&=\prod_{i=1}^n\Big[1-\frac{y_i-y_i^{m_i}}{h_i}\Big]^{1-r_i}
\end{aligned}\tag{25}$$

and we put $0^0=1$ in (25).

EXAMPLE 1. Let us consider the problem (1), (2) with the continuous function $f\colon\Omega\to\mathbb{R}$ satisfying the Lipschitz condition with respect to the functional argument and condition 6° of Theorem 3. Let $k_0\in\mathbb{N}$ satisfies the inequalities $1\leqq k_0\leqq n$ and let us assume that

$$\begin{aligned}
D_{q_i}f(Q)&\leqq 0,\quad Q\in\Omega,\ i=1,\dots,k_0,\\
D_{q_i}f(Q)&\geqq 0,\quad Q\in\Omega,\ i=k_0+1,\dots,n.
\end{aligned}$$

We denote $i(m) = m + e_i$, $-i(m) = m - e_i$, where $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ and 1 is standing on the i-th place $i = 1, \dots, n$. Define

$$\delta_i z(x, y^m) = h_i^{-1}(z(x, y^m) - z(x, y^{-i(m)})), \quad i = 1, \dots, k_0,$$
$$\delta_i z(x, y^m) = h_i^{-1}(z(x, y^{i(m)}) - z(x, y^m)), \quad i = k_0 + 1, \dots, n$$

for $x \in [0, a]$, $m \in \mathbb{Z}^n$. Consider the method of lines (4), (5) with the function $\Phi_h : \Theta_h \to \mathbb{R}$, $h \in I_0$, defined by

$$\Phi_h(x, y^m, z_{(x,m)}, q) = f(x, y^m, [T_h z]_{(x,y^m)}, q), \tag{26}$$

where $(x, y^m) \in E_h$, $z \in C(E_{0.h} \cup E_h, \mathbb{R})$, $q \in \mathbb{R}^n$. The function Φ_h and the difference operator δ satisfy (18).

We can define the difference operator in different ways. For example

$$\delta_i z(x, y^m) = h_i^{-1}(2z(x, y^{i(m)}) - 3z(x, y^m) + z(x, y^{-i(m)})),$$

$i = 1, \dots, n$, $x \in [0, a]$, $m \in \mathbb{Z}^n$. If we assume that $D_{q_i} f(Q) \geqq 0$ for $Q \in \Omega$, $i = 1, \dots, n$, then the above operator and the function Φ_h satisfy (18).

EXAMPLE 2. Consider the differential-integral equation

$$D_x z(x, y^m) = f\left(x, y, \int_{D'} z_{(x,y)}(s)ds, \delta z(x, y^m)\right), \quad (x, y) \in E, \tag{27}$$

with the initial condition

$$z(0, y) = \phi(y), \quad y \in \mathbb{R}^n, \tag{28}$$

where $D' = [-\tau, \tau]$ and $z_{(x,y)} : D' \to \mathbb{R}$ is given by

$$z_{(x,y)}(s) = z(x, y + s), \quad s \in D'.$$

One of the method of lines for Cauchy problem (27), (28) is the following

$$D_x z(x, y^m) = f\left(x, y^m, \int_{D'} [T_h z]_{(x,y^m)}(s)ds, \delta z(x, y^m)\right),$$

$x \in [0, a]$, $m \in \mathbb{Z}^n$,

$$z(0, y^m) = \phi(y^m), \quad m \in \mathbb{Z}^n.$$

We have obtained the Cauchy problem for infinite system of ordinary differential equations with

$$\int_{D'} [T_h z]_{(x,y^m)}(s)ds = \frac{1}{2^n} \prod_{i=1}^{n} h_i \sum_{l=-N}^{N-1} \sum_{r \in S_+} z(x, y^{m+l+r}),$$

where $N-1=(N_1-1,\dots,N_n-1)$.

EXAMPLE 3. Let $g\colon E\times\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}$, $\alpha\colon E\to[0,a]$, $\beta\colon E\to\mathbb{R}^n$ be given functions. Define $f(x,y,w,q)=g(x,y,w[\alpha(x,y)-x,\beta(x,y)-y],q)$. Equation (1) reduces to the differential equation with a deviated argument

$$D_x z(x,y)=g(x,y,z(\alpha(x,y),\beta(x,y)),D_y z(x,y)).$$

If we set

$$\Phi_h(x,y^m,z_{(x,m)},q)=g(x,y^m,[T_h z](\alpha(x,y^m),\beta(x,y^m)),q)$$

then we obtain (4).

EXAMPLE 4. Now we present a numerical example. Let us consider the following Cauchy problem

$$\begin{aligned} D_x z(x,y)=\frac{xy}{1+y^2}D_y z(x,y)-x(1+y^2)\int_{-\frac{y}{1+y^2}}^{0} z(x,y+s)\,ds \\ +xy\cdot z\Big(x,y-\frac{xy}{1+y^2}\Big)+F(x,y),\end{aligned}\tag{29}$$

$(x,y)\in[0,1]\times\mathbb{R}$,

$$z(0,y)=1+y,\quad y\in\mathbb{R},\tag{30}$$

where $F(x,y)=1+y-xy(1+x)(1+\frac{y}{2}-xy)(1+y^2)^{-1}$. The solution of the problem is $v(x,y)=1+x+y+xy$, $x\in[0,1]$, $y\in\mathbb{R}$. After applying the method of lines to the problem (29), (30) we get a Cauchy problem with a system of ordinary differential equations. If we apply the Euler method to the system then we get the following difference method:

$$z^{i+1,j}=z^{i,j}+h_0\Big\{\frac{x^i y^j}{1+(y^j)^2}\delta z^{i,j}-x^i(1+(y^j)^2)I^{i,j}+x^i y^j\cdot z^{i,k(i,j)}+F^{i,j}\Big\},\tag{31}$$

$i=0,\dots,N_0$, $j\in\mathbb{Z}$,

$$z^{0,j}=1+y^j,\quad j\in\mathbb{Z},\tag{32}$$

where we use the following notation: $x^i=ih_0$, $y^j=jh$, $z^{i,j}=z(x^i,y^j)$, $F^{i,j}=F(x^i,y^j)$, $h_0>0$ is a step of the time mesh and $h>0$ is a step of the spatial mesh, δ is the difference operator defined in the following way

$$\delta z^{i,j}=h^{-1}(z^{i,j}-z^{i,j-1}),\quad i=0,\dots,N_0,\ j\in\mathbb{Z},\ j>0,$$
$$\delta z^{i,j}=h^{-1}(z^{i,j+1}-z^{i,j}),\quad i=0,\dots,N_0,\ j\in\mathbb{Z},\ j\leqq 0,$$

and $I^{i,j}$ is the integral $\int_{-\frac{y^j}{1+(y^j)^2}}^{0} z(x^i, y^j+s)ds$ counted by using the following complex trapezium method

$$I^{i,j} = \frac{h}{2} \sum_{l=0}^{M(j)-1} (z(x^i, y^{j+l+1}) + z(x^i, y^{j+l})),$$

where $M(j)$ is the integer number such that $y^{M(j)-1} < -\frac{y^j}{1+(y^j)^2} \leqq y^{M(j)}$ and $k(i,j)$ is the integer number such that $y^{k(i,j)-1} < y^j - \frac{x^i y^j}{1+(y^j)^2} \leqq y^{k(i,j)}$. The natural number N_0 satisfies the condition $N_0 h_0 = 1$.

Denote by $u_{h_0 h}$ the solution of (31), (32) and define

$$e^j = v(0.5, y^j) - u_{h_0 h}(0.5, y^j), \quad \tilde{e}^j = v(1.0, y^j) - u_{h_0 h}(1.0, y^j), \quad j \in \mathbb{Z}.$$

Some values of the errors $e, \tilde{e}$ for the steps $h_0 = h = 0.001$ are listed in the tables.

Table of errors for x = 0.5

y^j:	−0.35	−0.2	−0.05	0.05	0.2	0.35
e^j:	−0.00688	−0.00265	−0.00049	−0.00013	−0.00241	−0.00680

Table of errors for x = 1.0

y^j:	−0.35	−0.2	−0.05	0.05	0.2	0.35
$\tilde{e}^j$:	−0.08248	−0.03247	−0.00388	−0.00202	−0.03316	−0.08950

The computation was performed by the computer IBM AT.

REFERENCES

[1] Bassanini, P. and Galaverni, M., Contrazioni multiple, sistemi iperbolici e problema del laser, *Atti Sem. Mat. Fis. Univ. Modena* **31** (1982), 32–50. *MR* **84m**:78034

[2] Brandi, P., Sul problema di Cauchy per une equazione differentiale non lineare alle derivate parziali del primo ordine con argomento funzionale, *Atti Sem. Mat. Fis. Univ. Modena* **28** (1979), 86–101. *MR* **81k**:35004

[3] Brandi, P. and Ceppitelli, R., On the existence of the solution of nonlinear functional partial differential equations of the first order, *Atti Sem. Mat. Fis. Univ. Modena* **29** (1980), 166–186. *MR* **82m**:35021

[4] Brandi, P. and Ceppitelli, R., Existence, uniqueness and continuous dependence for a hereditary non-linear functional partial differential equation of the first order, *Ann. Polon. Math.* **47** (1986), 121–136. *MR* **88e**:35180

[5] Brandi, P., Kamont, Z. and Salvadori, A., Approximate solutions of mixed problems for first order partial differential-functional equations, *Atti Sem. Mat. Fis. Univ. Modena* **39** (1991), 277–302. *MR* **92e**:35165

[6] Czyżewska–Ważewska, M., Lasota, A. and Mackey, M. C., Minimizing therapeutically induced anemia, *J. Math. Biol.* **13** (1981/82), 149–158. *MR* **83f**:92014

[7] DIEKMANN, O., HEIJMANS, H. J. and THIEME, H. R., On the stability of the cell-size distribution, *J. Math. Biol.* **19** (1984), 227–248. *MR* **86a**:92025

[8] EICHHORN, W. and GLEISSNER, W., On a functional differential equation arising in the theory of the distribution of wealth, *Aequationes Math.* **28** (1985), 190–198. *MR* **86g**:35196

[9] GASTALDI, L. and TOMARELLI, F., A nonlinear and nonlocal evolution equation describing the muscle contraction, *Nonlinear Anal.* **11** (1987), 163–182. *MR* **88b**:92013

[10] GURTIN, M. E. and MURPHY, L. F., On the optimal harvesting of persistent age-structured populations, *J. Math. Biol.* **13** (1981/82), 131–148. *MR* **84g**:92030

[11] GYŐRI, I., The method of lines for the solutions of some nonlinear partial differential equations, *Comput. Math. Appl.* **15** (1988), 635–658. *MR* **90h**:65194

[12] JARUSZEWSKA-WALCZAK, D., Existence of solutions of first order partial differential-functional equations, *Boll. Un. Mat. Ital. B* (7) **4** (1990), 57–82. *MR* **91h**:35318

[13] ŁOJASIEWICZ, S., Sur le problème de Cauchy pour les systèmes d'équations aux dérivées partielles du premier ordre dans le cas hyperbolique de deux variables indépendantes, *Ann. Polon. Math.* **3** (1956), 87–117. *MR* **18** #485d

[14] SANZ-SERNA, J. M. and VERWER, J. G., Convergence analysis of one-step schemes in the method of lines, *Appl. Math. Comput.* **31** (1989), 183–196. *MR* **90c**:65107

[15] SCHIESSER, W. E., *The numerical method of lines: Integration of partial differential equations*, Academic Press Inc., San Diego, CA, 1991. *MR* **92f**:65003

[16] VOIGT, A., Line method approximations to the Cauchy problem for nonlinear parabolic differential equations, *Numer. Math.* **23** (1974), 23–36. *MR* **50** #15370

[17] VOIGT, A., The method of lines for nonlinear parabolic differential equations with mixed derivatives, *Numer. Math.* **32** (1979), 197–207. *MR* **80e**:65087

[18] WALTER, W., *Differential and integral inequalities*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 55, Springer-Verlag, Berlin–New York, 1970. *MR* **42** #6391

[19] ACKER, A. and WALTER, W., On the global existence of solutions of parabolic differential equations with a singular nonlinear term, *Nonlinear Anal.* **2** (1978), 499–504. *MR* **80c**:35054

[20] ZUBIK-KOWAL, B., Convergence of the lines method for first-order partial differential-functional equations, *Numer. Methods Partial Differential Equations* **10** (1994), 395–409. *MR* **95g**:65127

(Received March 9, 1995)

INSTYTUT MATEMATYKI
UNIWERSYTET GDAŃSKI
UL. WITA STWOSZA 57
PL–80-952 GDAŃSK
POLAND

zubik@ksinet.univ.gda.pl

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 429–443

GENERALIZED UNIFORM SPACES AND APPLICATIONS TO FUNCTION SPACES

H. RENDER

Abstract

Using non-standard methods we investigate the class of small-set symmetric spaces. The importance of this class stems from the fact that many results being valid for uniform spaces carry over to this larger class, as the Ascoli Theorem, the exponential law and some further results. Examples show that our assumptions cannot be relaxed contradicting some results in the literature.

0. Introduction

One of the most intuitive applications of non-standard analysis is the description of nearness in a uniform space $(X, \mathcal{V})$: for x, y in the non-standard model ${}^*X$ we define y to be *near* to x, more briefly $y \approx x$, to mean that $(x, y) \in \mu := \mu(\mathcal{V}) := \cap_{V \in \mathcal{V}} {}^*V$ and the set $\mu[x] := \{y \in {}^*X : y \approx x\}$ is called the *monad* of x. Obviously, these definitions apply to *any* filter $\mathcal{V}$ on $X \times X$; it is well known that $\mathcal{V}$ is reflexive (symmetric, transitive) iff the relation $\approx$ has the corresponding property, cf. [12]. Thus $\mathcal{V}$ is uniform iff $\approx$ is an equivalence relation on ${}^*X$. But for topological applications it is very often sufficient to know that $\approx$ is only an equivalence relation on the nearstandard points of ${}^*X$. This leads to the concept of a small-set symmetric space which was introduced by N. Vakil as a *Wattenberg infinitesimal.* The importance of this class of generalized uniform spaces stems from the fact that many results being valid for uniform spaces carry over to this larger class, e.g. the Ascoli Theorem and further classical results as we will show in our paper. We give several (standard) characterizations for small-set symmetric spaces and relate our definition to some other concepts like point-symmetry and local symmetry known from the theory of quasi-uniform spaces. The definition of a small-set symmetric space can be used to give a characterization of a certain class of quasi-uniformities answering a question in [1, p. 9]. This is all done in the first section. In the second section we give very elegant and short proofs of results given in [10, 14, 20]. For example, we give a very simple construction of the splitting and jointly continuous topology for a locally bounded space, cf. the results in [10]. Moreover, we show by

1991 *Mathematics Subject Classification.* Primary 54J05; Secondary 54E15, 54C35.

Key words and phrases. Quasi-uniform spaces, small-set symmetric spaces, completeness.

examples that some results in [14] are false and we give the correct form of the corresponding results. In the last section we discuss completeness properties of locally transitive spaces. We show that the non-standard characterization "X complete if and only if $\mathrm{ns}^*X = \mathrm{pns}^*X$" still holds for locally transitive spaces. Furthermore non-standard proofs of some known results are given. Then we answer positively two questions posed by S. Naimpally: there exists a locally symmetric quasi-uniform compact Hausdorff space X such that the set of all (continuous) functions from X to X is not complete for the quasi-uniformity of uniform convergence. It also disproves Theorem 2.10 and 3.6 in [16]. On the other side, the set of all continuous functions is convergence complete with respect to the filter of compact convergence if the domain space is a k-space and the image space is a small-set symmetric, locally symmetric and complete quasi-uniform space. The reader should be familiar with the basic framework of non-standard topology as developed in [4, 12]. We assume a sufficiently saturated non-standard extension of the standard universe containing all the underlying spaces.

1. Small-set symmetric spaces

A filter $\mathcal{V}$ on the set $X \times X$ is *reflexive* if every $V \in \mathcal{V}$ contains the diagonal Δ of $X \times X$. Then $\mathcal{V}$ induces a topology $\tau_{\mathcal{V}}$ calling a set $T \subset X$ open if for every $x \in X$ there exists $V \in \mathcal{V}$ with $V[x] := \{y \in X : (x,y) \in V\} \subset T$. A *locally transitive filter* $\mathcal{V}$ is a reflexive filter such that for every $V \in \mathcal{V}$ and $x \in X$ there exists $W \in \mathcal{V}$ with $W \circ W[x] \subset V[x]$; if W only depends on V, then $\mathcal{V}$ is *transitive*. Thus every quasi-uniform space (cf. [1]) is locally transitive. Locally transitive spaces are also called locally quasi-uniform spaces [11]. It is easy to see that a reflexive filter $\mathcal{V}$ is locally transitive iff $\mathcal{V}$ satisfies the relation

(1) $\quad z \approx y$ and $y \approx {}^*x$ imply $z \approx {}^*x$ for all $z, y \in {}^*X$, $x \in X$.

As usual let $\mathcal{V}^{-1} := \{V^{-1} : V \in \mathcal{V}\}$. If (X, τ) is a topological space, then $m(x) := \cap_{U \in \tau, x \in U} {}^*U$ is the (topological) monad of $x \in X$ and we write $y \approx_\tau x$ for $y \in m(x)$. Moreover $\mathrm{ns}^*X := \cup_{x \in X} m(x)$ is the set of all *nearstandard points*. Further $\mathrm{cpt}^*X := \cup_{K \subset X \text{ compact}} {}^*K$ is called the set of all *compact points*. It is well known that X is compact iff $\mathrm{ns}^*X = {}^*X$ and that X is locally compact iff $\mathrm{ns}^*X = \mathrm{cpt}^*X$. A filter $\mathcal{V}$ on $X \times X$ is *compatible* if $\tau_{\mathcal{V}} = \tau$. Let $(X, \mathcal{U}), (Y, \mathcal{V})$ be reflexive spaces. A function $f \colon X \to Y$ is *uniformly continuous* if for every $V \in \mathcal{V}$ there exists $U \in \mathcal{U}$ with $(f(x), f(y)) \in V$ for all $(x, y) \in U$. Equivalently, this means: $y \approx x$ implies ${}^*f(y) \approx {}^*f(x)$ for all $x, y \in {}^*X$. Finally ${}^\sigma X := \{{}^*x : x \in X\}$ is the copy of X in the nonstandard model ${}^*X$.

Proposition 1.1. *Let $(X, \mathcal{V})$ be a locally transitive space. Then the system of all $V[x]$ with $V \in \mathcal{V}$ is a neighbourhood base of $x \in X$ and $\approx$ is an extension of $\approx_{\tau(\mathcal{V})}$.*

PROOF. Confer the proof of Lemma 1.3 in [24]. □

DEFINITION. Let $\mathcal{V}$ be a reflexive filter. We call $W \in \mathcal{V}$ semisymmetric in $x \in X$ if $(x,y) \in W$ implies $(y,x) \in W$ for all $y \in X$. $\mathcal{V}$ is called *small-set symmetric* if for all $V \in \mathcal{V}$ and $x \in X$ there exists a neighbourhood $U[x]$ with $U \in \mathcal{V}$ and $W \in \mathcal{V}$ being semisymmetric in $x \in X$ such that $W \circ W[y] \subset V[y]$ for all $y \in U[x]$.

In the following theorem we characterize small-set symmetric spaces. The equivalence of (a)–(d) and some further elementary properties are already proved in [22]. If $\mathcal{V}$ is a quasi-uniformity then Theorem 1.2 and Lemma 4 in [8] show that our definition of small-set symmetry coincides with the definition given in [8]. For further characterizations we refer to [2].

THEOREM 1.2. *Let $\mathcal{V}$ be a locally transitive filter. Then the following assertions are equivalent.*

(a) $\approx$ *is an equivalence relation on* $\mathrm{ns}^*X \times \mathrm{ns}^*X$.
(b) $y \approx {}^*x$ *and* $z \approx {}^*x$ *imply* $y \approx z$ *for all* $y, z \in {}^*X$, $x \in X$.
(c) $y \approx {}^*x$ *implies* $\mu[y] = \mu[{}^*x]$.
(d) *Every* $V \in \mathcal{V}$ *is a neighbourhood of* Δ.
(e) $\mathcal{V}$ *is small-set symmetric.*
(f) *There exists a compatible small-set symmetric filter* $\mathcal{U}$ *with* $\mathcal{V} \subset \mathcal{U}$.
If $\mathcal{V}$ is quasi-uniform then in addition are equivalent:
(g) $y \approx {}^*x$ *implies* ${}^*x \approx y$ *for all* $y \in {}^*X$, $x \in X$.
(h) $\tau_{\mathcal{V}^{-1}} \subset \tau_{\mathcal{V}}$.
(i) $\mathcal{V}$ *has a base of closed neighbourhoods of the diagonal.*
(j) *There exists a compatible uniform filter* $\mathcal{U}$ *with* $\mathcal{V} \subset \mathcal{U}$.

PROOF. (a)⇒(b) is trivial and the converse is straightforward using the local transitivity. The equivalence of (b) and (c) is obvious. For the equivalence of (c) and (d) observe that V is a neighbourhood of Δ iff $m(x) \times m(x) \subset {}^*V$ for every $x \in X$.

(a)⇒(e). Choose $W_0 \in {}^*\mathcal{V}$ with $W_0 \subset \mu$ and let $W := W_0 \cup (W_0[{}^*x] \times \{{}^*x\})$. Then $W \in {}^*\mathcal{V}$ and (a) yields $W \subset \mu$. Now let $V \in \mathcal{V}$. It is easy to see that $W \circ W[y] \subset V[y]$ for all $y \approx {}^*x$. Now straightforward arguments (via the transfer principle) yield (e). For (e)⇒(b) let $y \approx {}^*x$ and $z \approx {}^*x$ and $V \in \mathcal{V}$. Choose W, U as in the above Definition. Then $y \in {}^*U[x]$ and therefore ${}^*W \circ {}^*W[y] \subset {}^*V[y]$. Since $z \in {}^*W \circ {}^*W[y]$ we obtain $z \in {}^*V[y]$ for all $V \in \mathcal{V}$, i.e., that $y \approx z$. Obviously, (e)⇒(f) is trivial and for (f)⇒(b) consider the uniformly continuous identity map id: $(X,\mathcal{U}) \to (X,\mathcal{V})$.

(a)⇒(g) is trivial and for (g)⇒(h) observe that $y \approx_{\mathcal{V}} {}^*x \Rightarrow {}^*x \approx_{\mathcal{V}} y \Rightarrow y \approx_{\mathcal{V}^{-1}} {}^*x$.

For (h)⇒(i) apply the results in [1, p. 8] and (i)⇒(d) is trivial. For (h)⇒(j) consider the uniform space $\mathcal{V} \cap \mathcal{V}^{-1}$ generated by the system $\{V \cap V^{-1} : V \in \mathcal{V}\}$ and (j)⇒(f) is obvious. □

Theorem 1.2 (j) shows that a small-set symmetric quasi-uniform space has necessarily a completely regular topology, cf. Theorem 3.8 in [22]. But in general, a small-set symmetric space has only a regular topology; for the proof of regularity one needs only the property that for each $V \in \mathcal{V}$, $x \in X$ there exists $W \in \mathcal{V}$ being semisymmetric in $x \in X$ such that $W \circ W[x] \subset V[x]$, cf. [16, p. 770]. On the other side, let (X, τ) be a regular space; then the filter Δ_τ of all neighbourhoods of the diagonal Δ in $X \times X$ is a compatible small-set symmetric space, cf. Theorem 1.4 in [24] or Proposition 1.9 or Theorem 3.10 in [22].

A reflexive filter $\mathcal{V}$ is called *point-symmetric* [1, p. 36] if for each $V \in \mathcal{V}$ and $x \in X$ there exists a symmetric $W \in \mathcal{V}$ such that $W[x] \subset V[x]$.

PROPOSITION 1.3. *Let $(X, \mathcal{V})$ be a locally transitive space. Then the following assertions are equivalent:*

(a) *X is point-symmetric.*

(b) *${}^*x \approx y$ implies $y \approx {}^*x$ for all $y \in {}^*X$, $x \in X$.*

(c) $\tau_{\mathcal{V}} \subset \tau_{\mathcal{V}^{-1}}$.

PROOF. We prove only (b)⇒(a) since the implications (a)⇒(c)⇒(b) are obvious. Choose $W_0 \in {}^*\mathcal{V}$ with $W_0 \subset \mu$ and let $W := W_0 \cup W_0^{-1}$. It suffices to show that $W[{}^*x] \subset {}^*V[{}^*x]$ by the transfer principle. Let $y \in W[{}^*x]$. Now (b) shows that $y \approx {}^*x$ and therefore $y \in {}^*V[{}^*x]$. □

For a quasi-uniform space $(X, \mathcal{V})$ we have the following duality: $\mathcal{V}$ is small-set symmetric iff $\mathcal{V}^{-1}$ is point-symmetric. Of course, $\mathcal{V}^{-1}$ is in general not small-set symmetric if $\mathcal{V}$ is small-set symmetric, cf. Example 1.6.

COROLLARY 1.4. *Let X be a compact space. Then every compatible small-set symmetric space is uniform.*

PROOF. By compactness we have $\mathrm{ns}^*X = {}^*X$, hence $\approx$ is an equivalence relation on ${}^*X$. □

Similarly one may show that a compact locally transitive filter is transitive. The following result is a slight generalization of Theorem 1 in [9]; as shown in [9] it is not valid for arbitrary quasi-uniform spaces.

THEOREM 1.5. *Every continuous function from a compact locally transitive space into a small-set symmetric space is uniformly continuous.*

PROOF. Let $f : X \to Y$ be continuous and $y \approx x$ with $y, x \in {}^*X$. Since X is compact there exists $x_0 \in X$ with $x \approx {}^*x_0$ and by (1) we have $y \approx {}^*x_0$. The continuity of f yields ${}^*f(x) \approx {}^*f({}^*x_0)$ and ${}^*f(y) \approx {}^*f({}^*x_0)$. Since Y is small-set symmetric we obtain ${}^*f(y) \approx {}^*f(x)$. □

J. Williams has called a filter $\mathcal{V}$ *locally uniform* if $\mathcal{V}$ is symmetric and locally transitive. $\mathcal{V}$ is called an *NLU-space* if in addition for every $V \in \mathcal{V}$, $x \in X$ there exists $W \in \mathcal{V}$ with $W[x] \times W[x] \subset V$. Theorem 1.2 (d)⇒(e) shows that every NLU-space is small-set symmetric. Thus Corollary 1.4 can be seen as a generalization of Theorem 3.7 in [24]. A concept weaker than local uniformity is the following [11]: a reflexive filter $\mathcal{V}$ is *locally symmetric* provided that for every $V \in \mathcal{V}$ and $x \in X$ there exists a symmetric $W \in \mathcal{V}$ with $W \circ W[x] \subset V[x]$. Observe that we do not require the transitivity as in [1]. Example 1.6 shows that a small-set symmetric quasi-uniform space $\mathcal{V}$ need not be locally symmetric, in particular not locally uniform and not an NLU-space.

EXAMPLE 1.6. Let $X := \mathbb{R}$ and $V_n := \{(x,y) \in X \times X : |x-y| < \frac{1}{n}\}$. Consider the filter $\mathcal{V}$ which is induced by the sets $V_{n,x} := (V_n \setminus (\{x\} \times X)) \cup \{(x,x)\}$ with $n \in \mathbb{N}$, $x \in X$. It is easy to see that $y \approx x$ iff $y = x$ for $x \in {}^\sigma X := \{{}^*z : z \in X\}$ and $y \approx_{\mathbb{R}} x$ otherwise where $\approx_{\mathbb{R}}$ is defined by the uniform structure on the real numbers. Thus $(X, \mathcal{V})$ is quasi-uniform; it is small-set symmetric since the induced topology is discrete. But $(X, \mathcal{V})$ is not point-symmetric, in particular not locally symmetric.

PROPOSITION 1.7. *Let $(X, \mathcal{V})$ be a locally transitive space. Then the following assertions are equivalent:*

(a) *X is locally symmetric.*

(b) *$y \approx z$ and $y \approx {}^*x$ imply $z \approx {}^*x$ for all $y, z \in {}^*X$, $x \in X$.*

(c) *$\mu[z] \cap \mu[{}^*x] \neq \emptyset$ implies $\mu[z] \subset \mu[{}^*x]$ for all $z \in {}^*X$, $x \in X$.*

PROOF. (a)⇒(b). Let $V \in \mathcal{V}$ and choose $W \in \mathcal{V}$ symmetric with $W \circ W[x] \subset V[x]$. If $y \approx z$ we have $(z,y) \in {}^*W$ and therefore $(y,z) \in {}^*W$. Since $({}^*x, y) \in {}^*W$ we obtain $z \in {}^*W \circ {}^*W[{}^*x] \subset {}^*V[{}^*x]$. Since $V \in \mathcal{V}$ is arbitrary we have $z \approx {}^*x$. For the converse choose $W_0 \in {}^*\mathcal{V}$ with $W_0 \subset \mu$. Then $W := W_0 \cup W_0^{-1}$ is symmetric. It suffices to show that $W \circ W[{}^*x] \subset {}^*V[{}^*x]$ for every $V \in \mathcal{V}$. For $z \in W \circ W[{}^*x]$ there exists $y \in {}^*X$ with $({}^*x, y), (y,z) \in W$. If $({}^*x, y) \in W_0^{-1}$ then (b) implies $y \approx {}^*x$. Since $z \approx y$ or $y \approx z$ (1) resp. (b) yields $z \approx {}^*x$. The equivalence of (b) and (c) is obvious. □

Proposition 1.7 (b)⇒(a) has the following consequence: *A compact quasi-uniform space $(X, \mathcal{V})$ is locally symmetric iff $\tau_{\mathcal{V}}$ is regular.* Indeed, let $y \approx z$ and $y \approx {}^*x$. Since X is compact there exists $z_0 \in X$ with $z \approx {}^*z_0$. Thus $y \approx {}^*z_0$ and $y \approx {}^*x$. The regularity yields ${}^*z_0 \approx {}^*x$. By (1) we have $z \approx {}^*x$.

COROLLARY 1.8. *Let $(X, \mathcal{V})$ be a locally transitive space. Then the following assertions are equivalent:*

(a) *For each $V \in \mathcal{V}$, $x \in X$ there exists a symmetric $W \in \mathcal{V}$ and a neighbourhood $U[x]$ of x such that $W \circ W[y] \subset V[y]$ for all $y \in U[x]$.*

(b) *X is small-set symmetric and locally symmetric.*

If X is quasi-uniform then in addition are equivalent:

(c) *X is small-set symmetric and point-symmetric.*

(d) *$\tau_{\mathcal{V}} = \tau_{\mathcal{V}^{-1}}$.*

PROOF. (a)⇒(b) is trivial. For the converse choose $W_0 \in {}^*\mathcal{V}$ with $W_0 \subset \mu$ and let $W := W_0 \cup W_0^{-1}$. It is enough to show that $W \circ W[y] \subset {}^*V[y]$ for every $y \approx {}^*x$. For $z \in W \circ W[y]$ there exists $w \in {}^*X$ with $(y,w), (w,z) \in W$. Now the local symmetry shows that $w \approx {}^*x$ and therefore $z \approx {}^*x$. Since X is small-set symmetric and $y \approx {}^*x$ we obtain $z \approx y$, i.e., that $z \in {}^*V[y]$. Now let us show (c)⇒(b). Let $y \approx z$ and $y \approx {}^*x$. Then ${}^*x \approx y$ and therefore ${}^*x \approx z$. Since X is point-symmetric we infer $z \approx {}^*x$. The other implications are clear. □

A small-set symmetric locally symmetric quasi-uniformity is not necessarily symmetric (or equivalently locally uniform): Let $X = \mathbb{R}$ and consider the quasi-uniformity $\mathcal{V}$ generated by the sets $V_n := \Delta \cup \{(x,y) \in \mathbb{R} \times \mathbb{R} : n \leqq y \leqq x\}$ with $n \in \mathbb{N}$. Then $\tau_{\mathcal{V}} = \tau_{\mathcal{V}^{-1}}$ is the discrete topology but $\mathcal{V}$ is not symmetric.

Now we want to give a sufficient criterion for local transitivity and small-set symmetry. Therefore we need the following

DEFINITION. Let $(X, \mathcal{V})$ be a reflexive space. $\mathcal{V}$ is called *full* if $\approx_{\mathcal{V}}$ is an extension of $\approx_{\tau_{\mathcal{V}}}$ and for every $x \in X$ there exists a neighbourhood base $\tau_{\mathcal{V}}(x)$ of x such that for each $G_1, G_2 \in \tau_{\mathcal{V}}(x)$ with $G_1 \subset \overline{G_1} \subset G_2$ the set $(G_2 \times G_2) \cup ((X \setminus G_1) \times (X \setminus G_1))$ is in $\mathcal{V}$.

PROPOSITION 1.9. *Let $(X, \mathcal{V})$ be a full reflexive space. Then the following assertions are equivalent:*

(a) *For every $V \in \mathcal{V}$, $x \in X$ there exists a symmetric $W \in \mathcal{V}$ and a neighbourhood $U[x]$ of $x \in X$ such that $W \circ W[y] \subset V[y]$ for all $y \in U[x]$.*

(b) *$\mathcal{V}$ is small-set symmetric.*

(c) *$\tau_{\mathcal{V}}$ is regular and every $V \in \mathcal{V}$ is a neighbourhood of Δ.*

PROOF. (a)⇒(b)⇒(c) are clear. To prove (c)⇒(a), let $V \in \mathcal{V}$, $x \in X$. Since V is a neighbourhood of $(x,x) \in \Delta$ we can choose $G_3 \in \tau_{\mathcal{V}}(x)$ with $G_3 \times G_3 \subset V$. Since $\tau_{\mathcal{V}}$ is regular there exist $G_i \in \tau_{\mathcal{V}}(x)$ with $G_i \subset G_{i+1}$ for $i = 1,2$. Let $U_i := (G_{i+1} \times G_{i+1}) \cup ((X \setminus G_i) \times (X \setminus G_i)) \in \mathcal{V}$. Then $W := U_1 \cap U_2 \in \mathcal{V}$ is symmetric. If $y \in G_1$ and $z \in W \circ W[y]$ then there exists $r \in X$ with $(y,r) \in W \subset U_1$ and $(r,z) \in W \subset U_2$. Since $y \in G_1$ it follows that $r \in G_2$ and similarly $z \in G_3$. Hence $(y,z) \in G_3 \times G_3 \subset V$. □

Let (X, τ) be a topological space. It is not very difficult to show that the filter Δ_τ of all neighbourhoods of the diagonal Δ in $X \times X$ is full if τ is Hausdorff or regular, or more general, an R_0-space, cf. [1, p. 6]. Proposition 1.9 shows that Δ_τ is small-set symmetric iff τ is regular iff Δ_τ is locally transitive. In [13] it is proved that Δ_τ is quasi-uniform (or uniform) iff τ is almost 2-fully normal. In [1, p. 9] it is asked for which quasi-uniform spaces $\mathcal{V}$ the family of all τ-neighbourhoods of $\cap_{V \in \mathcal{V}} V$ is a compatible quasi-uniformity where τ is the supremum of $\tau_{\mathcal{V}}$ and $\tau_{\mathcal{V}^{-1}}$. Assume that $\tau_{\mathcal{V}}$ is Hausdorff. Then $\Delta = \cap_{V \in \mathcal{V}} V$ and the induced topology of Δ_τ is τ. Thus Δ_τ is compatible iff $\tau_{\mathcal{V}^{-1}} \subset \tau_{\mathcal{V}}$. We infer that among the Hausdorff spaces exactly the small-set symmetric quasi-uniform Hausdorff spaces which are almost 2-fully normal have the above property.

Now let us take a look at the famous Pervin quasi-uniformity $\mathcal{P}$ of a topological space (X, τ) which is by definition the filter generated by the sets $S_T := (T \times T) \cup ((X \setminus T) \times X)$ with $T \in \tau$. Since $\mu(\mathcal{P}) = \cap_{T \in \tau} {}^*S_T$ we obtain that $y \approx_{\mathcal{P}} x$ iff $y \in m(x) := \cap_{T \in \tau, x \in {}^*T} {}^*T$. Since $\approx_{\mathcal{P}}$ is obviously transitive the filter $\mathcal{P}$ is a (compatible) quasi-uniformity. This yields the well-known result that every topological space possesses a compatible quasi-uniformity; this improves Theorem 2.10 in [22]. Observe that the topology of $\mathcal{P}^{-1}$ is discrete if X is a T_1-space: $y \approx_{\mathcal{P}^{-1}} {}^*x$ implies ${}^*x \approx_{\mathcal{P}} y$. Choose $U := X \setminus \{x\}$ being open. If $y \neq {}^*x$ then $y \in {}^*U$ but ${}^*x \notin {}^*U$, a contradiction. Proposition 1.3 (c) shows that $\mathcal{P}$ is point-symmetric; Theorem 1.2 (e), (h) shows that $\mathcal{P}$ is small-set symmetric iff X is discrete. Finally we note that $\mathcal{P}$ is full since $S_{G_2} \cap S_{X\setminus\overline{G_1}} \subset (G_2 \times G_2) \cup ((X \setminus G_1) \times (X \setminus G_1))$.

Let (X, τ) be a topological T_1-space and consider the filter $\mathcal{C}$ generated by the sets $S_{F,U} := (F \times U) \cup ((X \setminus F) \times X)$ with F closed, U open and $F \subset U$. Then $y \approx_{\mathcal{C}} x$ iff $y \in c(x) := \cap_{F\subset U, x\in {}^*F, X\setminus F, U\in\tau} {}^*U$; in [23] $c(x)$ is called the *coarse monad* of $x \in {}^*X$. Obviously, $\mathcal{C}$ is symmetric and it is easy to see that every $S_{F,U}$ is a neighbourhood of Δ with respect to the product topology induced by τ. Thus $\mathcal{C}$ is locally transitive iff $\mathcal{C}$ is locally symmetric iff $\mathcal{C}$ small-set symmetric iff τ is regular, since regularity [normality resp.] obviously implies local transitivity [transitivity resp.] of $\approx_{\mathcal{C}}$. This improves Theorem 3.11 in [22]. Similarly we obtain that τ is normal iff $\mathcal{C}$ is uniform: indeed, if $\mathcal{C}$ is uniform then $c[x] \cap c[y] = \emptyset$ or $c[x] = c[y]$ and Lemma 2.5 (ii)$\Rightarrow$(i) in [23] yields the normality.

2. Applications to function spaces

Let X, Y be topological spaces, $C(X,Y)$ be the space of all continuous functions and τ_k the compact-open topology. For every compatible reflexive filter $\mathcal{V}$ on Y we define the equicontinuity of a family $H \subset C(X,Y)$: for every $x \in X$ and $V \in \mathcal{V}$ there exists a neighbourhood U of x such that $(f(y), f(x)) \in V$ for all $f \in H$, $y \in U$. A well-known non-standard characterization is the following:

(2) $$y \approx {}^*x \text{ implies } f(y) \approx f({}^*x) \text{ for all } f \in {}^*H,\ x \in X.$$

A family H is *pointwise bounded* if every image set $\{f(x) : f \in H\}$ is relatively compact, i.e., that the closure is a compact set. Recall that a *k-space* [*k_3-space* resp.] is a topological space X on which a Y-valued function is continuous if its restriction to each compact subspace is continuous for every topological [regular resp.] space Y.

THEOREM 2.1. *Let X be a k_3-space and Y a small-set symmetric space. Then $H \subset C(X,Y)$ is relatively compact if and only if H is equicontinuous and pointwise bounded.*

In fact, Theorem 2.1 is equivalent to the topological Ascoli-Theorem in [3] if we can show that equicontinuity is equivalent to the (weaker) topological concept of even continuity for a *pointwise bounded* family: Let H be evenly continuous and let $f \in {}^*H$, $x \in X$ and $y \approx {}^*x$. Since H is pointwise bounded there exists $z \in Y$ with $f(^*x) \approx z$. Now the non-standard characterization of even continuity yields $f(y) \approx z$, cf. [4, p. 162]. Thus $f(y) \approx z$ and $f(^*x) \approx z$. Since Y is small-set symmetric we infer $f(y) \approx f(^*x)$. □

The following example shows that Theorem 2.1 is *not* valid for an arbitrary quasi-uniform space $(Y, \mathcal{V})$ even if Y is a compact point-symmetric Hausdorff space. Moreover, it shows that some results in [14] are not correct; the error depends on the false Theorem 4.23 in [15] that has already been pointed out in its review *MR* **35** #2267. In a recent paper H. P. Künzi has given examples showing that in Theorem 2.1, Proposition 2.3 and Theorem 1.5 "Y small-set symmetric" cannot be replaced by "quiet", cf. [7], where also standard proofs of these theorems can be found.

EXAMPLE 2.2. Let $X = Y$ be an abelian topological group with a compact, non-discrete Hausdorff topology, e.g. $X = S^1$. Let $\mathcal{V}$ be the translation (left) invariant uniformity and $\mathcal{P}$ be the Pervin quasi-uniformity on Y. It is easy to see that the set $H := \{\tau_x : x \in X\}$ of all translations (defined by $\tau_x(y) = x + y$) is τ_k-compact. By 2.1 it is $\mathcal{V}$-equicontinuous and evenly continuous. We show that H is not $\mathcal{P}$-equicontinuous: obviously $y \approx {}^*x$ implies $-y \approx -{}^*x$. If H is $\mathcal{P}$-equicontinuous then $\tau_w(-y) \approx \tau_w(-{}^*x)$ for all $w \in {}^*Y$. With $w := y + {}^*x$ we obtain ${}^*x \approx_{\mathcal{P}} y$. By 1.2 (a) and $\mathrm{ns}^*Y = {}^*Y$ we infer that $\mathcal{P}$ is uniform. Then $(Y, \mathcal{P})$ is discrete [1, p. 43], a contradiction. This shows that Theorem 1.1, 1.3 and 1.5 in [14] are not correct (where $f \colon X \times X \to Y$ is defined by $f(x, y) := x + y$).

Let α be a family of subsets of X and let $(Y, \mathcal{V})$ be a reflexive space. Then the system of the sets $W(A, V) := \{(f, g) : (f(x), g(x)) \in V \text{ for all } x \in A\}$ induces a reflexive filter on the set $F(X, Y)$ of all functions $f \colon X \to Y$. Define $\alpha\,\mathrm{pt}\,{}^*X := \cup_{A \in \alpha} {}^*A$; note that $\mathrm{cpt}^*X = k\,\mathrm{pt}\,{}^*X$ for the system k of all compact subsets of X. The following characterization of the relation $\approx_\alpha$ of the induced filter is obvious:

$$f \approx_\alpha g \Leftrightarrow f(x) \approx_{\mathcal{V}} g(x) \quad \text{for all} \quad x \in \alpha\,\mathrm{pt}\,{}^*X. \tag{3}$$

The induced topology $\tau_\alpha(\mathcal{V})$ is called the *topology of $\mathcal{V}$-uniform convergence* on α. But in general the topology $\tau_\alpha(\mathcal{V})$ depends on the filter $\mathcal{V}$ even if α is the set k of all compact sets: It is not very hard to see that the set H in Example 2.2 is not $\tau_k(\mathcal{P})$-compact, thus we have $\tau_k(\mathcal{P}) \neq \tau_k = \tau_k(\mathcal{V})$, cf. Proposition 2.3. Thus Corollary 2.2, 2.4 and 2.5 in [14] are false. Note that $\tau_k \subset \tau_k(\mathcal{V})$ for any locally transitive filter. Proposition 2.4 yields further counterexamples.

In passing we note that the problem when two topologies of uniform convergence agree has a very nice non-standard solution. Let α, β be systems of closed subsets of a completely regular space X. Then it is not very difficult to show that $\tau_\alpha(\mathcal{V}) \subset \tau_\beta(\mathcal{V})$ on $C(X, \mathbb{R})$ if and only if $\alpha\,\mathrm{pt}\,{}^*X \subset \beta\,\mathrm{pt}\,{}^*X$; for details and some more general results see [17].

PROPOSITION 2.3. *Let $(Y,\mathcal{V})$ be a small-set symmetric space. Then $\tau_k = \tau_k(\mathcal{V})$.*

PROOF. The following non-standard characterization of τ_k is valid if X *or* Y is regular *or* Hausdorff (cf. [17]): $f \in {}^*C(X,Y)$ is near to $f_0 \in C(X,Y)$ iff $f(x) \approx f_0(x_0)$ for all $x \in \mathrm{cpt}^*X$, $x_0 \in X$ with $x \approx x_0$. Using (3) the proof is straightforward. □

PROPOSITION 2.4. *Let X be a locally compact Hausdorff space and let $\mathcal{P}$ be the Pervin quasi-uniformity on the Sierpiński space $\{0,1\}$. Then $\tau_k = \tau_k(\mathcal{P})$ on $C(X,\{0,1\})$ if and only if X is discrete.*

PROOF. $\mathcal{P}$ is generated by the set $V := \{(0,0),(1,0),(1,1)\}$. We identify the set $C(X,\{0,1\})$ with the space of all closed subsets of X. Let $x \in X$ and f be the characteristic function of $\{x\}$. It is easy to see that the $\tau_k(\mathcal{P})$-neighbourhoods of f are the sets $\{A : K \cap A \subset \{x\}\}$ where K is an arbitrary compact subset of X. If $(x_i)_i$ is a net in X converging to $x \in X$ then $\{x_i\}$ converges to $\{x\}$ in the compact-open topology. If $\tau_k = \tau_k(\mathcal{P})$ then $\{x_i\} \subset x$ for almost all $i \in I$. Thus X is discrete. The converse is left to the reader.□

The above-mentioned characterization of the compact-open topology on $C(X,Y)$ can be used to give very elegant and short proofs of standard results as the exponential law and the Ascoli Theorem, see [17, 23]. Recall that a topology τ on $C(X,Y)$ is *splitting* if every continuous function $f: T \times X \to Y$ induces a *continuous* function $\widehat{f}: T \to C(X,Y)$ [where $\widehat{f}(t)(x) := f(t,x)$] for *any* topological space T. The topology τ is jointly continuous if the evaluation $e: C(X,Y) \times X \to Y$ defined by $e(f,x) = f(x)$ is continuous. Here we want to illustrate some related results recently given by Lambrinos in [10]. At first we need some definitions: a subset A of a topological space X is *bounded* if every open cover of the whole space X has a finite subcover for the set A. It is easy to see that A is bounded iff ${}^*A \subset \mathrm{ns}^*X$. Thus every relatively compact set is bounded and for regular spaces the converse is also true. X is *locally bounded* if every $x \in X$ has a bounded neighbourhood, or equivalently, $\mathrm{ns}^*X = b\,\mathrm{pt}\,{}^*X$ where b is the system of all bounded subsets.

THEOREM 2.5. *Let X be locally bounded and Y regular. Then $\tau_b(\mathcal{V})$ is the unique jointly continuous and splitting topology on $C(X,Y)$ (and at the same time, the smallest jointly continuous one) where $\mathcal{V}$ is any compatible small-set symmetric filter.*

PROOF. To prove that $\tau_b(\mathcal{V})$ is splitting, let $t \approx t_0$ and $x \in b\,\mathrm{pt}\,{}^*X$. Then there exists $x_0 \in X$ with $x \approx x_0$. Therefore $(t,x) \approx (t_0,x_0)$ and the continuity of f implies ${}^*\widehat{f}(t)(x) = {}^*f(t,x) \approx f(t_0,x_0) = \widehat{f}(t_0)(x_0)$. Now (3) shows that ${}^*\widehat{f}(t) \approx_b \widehat{f}(t_0)$ proving the continuity of $\widehat{f}$. Further it is very easy to check that $\tau_b(\mathcal{V})$ is jointly continuous since $\mathrm{ns}^*X = b\,\mathrm{pt}\,{}^*X$. Finally, a topology τ on $C(X,Y)$ which is splitting and jointly continuous is uniquely determined and weaker than every jointly continuous topology, cf. [17]. □

REMARK 2.6. The regularity of Y is essential in Theorem 2.5. For example, let $X=\mathbb{R}$ and $Y=\{0,1\}$ the Sierpiński space then $\tau_b(\mathcal{V})=\tau_k(\mathcal{V})$ and $\tau_k(\mathcal{V})$ is jointly continuous. But $\tau_k(\mathcal{V})$ is not splitting since otherwise $\tau_k(\mathcal{V})=\tau_k$, a contradiction to Proposition 2.4.

Finally we give very easy non-standard proofs of results in [19, 20].

THEOREM 2.7. *Let X be a point-symmetric locally transitive space and let G be an equicontinuous group of homeomorphisms of X onto X. Then G is a topological group under the topology τ_p of pointwise convergence.*

PROOF. At first we show that the composition $\circ: G\times G\to G$ is continuous. Let $f\approx_p f_0$ and $g\approx_p g_0$ where $f,g\in {}^*G$, $f_0,g_0\in G$ and $\approx_p$ is the monad of τ_p. (3) yields $g({}^*x)\approx g_0(x)$ for any $x\in X$ and (2) shows $f(g({}^*x))\approx f({}^*g_0({}^*x))$. On the other side (3) implies $f({}^*g_0({}^*x))\approx f_0(g_0(x))$ and by (1) we obtain $f(g({}^*x))\approx f_0(g_0(x))$, i.e., $f\circ g\approx_p f_0\circ g_0$. For the continuity of the inversion it suffices to show that $f\approx_p \mathrm{id}$ implies $f^{-1}\approx_p \mathrm{id}$ where id is the identity element. It is easy to see that id is also the identity function on X. Thus $f({}^*x)\approx x$. Now (2) implies ${}^*x=f^{-1}(f({}^*x))\approx f^{-1}({}^*x)$. If Y is point-symmetric we obtain $f^{-1}({}^*x)\approx {}^*x$. The proof is complete. □

REMARK 2.8. Seyedin works in [20] with a topology τ weaker than the induced topology $\tau_{\mathcal{V}}$ of a locally symmetric quasiuniformity $\mathcal{V}$. But the assumptions of Theorem 6 in [20] always imply $\tau=\tau_{\mathcal{V}}$. Let $x\approx_\tau x_0$. Since $\mathrm{id}\in G$ we have by (3) that $x={}^*\mathrm{id}(x)\approx_{\mathcal{V}} {}^*x_0$, i.e., $\tau_{\mathcal{V}}\subset\tau$. Since every locally symmetric space is point-symmetric the main result Theorem 6 in [20] is covered by Theorem 3 in [19].

3. Completeness

A reflexive space $(X,\mathcal{V})$ is *precompact* if for every $V\in\mathcal{V}$ there exist $x_1,\dots,x_n\in X$ with $X=V[x_1]\cup\cdots\cup V[x_n]$. Let $\mathrm{pns}^*X:=\bigcap_{V\in\mathcal{V}}\bigcup_{x\in X}{}^*V[x]$ be the set of all *prenearstandard points.* Easy saturation arguments show that X is precompact if and only if ${}^*X=\mathrm{pns}^*X$. A filter $\mathcal{F}$ on the set X is a *Cauchy filter* if for every $V\in\mathcal{V}$ there exists $x\in X$ with $V[x]\in\mathcal{F}$. Let I be an index set in the standard universe; then ${}^*I_\infty$ denotes the set of all infinitely large $l\in{}^*I$. A net $(x_l)_{l\in I}$ is a Cauchy net if for every $V\in\mathcal{V}$ there exists $x\in X$, $l_0\in I$ such that $x_l\in V[x]$ for all $l\geqq l_0$. X is *complete* [*convergence complete*] if every Cauchy filter has an adherence [limit] point. It is easy to see that $m(\mathcal{F}):=\bigcap_{F\in\mathcal{F}}{}^*F$ is contained in pns^*X for every Cauchy Filter (briefly CF) $\mathcal{F}$. This yields the inclusion part of

$$\bigcup_{\mathcal{F}CF} m(\mathcal{F})=\mathrm{pns}^*X:=\bigcap_{V\in\mathcal{V}}\bigcup_{x\in X}{}^*V[x]. \tag{4}$$

For the converse let $y\in\mathrm{pns}^*X$ and consider the filter subbase $\{V[x]: V\in\mathcal{V},\ x\in X \text{ with } y\in{}^*V[x]\}$. Then the induced filter $\mathcal{F}_y$ is a Cauchy filter:

for $V \in \mathcal{V}$ there exists $x \in X$ with $y \in {}^*V[x]$ (since $y \in \mathrm{pns}^*X$), i.e., that $V[x] \in \mathcal{F}_y$. Since $y \in m(\mathcal{F})$ the proof is complete.

It is well known that a filter $\mathcal{F}$ in a topological space X has the limit [adherence] point $x \in X$ iff $m(\mathcal{F}) \subset m(x)$ [resp. $m(\mathcal{F}) \cap m(x) \neq \emptyset$].

THEOREM 3.1. *Let X be a locally transitive space. Then the following assertions are equivalent:*

(a) *X is complete.*

(b) *$m(\mathcal{F}) \cap \mathrm{ns}^*X \neq \emptyset$ for all Cauchy filters $\mathcal{F}$.*

(c) *$m(\mathcal{F}) \subset \mathrm{ns}^*X$ for all Cauchy filters $\mathcal{F}$.*

(d) $\mathrm{ns}^*X = \mathrm{pns}^*X$.

(e) *For every Cauchy net $(x_l)_{l \in I}$ in ${}^*X$ such that I is in the standard universe there exists $l \in {}^*I_\infty$ with $x_l \in \mathrm{ns}^*X$.*

PROOF. (a)⇒(b) and (d)⇒(c)⇒(a) are clear. Let us prove (b)⇒(d). Let $y \in \mathrm{pns}^*X$. Then the filter $\mathcal{F} := \{A \subset X : y \in {}^*A\}$ is a Cauchy filter. Assume that $y \notin \mathrm{ns}^*X$. For every $x \in X$ there exists $V \in \mathcal{V}$ such that $y \notin {}^*V[x]$. Thus $m(\mathcal{F}) \subset {}^*(X \setminus W[x])$ for some $W \in \mathcal{V}$ with $W \circ W[x] \subset V[x]$ and by Proposition 1.1 $m(\mathcal{F}) \subset {}^*X \setminus m(x)$. Since this holds for all $x \in X$ we obtain a contradiction to (b). The equivalence of (a) and (e) is left to the reader.□

The next example [1, p. 50] shows that completeness and convergence completeness are not the same.

EXAMPLE 3.2. Let $X = [0,1]$ and $V_\epsilon := \Delta \cup [\{0\} \times [0,\epsilon)] \cup [\{1\} \times (1-\epsilon, 1]] \cup [(1/2-\epsilon, 1/2) \times ((0,\epsilon) \cup (1-\epsilon, 1))]$. Then $\mathrm{pns}^*X = \mathrm{ns}^*X = {}^\sigma X \cup m(0) \cup m(1)$. Hence X is a point-symmetric complete space. But the filter $\mathcal{F}$ generated by the sets $(0,\epsilon) \cup (1-\epsilon, 1)$ is a non-convergent Cauchy filter.

If X is a uniform space it is well known that for every Cauchy filter $\mathcal{F}$ there exists $y \in {}^*X$ with $m(\mathcal{F}) \subset \mu[y]$ and y is necessarily in pns^*X. The last example shows that the latter property is not valid for quasi-uniform spaces. We refer to [18] for a discussion of further concepts of completeness via nonstandard techniques.

PROPOSITION 3.3. *Let X be a reflexive space and $\mathcal{F}$ be a filter on X. Then $\mathcal{F}$ is a Cauchy filter iff there exists $F \in {}^*\mathcal{F}$, $y \in {}^*X$ with $F \subset \mu[y]$.*

PROOF. Choose $F \in {}^*\mathcal{F}$ with $F \subset m(\mathcal{F})$ and consider $S_V := \{y \in {}^*X : F \subset {}^*V[y]\}$ for every $V \in \mathcal{V}$. If $\mathcal{F}$ is a Cauchy filter then S_V is non-empty and obviously $S_{V_1} \cap \cdots \cap S_{V_n} \supset S_{V_1 \cap \cdots \cap V_n}$. By saturation there exists $y \in {}^*X$ with $F \subset {}^*V[y]$ for all $V \in \mathcal{V}$. For the converse consider for every $V \in \mathcal{V}$ the following statement: $(\exists y \in {}^*X)(\exists F \in {}^*\mathcal{F})(F \subset {}^*V[y])$. □

We give now non-standard proofs of some classical results proved in [21] and [1].

THEOREM 3.4. *A compact locally transitive space is convergence complete.*

PROOF. Let $\mathcal{F}$ be a Cauchy filter on X, and let $F \in {}^*\mathcal{F}$, $y \in {}^*X$ with $F \subset \mu[y]$. By compactness there exists $x \in X$ with $y \approx {}^*x$. Since $z \approx y$ for all $z \in F$ we infer $F \subset m(x)$ by (1). The transfer principle shows that $\mathcal{F}$ converges to x. □

THEOREM 3.5. *A locally transitive space is compact if and only if it is both convergence complete and precompact.*

PROOF. Let X be compact. By Theorem 3.4 X is convergence complete and precompactness follows from $\mathrm{ns}^*X \subset \mathrm{pns}^*X \subset {}^*X = \mathrm{ns}^*X$. Conversely, we have ${}^*X = \mathrm{pns}^*X$ by precompactness and $\mathrm{pns}^*X \subset \mathrm{ns}^*X$ by completeness. □

PROPOSITION 3.6. *Let X be a locally symmetric space. If a Cauchy filter $\mathcal{F}$ has a cluster point $x \in X$ then $\mathcal{F}$ converges to $x \in X$. In particular, X is convergence complete if and only if it is complete.*

PROOF. Let $F \in {}^*\mathcal{F}$, $y \in {}^*X$ with $F \subset \mu[y]$. Since x is a cluster point we have $F \cap m(x) \neq \emptyset$. Thus $\mu[y] \cap \mu[{}^*x] \neq \emptyset$ and Proposition 1.7 yields $F \subset \mu[y] \subset \mu[{}^*x]$. Thus $\mathcal{F}$ converges to x. □

If $A \subset X$ is a subspace then it is easy to see that $\mathrm{pns}^*A \subset \mathrm{pns}\,{}^*X \cap {}^*A$. The following example [1, p. 48] shows that the inclusion may be proper even if X is complete.

EXAMPLE 3.7. Let $X = \{0\} \cup \{1/n : n \in \mathbb{N}\} \subset \mathbb{R}$ with the induced topology. For each $n \in \mathbb{N}$ define $V_n := \Delta \cup \{(0, 1/k) : k \in \mathbb{N}, k > n\}$. Then the induced filter is a compatible, locally symmetric, quasi-uniform compact Hausdorff space: it is obvious that for $x \in X$, $y \in {}^*X$ we have $y \approx {}^*x$ iff $y = {}^*x$ for $x \neq 0$ and $y \approx_{\mathbb{R}} 0$ for $x = 0$. Hence ${}^{\sigma}X \cup m(0) = \mathrm{ns}^*X = {}^*X$. For $A := \{1/n : n \in \mathbb{N}\}$ we obtain ${}^{\sigma}A = \mathrm{ns}^*A = \mathrm{pns}^*A \neq \mathrm{pns}^*X \cap {}^*A$. In particular, A is complete but not closed in X.

Let X be a set and Y be a complete uniform space. It is a well-known fact that $F(X, Y)$ is complete for the uniformity of uniform convergence. We give now an example in order to show that this result is not valid if Y is only a locally symmetric quasi-uniform (compact) space. This answers questions in [16] and disproves Theorem 2.10 and 3.6 in [16].

EXAMPLE 3.8. Let X as in the previous example and define for each $n \in \mathbb{N}$ continuous functions $g_n \in F(X, X)$ by $g_n(x) = x$ for $x \geq 1/n$ and $g_n(x) = 0$ for $x < 1/n$. Let $\mathcal{F}$ be the filter generated by the sets $F_n := \{g_k : k > n\}$. Since $F_n \subset W(X, V_n)[g_n]$ we infer that $\mathcal{F}$ is a Cauchy filter. Assume now that $F(X, X)$ or $C(X, X)$ is complete. Then there exists $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ with ${}^*g_N \in \mathrm{ns}^*F(X, X)$, i.e., that there exists $g \colon X \to X$ with ${}^*g_N(x) \approx {}^*g(x)$ for all $x \in {}^*X$. Hence $g(x) = x$ for all $x \in X$. Choose $x = 1/(2N)$. Then $g_N(x) = 0 \approx x$, a contradiction.

Observe that $F(X, X)$ and $C(X, X)$ in the last example are not point-symmetric (in particular not locally symmetric) although X is a locally symmetric quasi-uniform space. This is clear since ${}^*g \approx g_N$ but not $g_N \approx {}^*g$, where g is the identity on X.

PROPOSITION 3.9. *Let X be a topological space, $(Y, \mathcal{V})$ be a reflexive filter and let $C(X,Y)$ be endowed with the filter of compact $\mathcal{V}$-uniform convergence. Then the following assertions are true:*

(a) *If Y is small-set symmetric then $C(X,Y)$ is small-set symmetric.*
(b) *If Y is small-set symmetric and locally symmetric then $C(X,Y)$ is small-set symmetric and locally symmetric.*

PROOF. At first we prove that "Y small-set symmetric" implies that $C(X,Y)$ is locally transitive, i.e., satisfies (1): Let $f, g \in {}^*C(X,Y)$, $h \in C(X,Y)$ and let $f \approx g$ and $g \approx {}^*h$. Since $h \in C(X,Y)$ we know that ${}^*h(x) \in \mathrm{ns}^*Y$ for every $x \in \mathrm{cpt}^*X$. Using (3) it is easy to see that there exist $y \in Y$ with $f(x) \approx g(x) \approx {}^*h(x) \approx y$. Hence $f(x) \approx {}^*h(x)$ by 1.2 (a) and (1). Hence $f \approx {}^*h$. Further it is now obvious that $f \approx {}^*h$ and $g \approx {}^*h$ imply $f \approx g$. Similarly one proves property (b) of Theorem 1.2 for $C(X,Y)$. Statement (b) is also straightforward using 1.7. □

THEOREM 3.10. *Let X be a k-space and Y be a small-set symmetric, locally symmetric, complete quasi-uniform space. Then $C(X,Y)$ is convergence complete for the filter of compact $\mathcal{V}$-uniform convergence.*

PROOF. Let $f \in \mathrm{pns}^*C(X,Y)$. Assume that f is *standardizable* and cpt^*X*-continuous*, i.e., that (i) $f({}^*x_0) \in \mathrm{ns}^*Y$ for all $x_0 \in X$ and (ii) $f({}^*x_0) \approx {}^*y$ implies $f(x) \approx {}^*y$ for each $x \in \mathrm{cpt}^*X$, $x_0 \in X$ with $x \approx x_0$ and $y \in Y$. It is not very difficult to see that an internal, standardizable and cpt^*X-continuous function is contained in the set of all nearstandard points of $C(X,Y)$ with respect to $\tau_k = \tau_k(\mathcal{V})$ (Proposition 2.3) if X is a k-space; cf. [17] for details. Hence Theorem 3.1 and Proposition 3.6 show that $C(X,Y)$ is (convergence) complete.

Condition (i) is easily proved since $f({}^*x_0) \in \mathrm{pns}^*Y = \mathrm{ns}^*Y$. Now we prove (ii): let $f({}^*x_0) \approx {}^*y$ and K compact and $x \in {}^*K$, $x_0 \in K$ with $x \approx x_0$. We have to show that $f(x) \in {}^*V[{}^*y]$ for every $V \in \mathcal{V}$. Since Y is locally symmetric there exists a symmetric $V_1 \in \mathcal{V}$ such that $V_1 \circ V_1[y] \subset V[y]$. By transitivity there exists $U \in \mathcal{V}$ with $U \circ U \circ U \subset V_1$. As $f \in \mathrm{pns}^*C(X,Y)$ there exists $g_U \in C(X,Y)$ with $({}^*g_U, f) \in {}^*W(K,U)$ where $W(K,U) := \{(g,h) : \forall x \in K \ \ (g(x), h(x)) \in U\}$. Hence $({}^*g_U(x), f(x)) \in {}^*U$ and $({}^*g_U({}^*x_0), f({}^*x_0)) \in {}^*U$. Since g_U is continuous and $x \approx x_0$ we infer ${}^*g_U(x) \approx {}^*g_U({}^*x_0)$. Small-set symmetry implies ${}^*g_U({}^*x_0) \approx {}^*g_U(x)$ and ${}^*y \approx f({}^*x_0)$. Hence $({}^*g_U(x), {}^*g_U({}^*x_0)) \in {}^*U$ and $({}^*g_U({}^*x_0), f({}^*x_0)) \in {}^*U$ and $(f({}^*x_0), {}^*y) \in {}^*U$. Thus $({}^*g_U(x), {}^*y) \in {}^*V_1$. Since V_1 is symmetric we obtain $({}^*y, {}^*g_U(x)) \in {}^*V_1$. Now $({}^*g_U(x), f(x)) \in {}^*U \subset {}^*V_1$ implies that $({}^*y, f(x)) \in {}^*V_1 \circ {}^*V_1$. Thus $f(x) \in {}^*V_1 \circ {}^*V_1[{}^*y] \subset {}^*V[{}^*y]$. The proof is complete. □

REFERENCES

[1] Fletcher, P. and Lindgren, W. F., *Quasi-uniform spaces*, Lecture Notes in Pure and Applied Mathematics, 77, Marcel Dekker Inc., New York, 1982. *MR* **84h**:54026

[2] Fletcher, P. and Hunsaker, W., Symmetry conditions in terms of open sets, *Topology Appl.* **45** (1992), 39–47. *MR* **93h**:54020

[3] Fox, G. and Morales, P., A non-Hausdorff Ascoli theorem for k_3-spaces, *Proc. Amer. Math. Soc.* **39** (1973), 633–636. *MR* **52** #11826

[4] Hurd, A. E. and Loeb, P. A., *An introduction to nonstandard real analysis*, Pure and Applied Mathematics, 118, Academic Press, Orlando, 1985. *MR* **87d**:03184

[5] Johnson, D. R. and Mattson, D. A., Some applications of non-standard analysis to proximity spaces, *Colloq. Math.* **34** (1975), 17–24. *MR* **53** #9154

[6] Künzi, H.-P. A., Some remarks on quasi-uniform spaces, *Glasgow Math. J.* **31** (1989), 309–320. *MR* **91c**:54035

[7] Künzi, H.-P. A., On quasi-uniform convergence and quiet spaces, *Questions Answers Gen. Topology* **13** (1995), 87–92. *MR* **96e**:54019

[8] Künzi, H.-P., Mršević, M., Reilly, I. L. and Vamanamurthy, M. K., Convergence, precompactness and symmetry in quasi-uniform spaces, *Math. Japon.* **38** (1993), 239–253. *MR* **94b**:54080

[9] Lambrinos, P., A note on quasi-uniform continuity, *Bull. Austral. Math. Soc.* **8** (1973), 389–392. *MR* **47** #9550

[10] Lambrinos, P., The bounded-open topology on function spaces, *Manuscripta Math.* **36** (1981), 47–66. *MR* **83m**:54025

[11] Lindgren, W. F. and Fletcher, P., Locally quasi-uniform spaces with countable bases, *Duke Math. J.* **41** (1974), 231–240. *MR* **49** #6173

[12] Machover, M. and Hirschfeld, J., *Lectures on non-standard analysis*, Lecture Notes in Mathematics, Vol. 94, Springer-Verlag, Berlin – New York, 1969. *MR* **40** #2531

[13] Mansfield, M. J., Some generalizations of full normality, *Trans. Amer. Math. Soc.* **86** (1957), 489–505. *MR* **20** #273

[14] Morales, P., Non-Hausdorff Ascoli theory, *Dissertationes Math. (Rozprawy Mat.)* **119** (1974), 1–37. *MR* **53** #3996

[15] Murdeshwar, M. G. and Naimpally, S. A., *Quasi-uniform topological spaces*, P. Noordhoff Ltd., Groningen, 1966. *MR* **35** #2267

[16] Naimpally, S. A., Function spaces of quasi-uniform spaces, *Nederl. Akad. Wetensch. Proc. Ser. A* **68** = *Indag. Math.* **27** (1966), 768–771. *MR* **32** #4653

[17] Render, H., Nonstandard topology on function spaces with applications to hyperspaces, *Trans. Amer. Math. Soc.* **336** (1993), 101–119. *MR* **93j**:54036

[18] Render, H., Nonstandard methods of completing quasi-uniform spaces, *Topology Appl.* **62** (1995), 101–125. *MR* **96a**:54041

[19] Seyedin, M., Quasi-uniform spaces and topological homeomorphism groups, *Canad. Math. Bull.* **17** (1974), 97–98. *MR* **50** #8455

[20] Seyedin, M., R_3-quasi-uniform spaces and topological homeomorphism groups, *Proc. Amer. Math. Soc.* **52** (1975), 465–468. *MR* **51** #11441

[21] Sieber, J. L. and Pervin, W. J., Completeness in quasi-uniform spaces, *Math. Ann.* **158** (1965), 79–81. *MR* **30** #2449

[22] Vakil, N., Monadic binary relations and the monad systems at nearstandard points, *J. Symbolic Logic* **52** (1987), 689–697. *MR* **88j**:03049

[23] Wattenberg, F., Non-standard topology and extensions of monad systems to infinite points, *J. Symbolic Logic* **36** (1971), 463–476. *MR* **45** #2642

[24] WILLIAMS, J., Locally uniform spaces, *Trans. Amer. Math. Soc.* **168** (1972), 435–469. *MR* **45** #5950

(Received February 15, 1994)

GERHARD-MERCATOR-UNIVERSITÄT
GESAMTHOCHSCHULE DUISBURG
FACHBEREICH 11 MATHEMATIK
LOTHARSTRASSE 65
D–47057 DUISBURG
GERMANY

render@math.uni-duisburg.de

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 445–452

SEMICONTINUITY OF TOPOLOGICAL LIMITS OF MULTIVALUED MAPS

J. EWERT

Abstract

For topological limits of quasicontinuous multivalued maps we describe the sets of points of upper and lower semicontinuity.

In a topological space (X,T) by $\operatorname{Int} A$ and $\operatorname{Cl} A$ we denote the interior and the closure of a set $A\subset X$. A set A is called semi-open if $A\subset \operatorname{Cl}(\operatorname{Int} A)$, [9]; semi-closed if $X\setminus A$ is semi-open, [1]. The family

$$T_q=\{U\setminus H:\ U\in T,\ H \text{ is of the first category}\}$$

is a topology on X. The T_q-closure and the T_q-interior of A will be denoted by $\operatorname{Cl}_q A$ and $\operatorname{Int}_q A$, respectively. Then, if (X,T) is a Baire space, we have:

(1) the spaces (X,T) and (X,T_q) have the same classes of the first category sets, [6];

(2) a set $A\subset X$ is T_q-semi-open (T_q-semi-closed) iff it is of the form $A=B\setminus H$ (resp. $A=B\cup H$), where B is semi-open (semi-closed) and H is of the first category, [5];

(3) if A is T_q-open, then $\operatorname{Cl} A=\operatorname{Cl}_q A$, [6].

LEMMA 1. *Let (X,T) be a Baire space. Then:*

(a) *If a set $A\subset X$ is T_q-semi-open, then $\operatorname{Cl} A=\operatorname{Cl}_q A$.*

(b) *A set $A\subset X$ is T_q-semi-open if and only if it is of the form $A=(G\setminus H)\cup L$, where G is open, H, L are of the first category and $L\subset \operatorname{Cl} G$.*

PROOF. (a) If A is T_q-semi-open, then applying (3) to the set $\operatorname{Int}_q A$ we obtain $\operatorname{Int}_q A\subset A\subset \operatorname{Cl}_q(\operatorname{Int}_q A)=\operatorname{Cl}(\operatorname{Int}_q A)$, hence $\operatorname{Cl} A=\operatorname{Cl}(\operatorname{Int}_q A)=\operatorname{Cl}_q(\operatorname{Int}_q A)=\operatorname{Cl}_q A$.

(b) Let A be a T_q-semi-open set. According to (2) it is of the form $A=B\setminus H$, where B is semi-open and H is of the first category. Then $B=G\cup H_1$, where G is open and $H_1\subset(\operatorname{Cl} G)\setminus G$. Thus we obtain $A=(G\setminus H)\cup(H_1\setminus H)$ and it suffices to put $L=H_1\setminus H$. Conversely, assume that $A=(G\setminus H)\cup L$, where G is open, H, L of the first category with $L\subset \operatorname{Cl} G$. Then $G\setminus H\in T_q$

1991 *Mathematics Subject Classification.* Primary 54C60.
Key words and phrases. Multivalued map, quasicontinuity, topological limits.

Sponsored by KBN Research Grant (1992–94) No. 2, 1144–91–01.

and $\mathrm{Cl}\, G = \mathrm{Cl}(G \setminus H) = \mathrm{Cl}_q(G \setminus H)$. Thus $L \subset \mathrm{Cl}_q(G \setminus H)$ which implies $A \subset \mathrm{Cl}_q(G \setminus H) \subset \mathrm{Cl}_q(\mathrm{Int}_q A)$ and the proof is completed. □

For a sequence $\{A_n : n \geqq 1\}$ of subsets of a topological space the sets $\mathrm{Ls}\, A_n$ and $\mathrm{Li}\, A_n$ are defined as follows [7]:

$x \in \mathrm{Ls}\, A_n$ iff each neighbourhood of x intersects infinitely many sets A_n;

$x \in \mathrm{Li}\, A_n$ iff for each neighbourhood U of x there is n_0 such that $U \cap A_n \neq \emptyset$ for each $n \geqq n_0$.

If $\mathrm{Li}\, A_n = \mathrm{Ls}\, A_n$, then this set will be denoted as $\mathrm{Lt}\, A_n$.

Let X, Y be topological spaces; a multivalued map $F \colon X \to Y$ is a function defined on X and assuming non-empty values in the power set of Y. For a multivalued map $F \colon X \to Y$ and a set $W \subset Y$ we will write: $F^+(W) = \{x \in X : F(x) \subset W\}$ and $F^-(W) = \{x \in X : F(x) \cap W \neq \emptyset\}$. Furthermore, the symbols $C^+(F)$ and $C^-(F)$ will be used to denote the sets of all points at which F is upper or lower semicontinuous, respectively. A multivalued map $F \colon X \to Y$ is called upper (lower) quasi-continuous if for each open set $W \subset Y$ the set $F^+(W)$, (resp. $F^-(W)$) is semi-open [12]. Now, let $F, F_n \colon X \to Y$, $n \geqq 1$, be multivalued maps. We will write $F = \mathrm{Ls}\, F_n$, $F = \mathrm{Li}\, F_n$, $F = \mathrm{Lt}\, F_n$ if $F(x) = \mathrm{Ls}\, F_n(x)$, $F(x) = \mathrm{Li}\, F_n(x)$ or $F(x) = \mathrm{Lt}\, F_n(x)$ for each $x \in X$, respectively.

In [10] "Baire continuous" multivalued maps are considered; these are defined as follows: a map $F \colon X \to Y$ is called upper (lower) Baire continuous if for each open set $W \subset Y$ the set $F^+(W)$, (resp. $F^-(W)$) belongs to the class $\mathrm{Br} = \{A \subset X : A = (G \setminus H) \cup L$ where G is open, H, L of the first category, $L \subset \mathrm{Cl}\, G\}$. If X is a Baire space, then in virtue of Lemma 1, the upper (lower) Baire continuity coincides with the upper (lower) T_q-quasi-continuity. Thus, the result presented in [10] can be rewritten in the form:

(4) Let X be a Baire T_1 space and Y a compact metric one. If $F, F_n \colon X \to Y$, $n \geqq 1$, are multivalued maps with $F = \mathrm{Ls}\, F_n$ and F_n are lower T_q-quasi-continuous, then the set $X \setminus C^+(F)$ is of the first category.

An extension of this result will be given in Theorem 1.

A topological space Y is said to be perfect if each closed subset of Y is G_δ. We recall that a topological space Y is perfect normal (need not be T_1) iff for every open set $W \subset Y$ there exists a sequence $W_1, W_2, \ldots$ of open subsets of Y such that $W = \bigcup_{n=1}^{\infty} W_n$ and $\mathrm{Cl}\, W_n \subset W$ for $n \geqq 1$ [3, p. 73]. This fact and [2, p. 372] give the following characterization: a topological space Y is perfect normal iff each open set $W \subset Y$ is of the form $W = \bigcup_{n=1}^{\infty} W_n$, where W_n are open sets with $\mathrm{Cl}\, W_n \subset W_{n+1}$ for $n \geqq 1$.

LEMMA 2. *Let X be a non-empty set, Y a perfect normal space and let $F, F_n \colon X \to Y$, $n \geqq 1$, be multivalued maps with $F = \mathrm{Ls}\, F_n$. If for each $x \in X$*

there is $n \geqq 1$ *such that* $\mathrm{Cl}\Big(\bigcup_{j=n}^{\infty} F_j(x)\Big)$ *is compact, then for each open set* $U \subset Y$ *we have*

$$F^+(U) = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} F_j^+(\mathrm{Cl}\, W_k),$$

where W_k *are open sets with* $\mathrm{Cl}\, W_k \subset W_{k+1}$ *for* $k \geqq 1$*, and* $U = \bigcup_{k=1}^{\infty} W_k$*.*

PROOF. Let $W \subset Y$ be an open set and let $x \in F^+(W)$; then

$$\bigcap_{n=1}^{\infty} \mathrm{Cl}\Big(\bigcup_{j=n}^{\infty} F_j(x)\Big) \subset W.$$

There is an n_0 such that $\Big\{\mathrm{Cl}\Big(\bigcup_{j=n}^{\infty} F_j(x)\Big) : n \geqq n_0\Big\}$ is a decreasing sequence of compact sets, so we can choose n such that $\mathrm{Cl}\Big(\bigcup_{j=n}^{\infty} F_j(x)\Big) \subset W$. Hence, using that F has compact values, we obtain

$$F^+(W) \subset \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} F_j^+(W) \subset \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} F_j^+(\mathrm{Cl}\, W) \subset F^+(\mathrm{Cl}\, W).$$

Now, let $U \subset Y$ be an open set. It can be represented in the form

$$U = \bigcup_{k=1}^{\infty} W_k = \bigcup_{k=1}^{\infty} \mathrm{Cl}\, W_k,$$

where W_k are open sets with $\mathrm{Cl}\, W_k \subset W_{k+1}$ for each $k \geqq 1$. Thus

$$F^+(U) = \bigcup_{k=1}^{\infty} F^+(W_k) \subset \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} F_j^+(\mathrm{Cl}\, W_k) \subset \bigcup_{k=1}^{\infty} F^+(\mathrm{Cl}\, W_k) \subset F^+(U),$$

which completes the proof. □

THEOREM 1. *Let* (X,T) *be a Baire space,* Y *a separable metric one and let* $F, F_n \colon X \to Y$*,* $n \geqq 1$*, be multivalued maps with* $F = \mathrm{Ls}\, F_n$*. If for each* $x \in X$ *there is* $n \geqq 1$ *such that* $\mathrm{Cl}\Big(\bigcup_{j=n}^{\infty} F_j(x)\Big)$ *is compact and all* F_n *are lower* T_q*-quasi-continuous, then* $X \setminus C^+(F)$ *is of the first category.*

PROOF. Following Lemma 2, for an open set $U \subset Y$ it holds

$$F^+(U) = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{j=n}^{\infty} F_j^+(\mathrm{Cl}\, W_k),$$

where W_k are open sets with $\operatorname{Cl} W_k \subset W_{k+1}$ for $k \geqq 1$, and $U = \bigcup_{k=1}^{\infty} W_k$. Since the maps F_n are lower T_q-quasi-continuous, the sets $\bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k)$ are T_q-semi-closed; thus the sets

$$\bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k) \setminus \operatorname{Int} \bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k)$$

are of the first category. Furthermore, we have

$$\begin{aligned} &F^+(U) \setminus \operatorname{Int} F^+(U) \subset \\ &\subset \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \Big(\bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k) \setminus \bigcup_{m=1}^{\infty} \bigcup_{i=1}^{\infty} \operatorname{Int} \bigcap_{j=i}^{\infty} F_j^+(\operatorname{Cl} W_m) \Big) \subset \\ &\subset \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \Big(\bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k) \setminus \operatorname{Int} \bigcap_{j=n}^{\infty} F_j^+(\operatorname{Cl} W_k) \Big). \end{aligned}$$

From this it follows that $F^+(U) \setminus \operatorname{Int} F^+(U)$ is of the first category for each open set $U \subset Y$. Let $\mathcal{B}$ be a countable open base in Y. By $\{U_m : m \geqq 1\}$ we denote all finite sums of elements of $\mathcal{B}$; then — since F has compact values — we have

$$X \setminus C^+(F) = \bigcup_{n=1}^{\infty} (F^+(U_n) \setminus \operatorname{Int} F^+(U_n)).$$

Thus $X \setminus C^+(F)$ is of the first category which finishes the proof. □

As shown by the example given below, in Theorem 1 the limit Ls cannot be replaced by Li. Moreover, for upper T_q-quasi-continuous maps F_n the set $X \setminus C^-(\operatorname{Ls} F_n)$ need not be of the first category. Also it follows from [4] that for upper (lower) T_q-quasi-continuous maps F_n the sets $X \setminus C^+(\operatorname{Ls} F_n)$ and $X \setminus C^+(\operatorname{Lt} F_n)$ (resp. $X \setminus C^-(\operatorname{Li} F_n)$ and $X \setminus C^-(\operatorname{Lt} F_n)$) can be of the second category. But up to now the following problem formulated in [10] is not resolved: is the set $X \setminus C^-(F)$ of the first category if $F = \operatorname{Li} F_n$ and F_n are upper T_q-quasi-continuous? A partial answer will be given by Theorem 2.

EXAMPLE. Let μ denote the Lebesgue measure on the real line, $Q = \{q_j : j \geqq 1\}$ the set of rational numbers and $\mathbf{R}$ the space of real numbers with the usual metric. For each $n \geqq 1$ and $j \in \{1, 2, \ldots, n\}$ we fix numbers a_{nj}, b_{nj} satisfying

$$a_{nj} < q_j < b_{nj}$$

$$[a_{nj}, b_{nj}] \cap [a_{ni}, b_{ni}] = \emptyset \quad \text{for } i \neq j,\ i, j \in \{1, 2, \ldots, n\}$$

$$\mu\Big(\bigcup_{j=1}^{n} (a_{nj}, b_{nj}) \Big) < 2^{-n}.$$

Now we put:

$$A_n = \bigcup_{j=1}^{n} (a_{nj}, b_{nj}), \quad A = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \ \text{ and } \ B = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Then $\mu(A) = 0$, $Q \subset \bigcup_{k=n}^{\infty} A_k$ for each $n \geqq 1$, all sets A_n, $R \setminus A_n$ are semi-open. Hence A and $R \setminus A$ are dense sets and $R \setminus A$ is of the first category. Moreover, B is dense and $\mu(B) = 0$, so $R \setminus B$ is dense, too. Now we define multivalued maps $F_n \colon R \to [0, 2]$ assuming

$$F_n(x) = \begin{cases} \{0,1\} & \text{for } x \in A_n, \\ \{1,2\} & \text{for } x \in R \setminus A_n. \end{cases}$$

The maps F_n are lower and upper T_q-quasi-continuous and

$$\operatorname{Li} F_n(x) = \begin{cases} \{0,1\} & \text{for } x \in B, \\ \{1\} & \text{for } x \in A \setminus B, \\ \{1,2\} & \text{for } x \in R \setminus A, \end{cases}$$

$$\operatorname{Ls} F_n(x) = \begin{cases} \{0,1\} & \text{for } x \in B, \\ \{0,1,2\} & \text{for } x \in A \setminus B, \\ \{1,2\} & \text{for } x \in R \setminus A. \end{cases}$$

Then we have $C^+(\operatorname{Li} F_n) = \emptyset = C^-(\operatorname{Ls} F_n)$, so $R \setminus C^+(\operatorname{Li} F_n)$ and $R \setminus C^-(\operatorname{Ls} F_n)$ are of the second category. □

LEMMA 3. *Let X, Y be topological spaces and let $F, F_n \colon X \to Y$, $n \geqq 1$, be multivalued maps with $F = \operatorname{Lt} F_n$. Then:*

(a) $F^-(V) \subset \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(V)$
for each open set $V \subset Y$;

(b) $\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(M) \subset F^-(M)$
for each compact set $M \subset Y$;

(c) *if Y is a locally compact separable metric space, then for each open set $V \subset Y$ it holds*

$$F^-(V) = \bigcup_{j=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(\operatorname{Int} A_j) = \bigcup_{j=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(A_j),$$

where A_j are compact sets with $A_j \subset \operatorname{Int} A_{j+1}$ for $j \geqq 1$, and $V = \bigcup_{j=1}^{\infty} A_j$.

PROOF. If $V \subset Y$ is an open set, then — since $F = \operatorname{Li} F_n$ — the inclusion (a) is an immediate consequence of the definition of the limit Li.

Let $M \subset Y$ be a compact set and let $x \in \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(M)$. We choose n_0 such that $F_n(x) \cap M \neq \emptyset$ for each $n \geqq n_0$. Then we obtain $\emptyset \neq \operatorname{Ls}(F_n(x) \cap M) \subset (\operatorname{Ls} F_n(x)) \cap M = F(x) \cap M$, and this gives

$$\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(M) \subset F^-(M).$$

Now we are going to prove (c). Let $V \subset Y$ be an open set. Under the assumptions on Y it can be represented in the form $V = \bigcup_{j=1}^{\infty} A_j$, where A_j are compact sets and $A_j \subset \operatorname{Int} A_{j+1}$ for $j \geqq 1$ [7, p. 51]. Applying (a) and (b) we obtain

$$F^-(V) = \bigcup_{j=1}^{\infty} F^-(\operatorname{Int} A_j) \subset \bigcup_{j=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(\operatorname{Int} A_j) \subset$$
$$\subset \bigcup_{j=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(A_j) \subset \bigcup_{j=1}^{\infty} F^-(A_j) = F^-(V).$$

Thus we have shown (c) and the proof is completed. □

THEOREM 2. *Let (X,T) be a Baire space and Y a locally compact separable metric one. If $F, F_n \colon X \to Y$, $n \geqq 1$, are multivalued maps with $F = \operatorname{Lt} F_n$ and F_n are upper T_q-quasi-continuous, then the set $X \setminus C^-(F)$ is of the first category.*

PROOF. Let $V \subset Y$ be an open set with $\operatorname{Cl} V$ compact. Then from (a) and (b) in Lemma 3 we obtain

$$F^-(V) \setminus \operatorname{Int} F^-(\operatorname{Cl} V) \subset \Big(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(V)\Big) \setminus \operatorname{Int} \Big(\bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^-(V)\Big) \subset$$
$$\subset \bigcup_{n=1}^{\infty} \Big(\bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V) \setminus \bigcup_{k=1}^{\infty} \operatorname{Int} \bigcap_{m=k}^{\infty} F_m^-(\operatorname{Cl} V)\Big) \subset$$
$$\subset \bigcup_{n=1}^{\infty} \Big(\bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V) \setminus \operatorname{Int} \bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V)\Big).$$

Since F_n are upper T_q-quasi-continuous, the sets $\bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V)$ are T_q-semi-closed. Hence the sets

$$\bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V) \setminus \operatorname{Int} \bigcap_{m=n}^{\infty} F_m^-(\operatorname{Cl} V)$$

are of the first category, so also $F^-(V) \setminus \operatorname{Int} F^-(\operatorname{Cl} V)$ is of the first category. Now, by $\{V_j \colon j \geqq 1\}$ we denote a base of Y consisting of open sets with $\operatorname{Cl} V_j$

compact for $j \geqq 1$. Then by regularity of Y we have

$$X \setminus C^{-}(F) = \bigcup_{j=1}^{\infty} (F^{-}(V_j) \setminus \operatorname{Int} F^{-}(\operatorname{Cl} V_j)),$$

which finishes the proof. □

The properties (a) and (b) from Lemma 3 can be used to obtain also other results. We remind that the graph of a multivalued map $F: X \to Y$ is the set $\operatorname{Gr}(F) = \{(x,y) \in X \times Y : y \in F(x)\}$ and:

(5) Let X be a topological space, Y a locally compact one and let $F: X \to Y$ be a multivalued map with closed values. The graph of F is closed if and only if for each open set $W \subset Y$ with $Y \setminus W$ compact the set $F^{+}(W)$ is open [11].

THEOREM 3. *Let (X,T) be a Baire space, Y a locally compact separable metric one and let $F_n, F: X \to Y$, $n \geqq 1$, be multivalued maps with $F = \operatorname{Lt} F_n$. If F_n have closed values and the sets $\operatorname{Gr}(F_n)$ are closed for $n \geqq 1$, then the set $X \setminus C^{-}(F)$ is of the first category.*

PROOF. Let $V \subset Y$ be an open set. Under the assumptions on Y it can be represented in the form $V = \bigcup_{n=1}^{\infty} A_n$, where A_n are compact sets and $A_n \subset \operatorname{Int} A_{n+1}$ for $n \geqq 1$ [7, p. 51]. From Lemma 3(c) we have

$$F^{-}(V) = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^{-}(\operatorname{Int} A_j) = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{m=n}^{\infty} F_m^{-}(A_j).$$

According to (5) the sets $F_m^{-}(A_j)$ are closed, so $F^{-}(V)$ is an F_σ set for each open set $V \subset Y$. Furthermore

$$X \setminus C^{-}(F) = \bigcup_{n=1}^{\infty} (F^{-}(V_n) \setminus \operatorname{Int} F^{-}(V_n)),$$

where $\{V_n : n \geqq 1\}$ is an open base of Y, so the proof is completed. □

REMARK. It is easy to see that in Theorem 3 it suffices to suppose that $\operatorname{Gr}(F_n)$ are closed in the space $(X, T_q) \times Y$.

In the sequel ω_1 denotes the first uncountable ordinal number. A multivalued map $F: X \to Y$ is said to be of the upper (lower) class $\alpha < \omega_1$ if for each open set $W \subset Y$ the set $F^{+}(W)$ (resp. $F^{-}(W)$) is of the additive class α [8]. As simple consequences of Lemma 2 and 3(c) we get

COROLLARY 1. *Let X be a topological space, Y a perfect normal one and let $F_n, F: X \to Y$, $n \geqq 1$, be multivalued maps with $F = \operatorname{Ls} F_n$. If F_n are of* *lower class* α *and for each* $x \in X$ *there is* $n \geqq 1$ *such that* $\mathrm{Cl}\left(\bigcup_{j=n}^{\infty} F_j(x)\right)$ *is compact, then* F *is of the upper class* $\alpha + 1$. □

COROLLARY 2. *Let* X *be a topological space,* Y *a locally compact separable metric one and let* $F_n, F\colon X \to Y$, $n \geqq 1$, *be multivalued maps with* $F = \mathrm{Lt}\, F_n$. *Then:*

(a) *if* F_n *are of an upper class* α, *then* F *is of the lower class* $\alpha + 1$;

(b) *if* F_n *are of a lower class* α, *then* F *is of the lower class* $\alpha + 2$. □

REFERENCES

[1] CROSSLEY, S. G. and HILDEBRAND, S. K., Semi-closed sets and semicontinuity in topological spaces, *Texas J. Sci.* **22** (1971), 123–126.

[2] CSÁSZÁR, Á., *General topology*, Adam Hilger Ltd., Bristol, 1978. *MR* **57** #13812

[3] ENGELKING, R., *General topology*, Monografie Matematyczne, Tom 60, PWN – Polish Scientific Publishers, Warszawa, 1977. *MR* **58** #18316b

[4] EWERT, J., Semicontinuity of maps which are limits of sequences of quasi-continuous multivalued maps, *Mat. Vesnik* **39** (1987), 47–52. *MR* **88j**:54029

[5] EWERT, J., Quasi-continuity of multivalued maps with respect to the qualitative topology, *Acta Math. Hungar.* **56** (1990), 39–44. *MR* **92b**:54035

[6] HASHIMOTO, H., On the *-topology and its application, *Fund. Math.* **91** (1976), 5–10. *MR* **54** #1179

[7] KURATOWSKI, K., *Topology*, Vol. 1, New edition, revised and augmented, Panstwowe Wydawnictwo Naukowe, Warszawa; Academic Press, New York–London, 1966. *MR* **36** #840

[8] KURATOWSKI, K., Some remarks on the relation of classical set-valued mappings to the Baire classification, *Colloq. Math.* **42** (1979), 273–277. *MR* **81c**:54024

[9] LEVINE, N., Semi-open sets and semi-continuity in topological spaces, *Amer. Math. Monthly* **70** (1963), 36–41. *MR* **29** #4025 and **51** #4130

[10] MATEJDES, M., Semicontinuity of maps which are limits of sequences of Baire continuous multivalued maps, *Mat. Vesnik* **42** (1990), 217–220. *MR* **93g**:54028

[11] NEUBRUNN, T., C-continuity and closed graphs, *Časopis Pěst. Mat.* **110** (1985), 172–178. *MR* **87f**:54026

[12] NEUBRUNN, T., Quasi-continuity, *Real Anal. Exchange* **14** (1988/89), 259–306. *MR* **90f**:54025

(Received November 8, 1994)

WYŻSZA SZKOŁA PEDAGOGICZNA
WYDZIAL MATEMATYCZNO-PRZYRODNICZY
INSTYTUT MATEMATYKI
UL. ARCISZEWSKIEGO 22A
PL–76-200 SŁUPSK
POLAND

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 453–473

TIME DEPENDENT ANALYSIS OF T-POLICY $M/M/1$ QUEUES — A NEW APPROACH

KANWAR SEN and RITU GUPTA

Abstract

This paper demonstrates a simple and elegant lattice path combinatoric technique for computing transient probabilities concerning $M/M/1$ queueing models. Through this lattice path approach time dependent analysis of T-policy $M/M/1$ queue is presented. The transient probabilities computed herein are free from modified Bessel function and are amenable to pragmatic probabilistic interpretations. As a special case the results for ordinary $M/M/1$ queues are checked.

1. Introduction

Consider a T-policy $M/M/1$ queueing model which activates the server T time units after the end of a busy period to determine if customers are present. If no customers are found when the server scans the queue, it is turned off, and the system is scanned again after an interval of length T. This procedure is repeated until the server finds at least one customer waiting, after which the server is kept in active state until the system becomes empty. This model can also be viewed as one where the server takes a sequence of vacations each of duration T, at the end of busy period (see Doshi [4]). Henceforth T-policy $M/M/1$ queueing model will be referred to as $M/M/1(T)$.

Different aspects of T-policy queues were studied by Heyman [8] (see also Teghem [21], Takagi [20]). However, little effort was made to find the transient solution of this model (see [20]). As opposed to classical method which entails formulation of tedious unwieldy difference-differential equations, in this paper the lattice path approach — a new combinatorial technique is adopted for studying transient behaviour of $M/M/1(T)$ queues. Starting initially with k ($\geqq 0$) units, the probability of i arrivals and j departures up to time t is found for the $M/M/1(T)$ queue. This probability in turn leads to the probability of the number of units in the system up to time t. Over the years combinatorial techniques have been successfully employed in solving queueing problems (refer to Takács [18], [19]). Recently, using lattice path combinatorics Mohanty and Panny [15], Böhm and Mohanty [2], Kanwar

1991 *Mathematics Subject Classification.* Primary 60K25; Secondary 60J15.

Key words and phrases. T-policy, vacation period, busy period, $M/M/1(T)$ model, discretized $M/M/1(T)$ model, slot, lattice path, transient probability.

Sen, Jain and Gupta [12] obtained transient solutions of $M/M/1$ queues, $M/M/1$ queues under (M, N) and (O, K) control policies, respectively.

In the lattice path approach transient probabilities are computed by using a four stage procedure:

(a) discretizing the continuous model and then representing it by a lattice path;

(b) computing the number of lattice paths stipulated by the process;

(c) computing probabilities associated with such paths;

(d) applying limiting process to these probabilities so as to obtain the transient probabilities of $M/M/1(T)$ queues.

In addition to evolving a simple solution to the said problems this method leads to the results which are conducive to significant probabilistic interpretations and provide meaningful insight into the nature of the process involved.

2. Lattice path approach

For determining the transient solution we first propose a discrete time analogue of $M/M/1(T)$ queueing process. Assume that the time interval $(0, t)$ is segmented into a sequence of t/h time intervals (slots) each of duration h (> 0) such that t/h and T/h are integers. Consequently, scanning period for T-policy will be T/h time slots. We further assume that

(i) No more than one customer may arrive or finish being served in a given slot;

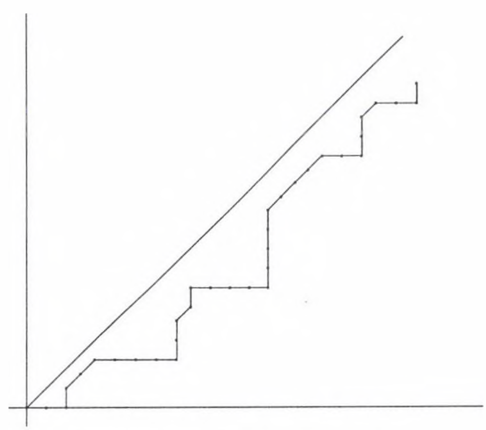

Fig. 1. Lattice path

(ii) Events in different slots are independent. Observe the system at epoch $1, 2, \ldots, t/h$ each of which marks the end of an interval of duration h.

Associate a lattice path with the queueing system representing an arrival, a departure, and a stay in a slot by a unit horizontal step, a unit vertical step, and by a unit diagonal step, respectively (see Figure 1).

If the state of the system is described by the pair (x, y) $(x \geqq y,\ x \geqq k)$, where $x - k$ denotes the number of arrivals and stays and y denotes the number of departures and stays. The transition probabilities associated with the queueing system are

(i) Under busy period

$$\begin{aligned} P[(x,y) \to (x+1,y)] &= P[\text{an arrival in a slot}] \\ &= \lambda h + o(h) \\ P[(x,y) \to (x,y+1)] &= P[\text{a departure in a slot}] \\ &= \begin{cases} \mu h + o(h), & x \neq y, \\ 0, & x = y, \end{cases} \\ P[(x,y) \to (x+1,y+1)] &= P[\text{stay in a slot}] \\ &= \begin{cases} 1 - \lambda h - \mu h + o(h), & x \neq y \\ 1 - \lambda h + o(h), & x = y. \end{cases} \end{aligned} \tag{2.1}$$

(ii) Under vacation period

$$\begin{aligned} P[(x,y) \to (x+1,y)] &= P[\text{an arrival in a slot}] \\ &= \lambda h + o(h) \\ P[(x,y) \to (x+1,y+1)] &= P[\text{stay in a slot}] = 1 - \lambda h + o(h). \end{aligned} \tag{2.2}$$

Let the probability that the discretized $M/M/1(T)$ queueing process encounters i arrivals and j departures in t/h time slots starting initially with k $(\geqq 0)$ units be denoted by:

$Q_{i,j;k}(t/h)$ When the system encounters no vacations (i.e. without being empty in-between).

$P_{i,j;k}(t/h, T/h)$ When the system encounters at least one vacation and the queue length attained is less than or equal to the number of arrivals encountered in the last sequence of vacations.

${}_LP_{i,j;k}(t/h, T/h)$ When the system encounters at least one vacation and the queue length attained is greater than or equal to the number of arrivals encountered in the last sequence of vacations.

Further let

${}_TP_{i,j;k}(t/h, T/h)$ denote the probability that $M/M/1(T)$ queueing process encounters i arrivals and j departures in t/h time slots.

The discretized $M/M/1(T)$ model leads to binomial probability distributions of arrivals and departures within a given period of time. And by a suitable limiting process ($h \to 0$; $t/h \to \infty$; $\lambda h \to 0$, $\mu h \to 0$), these distributions tend to Poisson distributions. Hence the transient probabilities for the continuous time model following Poisson distribution can be obtained from discrete time analogue by passing on to the limit as $h \to 0$ (Meisling [13]), Whittle [22]). Further, let $Q^*_{i,j;k}(t)$, $P^*_{i,j;k}(t,T)$ and ${}_1P^*_{i,j;k}(t,T)$ and ${}_TP^*_{i,j;k}(t)$ be the respective continuous time analogue.

For computing the probabilities defined above we require the number of lattice paths stipulated by the discretized $M/M/1(T)$ queueing process. Counting of lattice paths is performed in two stages. Firstly, we delete all diagonal steps to construct new path called skeleton path (see Figure 2).

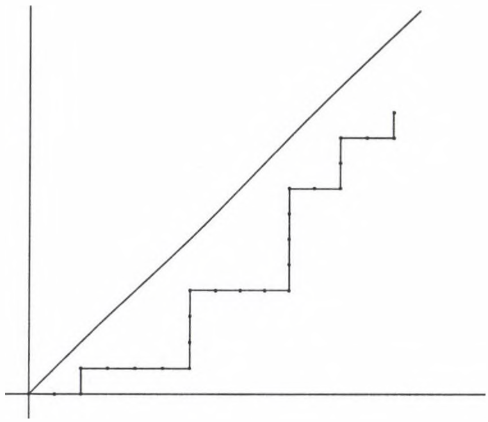

Figure 2. Skeleton lattice path

Lemmas 1, 2 and 3 stated below give the number of skeleton lattice paths nomenclatured as $LP_{d_1,d_2,\dots,d_r}(k,r;m,n)$ from $(k,0)$ to (m,n) comprising of $m-k$ horizontal steps, n vertical steps, touching r times the barrier $y=x$ and having at least d_i $(\geqq 1)$ horizontal steps preceding a vertical step after the ith $(i=1,2,\dots,r)$ touch with the barrier $x=y$, respectively. These could be computed following arguments and constructions in Csáki and Vincze [3], Kanwar Sen [9], Kanwar Sen, Jain and Gupta [12], Mohanty [14].

LEMMA 1. *For* $k>0$, $1 \leqq m-n \leqq d_r$, $d_i \geqq 1$ $(i=1,2,\dots,r)$

$$LP_{d_1,d_2,\dots,d_r}(k,r;m,n) = \left[\binom{m+n-k-d}{n} - \binom{m+n-k-d}{m}\right]. \tag{2.3}$$

For $k>0$, $m-n \geqq d_r$, $d_i \geqq 1$ $(i=1,2,\dots,r)$

$$LP_{d_1,d_2,\dots,d_r}(k,r;m,n) = \left[\binom{m+n-k-d}{m-d_r} - \binom{m+n-k-d}{m}\right], \tag{2.4}$$

where $d=d_1+d_2+\dots+d_r$.

PROOF. Consider a skeleton path R envisaged in Lemma 1 (see Figure 3).

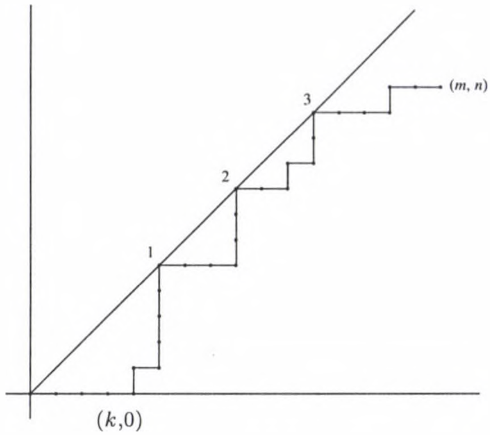

Figure 3. Skeleton lattice path; $k=4,\ m=16,\ n=12,\ r=3$

Make the following transformations on R.

(i) Remove k horizontal steps starting from the origin.

(ii) Remove $d_1, d_2, \dots, d_r$ horizontal steps immediately following $1^{\text{st}}, 2^{\text{nd}}, \dots, r^{\text{th}}$ touch point, respectively.

(iii) Concatenate the truncated path segments in succession. The transformed path is shown in Figure 4.

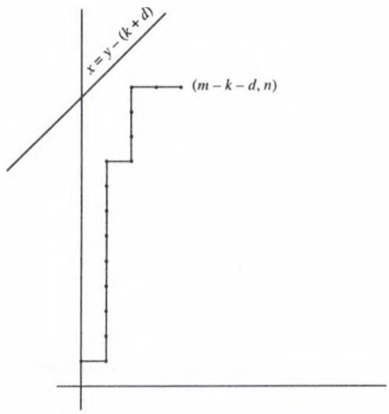

Fig. 4

This path from $(0,0)$ to $(m-k-d, n)$ is characterized by not crossing the barrier $X = Y-(k+d)$. The number of such paths as given by reflection principle (see Feller [6], Mohanty [14]) is

$$\left[\binom{m+n-k-d}{n} - \binom{m+n-k-d}{m}\right]. \tag{2.5}$$

Hence (2.3) is established.

Further, (2.4) can be proved by using similar transformations on skeleton lattice path stipulated. The transformed skeleton path in this case is a path from $(0,0)$ to $(m-k-d,n)$, which touches or crosses the barrier $X = Y-(k+d-d_r)$, and remains below the barrier $X = Y-(k+d)$. Therefore the number of such paths

(2.6) = number of skeleton paths from $(0,0)$ to $(m-k-d,n)$, which touch or cross the barrier $X=Y-(k+d-d_r)-$ number of paths from $(0,0)$ to $(m-k-d,n)$ which touch or cross the barrier $X=Y-(k+d)$.

$$= \left[\binom{m+n-k-d}{m-d_r} - \binom{m+n-k-d}{m}\right].$$

Hence (2.4) follows.

LEMMA 2. *For* $k=0,\ 1<m-n\leqq d_r,\ d_i\geqq 1\ (i=1,2,\ldots,r)$

$$LP_{d_1,d_2,\ldots,d_r}(0,r;m,n) = \left[\binom{m+n-d-1}{n} - \binom{m+n-d-1}{m}\right]. \tag{2.7}$$

For $k=0,\ m-n\geqq d_r,\ d_i\geqq 1\ (i=1,2,\ldots,r)$

$$LP_{d_1,d_2,\ldots,d_r}(0,r;m,n) = \left[\binom{m+n-d-1}{m-d_r} - \binom{m+n-d-1}{m}\right]. \tag{2.8}$$

LEMMA 3. *Lattice paths envisaged with $m=n$ will not encounter any horizontal step after the r^{th} contact point.*

Hence for $m=n,\ k>0,\ d_i\geqq 1\ (i=1,2,\ldots,r-1)$

$$LP_{d_1,d_2,\ldots,d_{r-1}}(k,r;m,n) = \left[\binom{2m-k-d-1}{m-1} - \binom{2m-k-d-1}{m}\right], \tag{2.9}$$

and for $m=n,\ k=0,\ d_i\geqq 1\ (i=1,2,\ldots,r-1)$

$$LP_{d_1,d_2,\ldots,d_{r-1}}(0,r;m,n) = \left[\binom{2m-d-2}{m-1} - \binom{2m-d-2}{m}\right]. \tag{2.10}$$

Lemmas 2 and 3 can be proved by following constructions as in Lemma 1.

Lastly, the number of lattice paths stipulated by the process is obtained by inserting diagonal steps at the intervening vertices of the skeleton lattice path (see, for example, Dua, Khadilkar and Kanwar Sen [5], Kanwar Sen and Ahuja [10]). Then on associating transition probabilities defined in (2.1) and (2.2) with the number of lattice paths stipulated and passing on to the limit $h\to 0$, we get the probabilities for $M/M/1(T)$ queueing process.

3. Probability for discrete time model

Transient probabilities for discretized $M/M/1(T)$ queueing process (see [15]) are developed in Theorems below.

THEOREM 1.

$$\begin{aligned}(3.1)\quad & P_{i,j;k}[t/h, T/h] \\ &= \sum_{r=1}^{j-k+1} \sum_{R_1}\sum_{R_2}\sum_{R_3} \binom{s-1}{r-1} \prod_{q=1}^{r} \binom{T/h}{d_q} \left[\binom{i+j-d}{j} - \binom{i+j-d}{i+k}\right] \\ &\times \left[\begin{matrix} t/h - sT/h \\ i+j-d \end{matrix}\right] (\lambda h + o(h))^i (\mu h + o(h))^j \\ &\times (1-\lambda h + o(h))^{(sT/h)-d} \\ &\times (1-\lambda h - \mu h + o(h))^{(t/h)-(sT/h)-(i+j-d)}\end{aligned}$$

and

$$\begin{aligned}(3.2)\quad & {}_1P_{i,j;k}(t/h, T/h) \\ &= \sum_{r=1}^{j-k+1} \sum_{R_1}\sum_{R_2}\sum_{R_3} \binom{s-1}{r-1} \prod_{q=1}^{r} \binom{T/h}{d_q} \left[\binom{i+j-d}{i+k-d_r} - \binom{i+j-d}{i+k}\right] \\ &\times \left[\begin{matrix} t/h - sT/h \\ i+j-d \end{matrix}\right] (\lambda h + o(h))^i (\mu h + o(h))^j \\ &\times (1-\lambda h + o(h))^{(sT/h)-d} \\ &\times (1-\lambda h - \mu h + o(h))^{(t/h)-(sT/h)-(i+j-d)},\end{aligned}$$

where

R_1: *Summation over* d *from* r *to* $\min\left(\frac{T}{h}r, i\right)$;

R_2: *Multiple summation over* $d_1, d_2, \dots, d_r$ *such that* $1 \leqq d_i \leqq \frac{T}{h}$, $(i = 1, 2, \dots, r-1)$, $1 \leqq i+k-j \leqq d_r \leqq \frac{T}{h}$ *and* $d = d_1 + d_2 + \ldots + d_r$;

R_3: *Summation over* s *from* r *to* $\left[\frac{t}{T} - (i+j-d)\frac{h}{T}\right]$;

R_4: *Multiple summation over* $d_1, d_2, \dots, d_r$ *such that* $1 \leqq d_i \leqq \frac{T}{h}$, $(i = 1, 2, \dots, r)$, $1 \leqq d_r \leqq \min\left(i+k-j, \frac{T}{h}\right)$ *and* $d = d_1 + d_2 + \ldots + d_r$.

$\binom{n}{r} = 0$ *for* $r < 0$ *and* $[x]$ *denotes the largest integer in* x.

PROOF. Lattice paths stipulated in Theorem 1 starting from the point $(k, 0)$ to $(t/h - j + k, t/h - i)$ will have i horizontal steps, j vertical steps,

$\left(\frac{t}{h}-i-j\right)$ diagonal steps and will encounter vacations of fixed duration $\frac{T}{h}$ at each of the contact points with the barrier. On removing the diagonal steps the lattice paths will reduce to skeleton lattice paths from $(k,0)$ to $(i+k,j)$ having i horizontal steps and j vertical steps. Vacations encountered are of two types.

I. Vacations with no arrivals.

II. Vacations with at least one arrival.

In the sequence of the vacations encountered at each contact point only the last vacation will be of II$^{\text{nd}}$ type. Consequently, at the i^{th} contact point with the barrier d_i arrivals could occur in a vacation of II$^{\text{nd}}$ type in

$$\binom{T/h}{d_i} \qquad (i=1,2,\ldots,r) \tag{3.3}$$

ways. Taking the total number of vacations to be s and r the number of touches with the barrier there will be $(s-r)$ vacations of I$^{\text{st}}$ type which can be distributed at r contact points on the barrier by using "balls into cells technique" (see [6]) in

$$\binom{s-1}{r-1} \tag{3.4}$$

ways. Further $\left(\frac{t}{h}-\frac{sT}{h}-i-j+d\right)$ stays or diagonal steps in a busy period occur at $(i+j-d+1)$ vertices of the skeleton lattice path in

$$\binom{\frac{t}{h}-\frac{sT}{h}}{i+j-d} \tag{3.5}$$

ways. The number of skeleton lattice paths envisaged in the Theorem consisting of horizontal and vertical steps only, can be obtained from (2.3) on replacing m by $i+k$ and n by j given by

$$\left[\binom{i+j-d}{j}-\binom{i+j-d}{i+k}\right] \tag{3.6}$$

which on multiplying by (3.3), (3.4) and (3.5) gives the number of lattice paths stipulated. This number multiplied by appropriate probabilities in (2.1) and (2.2), summed over $r, d_1, d_2, \ldots, d_r$ and s yields (3.1).

(3.2) can be proved similarly.

THEOREM 2.

$$P_{i,j;0}\left(\frac{t}{h},\frac{T}{h}\right)$$

$$= \sum_{r=1}^{j}\sum_{R_1}\sum_{R_2}\sum_{R_3}\binom{s-1}{r-1}\prod_{q=1}^{r}\binom{T/h}{d_q}\left[\binom{i+j-d-1}{j}-\binom{i+j-d-1}{i}\right]$$

$$\times\sum_{R_5}\left[\begin{matrix} t/h-(sT/h)-m-1 \\ i+j-d-1\end{matrix}\right](\lambda h+o(h))^i(\mu h+o(h))^j \tag{3.7}$$

$$\times(1-\lambda h+o(h))^{m+(sT/h)-d}$$

$$\times(1-\lambda h-\mu h+o(h))^{(t/h)-(sT/h)-(i+j-d)-m}$$

and

$${}_1P_{i,j;0}\left(\frac{t}{h},\frac{T}{h}\right)$$

$$= \sum_{r=1}^{j}\sum_{R_1}\sum_{R_4}\sum_{R_3}\binom{s-1}{r-1}\prod_{q=1}^{r}\binom{T/h}{d_q}\left[\binom{i+j-d-1}{i-d_r}-\binom{i+j-d-1}{i}\right]$$

$$\times\sum_{R_5}\left[\begin{matrix} t/h-(sT/h)-m-1 \\ i+j-d-1\end{matrix}\right](\lambda h+o(h))^i(\mu h+o(h))^j \tag{3.8}$$

$$\times(1-\lambda h+o(h))^{m+(sT/h)-d}$$

$$\times(1-\lambda h-\mu h+o(h))^{(t/h)-(sT/h)-m-(i+j-d)},$$

where

R_1: *Summation over d from r to* $\min\left(r\frac{T}{h}, i-1\right)$;

R_2: *Multiple summation over $d_1, d_2, \ldots, d_r$ such that* $1 \leqq d_i \leqq \frac{T}{h}$, $(i=1,2,\ldots,r)$,

$$1 \leqq i-j \leqq d_r \leqq \frac{T}{h} \quad \textit{and} \quad d = d_1+d_2+\cdots+d_r;$$

R_3: *Summation over s from r to* $\left[\frac{t}{T}-(i+j-d)\frac{h}{T}\right]$;

R_4: *Multiple summation over $d_1, d_2, \ldots, d_r$ such that* $1 \leqq d_i \leqq \frac{T}{h}$, $(i=1,2,\ldots,r)$,

$$d_r \leqq \min\left(i-j, \frac{T}{h}\right) \quad \textit{and} \quad d = d_1+d_2+\cdots+d_r;$$

R_5: *Summation over m from 0 to* $\left[\frac{t}{h} - \frac{sT}{h} - (i+j-d)\right]$.

Theorem 2 can easily be proved by following decomposition of stays (or diagonal steps) as in Theorem 1 and Lemma 2.

REMARK 1. For $\frac{T}{h} = 1$, which in turn reduces the value of $d_1, d_2, \ldots, d_r$ to 1 each, results stated in Theorems 1 and 2 yield those for the case of ordinary discretized $M/M/1$ queue.

For $i + k = j$ the last contact point coincides with the final state reached (i.e. a vacation is encountered necessarily) hence this results in a system ending with a vacation. Further in some cases this situation may lead to an incomplete vacation. Let T_1 denote the duration of an incomplete vacation and n the number of arrivals in time T_1. Define ${}_vP_{n,j;k}\left(\frac{t}{h}, \frac{T}{h}, \frac{T_1}{h}\right)$ the probability that the discretized $M/M/1(T)$ process ending with vacation, encountering j departures and an incomplete vacation of duration $\frac{T_1}{h}$ attains queue length n in $\frac{t}{h}$ slots, starting initially with k (≥ 0) units. This probability summed over $\frac{T_1}{h}$, yields the probability of occurrence of $j+n-k$ arrivals, j departures in $\frac{t}{h}$ time slots and ending with a vacation denoted by ${}_vP^*_{j+n-k,j;k}\left(\frac{t}{h}, \frac{T}{h}\right)$. The corresponding probabilities for the continuous time model are denoted by ${}_vP^*_{n,j;k}(t, T, T_1)$ and ${}_vP^*_{j+n-k,j;k}(t, T)$, respectively.

THEOREM 3. *If the system is in vacation at time $\frac{t}{h}$, then for $k > 0$,*

$$(3.9) \qquad {}_vP_{j+n-k,j;k}\left(\frac{t}{h}, \frac{T}{h}\right) = \sum_{T_1/h}^{T/h} {}_vP_{n,j;k}(t/h, T/h, T_1/h)$$

$$(3.10) \qquad \begin{aligned} &= \sum_{\frac{T_1}{h}=0}^{T/h} \sum_{r=1}^{j-k+1} \sum_{R_1} \sum_{R_2} \sum_{R_3} \binom{s}{r-1} \prod_{q=1}^{r} \binom{T/h}{d_q} \binom{T_1/h}{n} \\ &\times \left[\binom{2j-k-d-1}{j-1} - \binom{2j-k-d-1}{j}\right] \\ &\times \left[\begin{matrix} t/h - (s/T/h) - (T_1/h) - 1 \\ 2j-k-d-1 \end{matrix}\right] (\lambda h + o(h))^{j+n-k} \\ &\times (\mu h + o(h))^j (1 - \lambda h + o(h))^{(sT/h)-(T_1/h)-d-n} \\ &\times (1 - \lambda h - \mu h + o(h))^{(t/h)-(sT/h)-(T_1/h)-(2j-k-d)} \end{aligned}$$

and for $k=0$

$$ {}_vP_{n,j;0}\Big(\frac{t}{h},\frac{T}{h}\Big)=\sum_{\frac{T_1}{h}=0}^{T/h} {}_vP_{n,j;0}\Big(\frac{t}{h},\frac{T}{h},\frac{T_1}{h}\Big) \tag{3.11}$$

$$\begin{aligned}
&=\sum_{\frac{T_1}{h}=0}^{T/h}\sum_{r=1}^{j-k+1}\sum_{R_6}\sum_{R_2}\sum_{R_4}\binom{s}{r-1}\prod_{q=1}^{r-1}\binom{T/h}{d_q}\binom{T_1/h}{n}\\
&\quad\times\left[\binom{2j-d-2}{j-1}-\binom{2j-d-2}{j}\right]\\
&\quad\times\sum_{R_5}\left[\begin{matrix}t/h-(sT/h)-(T_1/h)-m-2\\ 2j-d-2\end{matrix}\right](\lambda h+o(h))^{j+n}\\
&\quad\times(\mu h+o(h))^j(1-\lambda h+o(h))^{(sT/h)-(T_1/h)-d-n+m}\\
&\quad\times(1-\lambda h-\mu h+o(h))^{(t/h)-(sT/h)-(T_1/h)-m-2j+d},
\end{aligned}\tag{3.12}$$

where

R_1: *Summation over* d *from* $(r-1)$ *to* $\min\Big((r-1)\frac{T}{h},j\Big)$;

R_2: *Multiple summation over* $d_1,d_2,\dots,d_{r-1}$ *such that* $1\leqq d_i\leqq\frac{T}{h}$, $(i=1,2,\dots,r-1)$, *and* $d=d_1+d_2+\dots+d_{r-1}$;

R_3: *Summation over* s *from* $(r-1)$ *to* $\Big[\frac{(t-T_1)}{T}-(2j+k-d)\frac{h}{T}\Big]$;

R_4: *Summation over* s *from* $(r-1)$ *to* $\Big[\frac{(t-T_1)}{T}-(2j-d)\frac{h}{T}\Big]$;

R_5: *Summation over* m *from* 0 *to* $\Big[\frac{(t-T_1)}{h}-\frac{sT}{h}-2j+d\Big]$;

R_6: *Summation over* d *from* $(r-1)$ *to* $\min\Big((r-1)\frac{T}{h}-1,(j-1)\Big)$.

Proof of Theorem 3 can be constructed by following the decomposition of stays (or diagonal steps) as in Theorems 1–2 and then using Lemma 3.

REMARK 2. For $\frac{T_1}{h}=0$, $\frac{T}{h}=1$, which in turn reduces the values of $d_1,d_2,\dots,d_{r-1}$ to 1 each and $n=0$, results stated in Theorem 3 would yield those for the case of ordinary discretized $M/M/1$ queue.

THEOREM 4. *For* $k>0$

$$\begin{aligned}
Q_{i,j;k}(t/h)&=\left[\binom{i+j}{i}-\binom{i+j}{i+k}\right]\binom{t/h}{i+j}(\lambda h+o(h))^i\\
&\quad\times(\mu h+o(h))^j(1-\lambda h-\mu h+o(h))^{(t/h)-i-j}
\end{aligned}\tag{3.13}$$

and

$$Q_{i,j;0}(t/h) = \sum_{m=0}^{[t/h-i-j]} \left[\binom{i+j-1}{i-1} - \binom{i+j-1}{i} \right] \binom{t/h-m-1}{i+j-1} \times (\lambda h + o(h))^i (\mu h + o(h))^j \times (1-\lambda h - \mu h + o(h))^{(t/h)-i-j-m} (1-\lambda h + o(h))^m. \tag{3.14}$$

REMARK 3. The probability ${}_TP_{i,j;k}(t/h, T/h)$, $k > 0$ can be obtained by considering the following two mutually exclusive cases:

(a) System encounters no vacation in t/h time epoch;

(b) System encounters at least one vacation in t/h time epoch.

The probability for the former case is given by (3.13) and for the later case is given by (3.1) or (3.2), accordingly as the queue length attained is < or $\geqq 1$, the number of arrivals encountered in the last sequence of vacations. Thus by adding up the probabilities for the two cases we get the required probability which is given by

$$\left[\binom{i+j}{i} - \binom{i+j}{i+k} \right] \binom{t/h}{i+j} (\lambda h + o(h))^i (\mu h + o(h))^j \times \times (1-\lambda h - \mu h + o(h))^{t/h-i-j}$$

$$+ \sum_{r=1}^{j-k+1} \sum_{R_1} \sum_{R_2} \sum_{R_3} \binom{s-1}{r-1} \prod_{q=1}^{r} \binom{T/h}{d_q} \mathbb{A}(i,j,k,d,d_r) \binom{t/h - sT/h}{i+j-d} \times \times (\lambda h + o(h))^i (\mu h + o(h))^j (1-\lambda h + o(h))^{sT/h-d} (1-\lambda h - \mu h + o(h))^{t/h - sT/h - i-j+d}, \tag{3.15}$$

where R_1 and R_3 are summations as defined in Theorem 1 and

R_2: Multiple summation over $d_1, d_2, \dots, d_r$, $1 \leqq d_i \leqq T/h$ $(i = 1, 2, \dots, r)$ and $d = d_1 + d_2 + \dots + d_r$. Further,

$$\mathbb{A}(i,j,k,d,d_r) = \begin{cases} \left[\binom{i+j-d}{j} - \binom{i+j-d}{i+k} \right], & i+k-j < d_r, \\ \left[\binom{i+j-d}{i+k-d_r} - \binom{i+j-d}{i+k} \right], & i+k-j \geqq d_r. \end{cases}$$

Similarly, we can obtain for the case $k = 0$.

4. Transient probabilities

Lemma 4 stated below gives limiting results which are employed to compute the transient probabilities for the continuous time model from its discrete analogue.

LEMMA 4. *If*

(4.1)
$$\psi_1\Big(\frac{t}{h},\frac{sT}{h},p\Big)=h^p\sum_{m=0}^{[\frac{t}{h}-\frac{sT}{h}-p]}\binom{\frac{t}{h}-\frac{sT}{h}-m-1}{p-1}(1-\lambda h+o(h))^{m+(sT/h)-d}\times(1-\lambda h-\mu h+o(h))^{(t/h)-(sT/h)-p-m},$$

(4.2)
$$\psi_2\Big(\frac{t}{h},\frac{sT}{h},\frac{T_1}{h},p\Big)=h^{p-1}\sum_{m=0}^{[\frac{t}{h}-\frac{sT}{h}-\frac{T_1}{h},p]}\binom{\frac{t}{h}-\frac{sT}{h}-\frac{T_1}{h}-m-2}{p-2}\times\times(1-\lambda h+o(h))^{m+(sT/h)+(T_1/h)-d}\times(1-\lambda h-\mu h+o(h))^{(t/h)-(sT/h)-(T_1/h)-m-(2i-d)},$$

$$\psi_3(r,p,t,T)=T^r\sum_{s=r}^{[\frac{t}{T}]}\binom{s-1}{r-1}\frac{(t-sT)^p}{p!}e^{-\mu(t-sT)},\tag{4.3}$$

(4.4)
$$\psi_4(r,p,t,T,T_1)=T^r\sum_{s=r-1}^{[\frac{t-T_1}{T}]}\binom{s}{r-1}\frac{(t-sT-T_1)^p}{p!}e^{-\mu(t-sT-T_1)},$$

then

$$\operatorname*{Lim}_{h\to0}\psi_1\Big(\frac{t}{h},\frac{sT}{h},p\Big)=\frac{e^{-\lambda t}}{\mu^p}\Big(1-e^{-\mu(t-sT)}\sum_{n=0}^{p-1}\frac{(\mu(t-sT))^n}{n!}\Big),\tag{4.5}$$

$$\operatorname*{Lim}_{h\to0}\psi_2\Big(\frac{t}{h},\frac{sT}{h},\frac{T_1}{h},p\Big)=\frac{e^{-\lambda t}}{\mu^{p-1}}\Big(1-e^{-\mu(t-sT-T_1)}\sum_{n=0}^{p-2}\frac{(\mu(t-sT-T_1))^n}{n!}\Big),\tag{4.6}$$

$$\operatorname*{Lim}_{T\to0}\psi_3(r,p,t,T)=\sum_{m=0}^{r-1}\frac{(t)^{r-1-m}(-1)^m}{(r-1-m)!}\binom{m+p}{p}\Big(\frac{1}{\mu}\Big)^{m+p+1}\times\times\Big(1-\sum_{n=0}^{m+p}e^{-\mu t}\frac{(\mu t)^n}{n!}\Big),\tag{4.7}$$

$$\begin{aligned}\lim_{\substack{T\to 0\\ T_1\to 0}} \psi_4(r,p,t,T,T_1) = \sum_{m=0}^{r-1} \frac{(t)^{r-1-m}(-1)^m}{(r-1-m)!}\binom{m+p}{p}\left(\frac{1}{\mu}\right)^{m+p+1}\times \\ \times\left(1-\sum_{n=0}^{m+p} e^{-\mu t}\frac{(\mu t)^n}{n!}\right).\end{aligned} \tag{4.8}$$

Result (4.5) can be proved by differentiating partially a finite geometric series of the type $\{x^m y^{\frac{t}{h}-\frac{sT}{h}-m-1}\}$, m varies from 0 to $\left[\frac{t}{h}-\frac{sT}{h}-p\right]$, w.r.t. x $(p-1)$ times where $x=1-\lambda h+o(h)$, $y=1-\lambda h-\mu h+o(h)$ and taking limit as $h\to 0$. Results (4.6)–(4.8) can be proved similarly by considering appropriate series (see [1]).

4.1. *Bivariate distribution of the number of arrivals and departures up to time t*

THEOREM 5. *For $k>0$*

$$\begin{aligned}&P^*_{i,j;k}(t,T) \\ &= \sum_{r=1}^{j-k+1}\sum_{R_1}\sum_{R_2}\sum_{R_3}\left\{e^{-\lambda T(s-r)}\binom{s-r}{r-1}\prod_{q=1}^{r}\left(\frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!}\right)\right\}\times \\ &\times\left\{(\lambda(t-sT))^{i-d}(\mu(t-sT))^j\frac{e^{-(\lambda+\mu)(t-sT)}}{(i+j-d)!}\times\right. \\ &\left.\times\left[\binom{i+j-d}{j}-\binom{i+j-d}{i+k}\right]\right\},\end{aligned} \tag{4.9}$$

and for $k>0$

$$\begin{aligned}&P^*_{i,j;k}(t,T) \\ &= \sum_{r=1}^{j-k+1}\sum_{R_1}\sum_{R_4}\sum_{R_3}\left\{e^{-\lambda T(s-r)}\binom{s-1}{r-1}\prod_{q=1}^{r}\left(\frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!}\right)\right\}\times \\ &\times\left\{(\lambda(t-sT))^{i-d}(\mu(t-sT))^j\frac{e^{-(\lambda+\mu)(t-sT)}}{(i+j-d)!}\times\right. \\ &\left.\times\left[\binom{i+j-d}{i+k-d_r}-\binom{i+j-d}{i+k}\right]\right\},\end{aligned} \tag{4.10}$$

where

R_1: *Summation over d from r to i;*

R_2: *Multiple summation over* $d_1, d_2, \ldots, d_r$ *such that* $1 \leqq i+k-j \leqq d_r$ *and* $d = d_1 + d_2 + \cdots + d_r$;

R_3: *Summation over* s *from* r *to* $\left[\frac{t}{T}\right]$;

R_4: *Multiple summation over* $d_1, d_2, \ldots, d_r$ *such that* $i+k-j \geqq d_r$, *and* $d = d_1 + d_2 + \cdots + d_r$.

THEOREM 6.

$$\begin{aligned} &P^*_{i,j;0}(t,T) \\ &= \sum_{r=1}^{j} \sum_{R_1} \sum_{R_2} \sum_{R_3} \left\{ e^{-\lambda T(s-r)} \binom{s-1}{r-1} \prod_{q=1}^{r} \left(\frac{e^{-\lambda T} (\lambda T)^{d_q}}{(d_q)!} \right) \right\} \times \\ &\quad \times \left\{ (e^{-\lambda(t-sT)} \left(\frac{\lambda}{\mu}\right)^{i-d} \left(1 - e^{-\mu(t-sT)} \sum_{n=0}^{i+j-d-1} \frac{(\mu(t-sT))^n}{n!} \right) \times \right. \\ &\quad \left. \times \left[\binom{i+j-d-1}{i} - \binom{i+j-d-1}{j} \right] \right\}, \end{aligned} \tag{4.11}$$

and

$$\begin{aligned} &{}_1P^*_{i,j;0}(t,T) \\ &= \sum_{r=1}^{j} \sum_{R_1} \sum_{R_4} \sum_{R_3} \left\{ e^{-\lambda T(s-r)} \binom{s-1}{r-1} \prod_{q=1}^{r} \left(\frac{e^{-\lambda T} (\lambda T)^{d_q}}{(d_q)!} \right) \right\} \times \\ &\quad \times \left\{ (e^{-\lambda(t-sT)} \left(\frac{\lambda}{\mu}\right)^{i-d} \left(1 - e^{-\mu(t-sT)} \sum_{n=0}^{i+j-d-1} \frac{(\mu(t-sT))^n}{n!} \right) \times \right. \\ &\quad \left. \times \left[\binom{i+j-d-1}{i-d_r} - \binom{i+j-d-1}{i} \right] \right\}, \end{aligned} \tag{4.12}$$

where

R_1: *Summation over* d *from* r *to* $i-1$;

R_2: *Multiple summation over* $d_1, d_2, \ldots, d_r$ *such that* $1 \leqq i-j < d_r$, $d_i \geqq 1$ $(i = 1, 2, \ldots, r)$, *and* $d = d_1 + d_2 + \cdots + d_r$;

R_3: *Summation over* s *from* r *to* $\left[\frac{t}{T}\right]$;

R_4: *Multiple summation over* $d_1, d_2, \ldots, d_r$ *such that* $d_r \leqq i-j$, $d_i \geqq 1$, $(i = 1, 2, \ldots, r)$ *and* $d = d_1 + d_2 + \cdots + d_r$.

(4.9) can be obtained on substituting the value from (3.1) and taking the limit as $h \to 0$. Similarly, (4.10) can be proved on using (3.2). (4.11) and (4.12) can be proved as a limiting case of Theorem 2 on using (4.6).

REMARK 4. Results in Theorems 5–6 can be interpreted probabilistically by tracing the dissection of time t, through the following bijective transformation in Figure 3. Collect the segments of $d_1, d_2, \ldots, d_r$ horizontal steps after $1^{\text{st}}, 2^{\text{nd}}, \ldots, r^{\text{th}}$ contact points, respectively. Concatenate these segments in succession and append the resulting path segment at $(k, 0)$. Collate the remaining truncated segments and append the path segment at $(k+d, 0)$. The resulting path is shown in Figure 4.

Considering the division of time t as suggested by Figure 3 referring to (4.9) for fixed values of $r, d_1, d_2, \ldots, d_r, T$ and s the two terms within parentheses inside the summations are the probabilities of occurrence of $(s-r)$ vacations of first type and r vacations of second type with d_i arrivals in the i^{th} vacation $(i = 1, 2, \ldots, r)$ and of queue length being $(i + k - j)$ in time $(t - sT)$ starting with $k + d$ units without being empty in between (see [11]).

REMARK 5. To obtain the transient solution in continuous time for ordinary $M/M/1$ queue from Theorem 5, we will have to pass to the limit $T \to 0$ and set $d_1 = d_2 = \cdots = d_r = 1$, an obvious implication of the former (Greenberg and Greenberg [7]). Under same conditions Theorem 6, on using (4.7), reduces to

$$
\begin{aligned}
\lim_{T\to 0} P^*_{i,j;0}(t,T) &= e^{-\lambda t}\left(\frac{\lambda}{\mu}\right)^i \sum_{r=1}^{j} \left[\binom{i+j-r-1}{i-1} - \binom{i+j-r-1}{i}\right] \times \\
&\quad \times \left\{\frac{(\mu t)^r}{r!} - \sum_{n=0}^{i+j-r-1} \sum_{m=0}^{r-1} \binom{m+n}{n} \frac{(-1)^m}{(\mu t)^{m+1-r}} \times \right. \\
&\quad \left. \times \frac{1}{(r-1-m)!}\left(1 - e^{-\mu t} \sum_{a=0}^{m+n} \frac{(\mu t)^a}{a!}\right)\right\} \\
&= e^{-\lambda t}\left(\frac{\lambda}{\mu}\right)^i \sum_{r=1}^{j} \sum_{m=0}^{r} \frac{(-1)^m (\mu t)^{r-m}}{(r-m)!} \binom{i+j-r+m-1}{m} \times \\
&\quad \times \left[\binom{i+j-r-1}{i-1} - \binom{i+j-r-1}{i}\right] \times \\
&\quad \times \left(1 - e^{-\mu t} \sum_{a=0}^{m+i+j-r-1} \frac{(\mu t)^a}{a!}\right),
\end{aligned}
\tag{4.13}
$$

which checks with the result $P_{ij}(t)$ (see (10) in Pegden and Rosenshine [16] of ordinary $M/M/1$ queues).

THEOREM 7. *If the system is in vacation at time t, then for $n \geqq 0$, $j \geqq k > 0$*

$${}_vP^*_{j+n-k,j;k}(t,T)$$

$$
\begin{aligned}
&= \int_0^T \sum_{r=1}^{j-k+1} \sum_{R_1}\sum_{R_2}\sum_{R_3} \left\{ e^{-\lambda T(s-r+1)} \binom{s}{r-1} \frac{e^{-\lambda T_1}(\lambda T_1)^n}{n!} \times \right. \\
&\quad \times \prod_{q=1}^{r-1} \frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!} \Bigg\} \Bigg\{ \left[\binom{2j-k-d-1}{j-1} - \binom{2j-k-d-1}{j} \right] \times \\
&\quad \times \frac{e^{-(\lambda+\mu)(t-sT-T_1)}}{(2j-k-d-1)!} \mu \times \\
&\quad \left. \times (\lambda(t-sT-T_1))^{j-d-k} (\mu(t-sT-T_1))^{j-1} \right\} dT_1
\end{aligned} \tag{4.14}
$$

and for $n \geqq 0$

$$
\begin{aligned}
&{}_vP^*_{j+n,j;0}(t,T) \\
&= \int_0^T \sum_{r=1}^{j} \sum_{R_4}\sum_{R_2}\sum_{R_3} \left\{ e^{-\lambda T(s-r+1)} \binom{s}{r-1} \frac{e^{-\lambda T_1}(\lambda T_1)^n}{n!} \times \right. \\
&\quad \times \prod_{q=1}^{r-1} \frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!} \Bigg\} \Bigg\{ \left(\frac{\lambda}{\mu}\right)^{j-d} \mu e^{-(\lambda+\mu)(t-sT-T_1)} \times \\
&\quad \times \left[\binom{2j-d-2}{j-1} - \binom{2j-d-2}{j} \right] \times \\
&\quad \left. \times \left(1 - e^{-\mu(t-sT-T_1)} \sum_{m=0}^{2j-d-2} \frac{(\mu(t-sT-T_1))^m}{m!} \right) \right\} dT_1,
\end{aligned} \tag{4.15}
$$

where

R_1: *Summation over* d *from* $(r-1)$ *to* j;

R_2: *Multiple summation over* $d_1, d_2, \ldots, d_{r-1}$ *such that* $d_i \geqq 1$ $(i = 1, 2, \ldots, r-1)$, *and* $d = d_1 + d_2 + \cdots + d_{r-1}$;

R_3: *Summation over* s *from* $(r-1)$ *to* $\left[\frac{(t-T_1)}{T}\right]$;

R_4: *Summation over* d *from* $(r-1)$ *to* j.

Theorem 7 can be proved as a limiting case of Theorem 3 on using (4.6).

REMARK 6. In order to obtain probabilistic interpretations of (4.14) and (4.15), referring back to Figure 4, we observe that for fixed values of $r, d_1, d_2, \ldots, d_{r-1}$, the number of complete vacations is s each of duration T and an incomplete vacation of duration T_1. The two terms within parentheses inside the summations are the probabilities of occurrence of $(s-r)$ vacations of first type, an incomplete vacation of duration T_1 with n arrivals and r

vacations of second type with d_i arrivals in the i^{th} vacation $(i=1,2,\dots,r-1)$ and of dropping the level $k+d$ to level zero in time $(t-sT-T_1)$ encountering j services without being empty in-between (see [11]).

REMARK 7. In order to obtain transient solutions in continuous time for ordinary $M/M/1$ queues from (3.9) and (3.11) we pass to the limit as $h\to 0$ and $T\to 0$ by using (4.8), (4.6) and set $d_1=d_2=\cdots=d_{r-1}=1$, $T_1=0$ and $n=0$ an obvious implication of the later (see [7], [16]).

THEOREM 8. *For* $k>0$

$$Q^*_{i,j;k}(t)=\left[\binom{i+j}{i}-\binom{i+j}{i+k}\right]e^{-(\lambda+\mu)t}(\lambda t)^i(\mu t)^j \tag{4.16}$$

and for $k=0$

$$\begin{aligned}Q^*_{i,j;0}(t)=&\left[\binom{i+j-1}{i-1}-\binom{i+j-1}{i}\right]e^{-\lambda t}\left(\frac{\lambda}{\mu}\right)^i\times\\&\times\left(1-e^{-\mu t}\sum_{n=0}^{i+j-1}\frac{(\mu t)^n}{n!}\right).\end{aligned}\tag{4.17}$$

REMARK 8. ${}_TP^*_{i,j;k}(t,T)$ can be obtained as a limiting case of ${}_TP^*_{i,j;k}(t/h,T/h)$ computed in Remark 3.

For $k>0$

$$\begin{aligned}&{}_TP^*_{i,j;k}(t,T)\\&=\left[\binom{i+j}{i}-\binom{i+j}{i+k}\right]e^{-(\lambda+\mu)t}(\lambda t)^i(\mu t)^j\\&+\sum_{r=1}^{j-k+1}\sum_{R_1}\sum_{R_2}\sum_{R_3}\left\{e^{-\lambda T(s-r)}\binom{s-r}{r-1}\prod_{q=1}^{r}\frac{e^{-\lambda T}(\lambda T)^{d_q}}{d_q!}\right\}\times\\&\times\left\{\frac{(\lambda(t-sT))^{i-d}(\mu(t-sT))^j}{(i+j-d)!}e^{-(\lambda+\mu)(t-sT)}\mathbb{A}(i,j,k,d,d_r)\right\},\end{aligned}\tag{4.18}$$

where R_1 and R_3 are summations as defined in Theorem 5, $\mathbb{A}(i,j,k,d,d_r)$ is defined in (3.15),

R_2: Multiple summation over $d_1,d_2,\dots,d_r$, $d_i\ (\geqq 1)$, $i=1,2,\dots,r$ and $d=d_1+d_2+\cdots+d_r$.

4.2. *Queue length distribution*

The probability that starting initially with k units $(k\geqq 0)$ there are n units in the system at time t after having been empty at least once in-between, denoted by $P^*_{n;k}(t,T)$ is given by

$$P^*_{n;k}(t,T) = \sum_{j=k}^{\infty} P^*_{j+n-k,j;k}(t,T). \tag{4.19}$$

In particular, (4.19) on using (4.9), reduces to

$$\begin{aligned} P^*_{n;k}(t,T) = \sum_{r=1}^{\infty}\sum_{R_1}\sum_{R_2}\sum_{R_3}\left\{e^{-\lambda T(s-r)}\binom{s-1}{r-1}\prod_{q=1}^{r}\left(\frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!}\right)\right\}\times \\ \times\left\{e^{-(\lambda+\mu)(t-sT)}\left(\frac{\lambda}{\mu}\right)^{\frac{n-k-d}{2}}\times \right. \\ \left. \times\left[I_{n-k-d}(2\sqrt{\lambda\mu}(t-sT)) - I_{n+k+d}(2\sqrt{\lambda\mu}(t-sT))\right]\right\}, \end{aligned} \tag{4.20}$$

where summations are as defined in respective Theorems.

REMARK 9. In (4.20) the term in the last parentheses denotes the probability that starting with $(k+d)$ units at time $t=0$, there are n (>0) units in the system at time $(t-sT)$ in the presence of an absorbing barrier at the origin (see [11]). Thus $P^*_{n;k}(t,T)$ is amenable to identical probabilistic interpretation.

The probability that $M/M/1(T)$ system ending with vacation attains queue length n in time t starting initially with k $(\geqq 0)$ units, denoted by ${}_vP^*_{n;k}(t,T)$ is given by

$${}_vP^*_{n;k}(t,T) = \sum_{j=k}^{\infty} {}_vP^*_{j+n-k,j;k}(t,T). \tag{4.21}$$

Then for $k>0$, $n\geqq 0$,

$$\begin{aligned} {}_vP^*_{n;k}(t,T) = \sum_{r=1}^{\infty}\int_0^T\sum_{R_1}\sum_{R_2}\sum_{R_3}\left\{e^{-\lambda T(s-r+1)}\binom{s}{r-1}\frac{e^{-\lambda T_1}(\lambda T_1)^n}{n!}\right. \\ \left.\times\prod_{q=1}^{r-1}\left(\frac{e^{-\lambda T}(\lambda T)^{d_q}}{(d_q)!}\right)\right\}\left\{e^{-(\lambda+\mu)(t-sT-T_1)}\left(\frac{k+d}{t-sT-T_1}\right)\times \right. \\ \left.\times\left(\frac{\lambda}{\mu}\right)^{-\frac{(k+d)}{2})} I_{k+d}(2\sqrt{\lambda\mu}(t-sT-T_1))\right\}dT_1, \end{aligned} \tag{4.22}$$

where summations are as defined in respective theorems.

REMARK 10. In (4.22) the terms in the last parentheses denote the probability that starting with $(k+d)$ units at time $t=0$, the system becomes empty at time $(t - sT - T_1)$ (see [17]). (4.22) can then be interpreted in a routine way.

REMARK 11. $P^*_{i,j;k}(t,T)$ and $P^*_{i,j;k}(t,T,T_1)$ are expressions involving only finite sums. Hence computations of $P^*_{n;k}(t,T)$ and $P^*_{0;k}(t,T,T_1)$ could be comfortably handled from (4.19) and (4.21) since the infinite sums does not involve Bessel functions explicitly.

ACKNOWLEDGEMENT. The authors express their thanks to the referee for his valuable comments.

REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. A. (eds.), *Handbook of mathematical functions, with formulas, graphs and mathematical tables*, Dover Publications, New York, 1966. *MR* **34** #8606

[2] BÖHM, W. and MOHANTY, S. G., The transient solution of $M/M/1$ queues under (M,N)-policy. A combinatorial approach, *J. Statist. Plann. Inference* **34** (1993), 23–33. *MR* **94c**:60152

[3] CSÁKI, E. and VINCZE, I., On some problems connected with the Galton-test, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6** (1961), 97–109. *MR* **26** #3138

[4] DOSHI, B. T., Queueing systems with vacations – a survey, *Queueing Systems Theory Appl.* **1** (1986), 29–66. *MR* **89b**:60212

[5] DUA, S., KHADILKAR, S. and SEN, K., A modified random walk in the presence of partially reflecting barriers, *J. Appl. Probability* **13** (1976), 169–175. *MR* **53** #4235

[6] FELLER, W., *An introduction to probability theory and its applications*, Vol. 1, 3rd ed., John Wiley, New York – London, 1968. *MR* **37** #3604

[7] GREENBERG, H. and GREENBERG, I., The number served in a queue, *Operations Res.* **14** (1966), 137–144. *MR* **32** # 4756

[8] HEYMAN, D. P., The T-policy for the $M/G/1$ queue, *Management Sci.* **23** (1977), 775–778. *Zbl* **357**. 60022

[9] SEN, K., On some combinatorial relations concerning the symmetric random walk, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9** (1965), 335–357. *MR* **33** #6715

[10] SEN, K. and AHUJA, Y. L., On certain probability model in random walk, *Calcutta Statist. Assoc. Bull.* **21** (1972), 71–76. *MR* **50** #1341

[11] SEN, K. and JAIN, J. L., Combinatorial approach to Markovian queueing models, *J. Statist. Plann. Inference* **34** (1993), 269–279. *MR* **93j**:60148

[12] SEN, K., JAIN, J. L. and GUPTA, J. M., Lattice path approach to transient solution of $M/M/1$ with $(0,k)$ control policy, *J. Statist. Plann. Inference* **34** (1993), 259–267. *MR* **93j**:60147

[13] MEISLING, T., Discrete-time queueing theory, *Operations Res.* **6** (1958), 96–105. *MR* **19**, 1092e

[14] MOHANTY, S. G., *Lattice path counting and applications*, Probability and Mathematical Statistics, Academic Press, New York – London, 1979. *MR* **81f**:60020

[15] MOHANTY, S. G. and PANNY, W., A discrete-time analogue of the $M/M/1$ queue and the transient solution: a geometric approach, *Sankhyā Ser. A* **52** (1990), 364–370. *MR* **93f**:60146

[16] PEGDEN, C. D. and ROSENSHINE, M., Some new results for the $M/M/1$ queue, *Management Sci.* **28** (1982), 821–828. *MR* **83i**:60115

[17] SAATY, T. L., *Elements of queueing theory, with applications*, McGraw Hill, New York – Toronto – London, 1961. *MR* **24** #A3010

[18] TAKÁCS, L., Combinatorial methods in the theory of queues, *Rev. Inst. Internat. Statist.* **32** (1964), 207–219. *MR* **31** #2770
[19] TAKÁCS, L., *Combinatorial methods in the theory of stochastic processes*, John Wiley, New York – London – Sydney, 1967. *MR* **36** #947
[20] TAKAGI, H., *Queueing analysis: a foundation of performance evaluation, Vol. 1: Vacation and priority systems.* Part 1, North-Holland Publ. Co., Amsterdam, 1991. *MR* **93e**:60188
[21] TEGHEM, J., JR., Control of the service process in a queueing system, *European J. Oper. Res.* **23** (1986), 141–158. *MR* **88e**:90035
[22] WHITTLE, P., *Optimization over time, Vol. 2: Dynamic programming and stochastic control*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley, Chichester, 1983. *MR* **85e**:90062

(Received August 12, 1992; in revised form October 16, 1996)

DEPARTMENT OF STATISTICS
UNIVERSITY OF DELHI
FACULTY OF MATHEMATICAL SCIENCES
IND–110007 DELHI
INDIA
manju_agarwal@hotmail.com

Studia Scientiarum Mathematicarum Hungarica 34 (1998), 475–486

ON THE SIMULTANEOUS APPROXIMATION OF FUNCTIONS AND THEIR DERIVATIVES BY THE GAMMA OPERATORS

S. GUO and Q. QI

Abstract

In this paper, we investigate the degree of approximation by Gamma operator for functions whose derivatives have only discontinuity points of the first kind on $[0,\infty)$. Our estimates are essentially the best possible.

1. Introduction

Let f be a function defined on the interval $[0,\infty)$. The Gamma operator $G_n(f,x)$ is defined as follows:

$$G_n(f,x)=\int_0^\infty g_n(x,u)f\left(\frac{n}{u}\right)du, \tag{1.1}$$

where $g_n(x,u)=e^{-xu}x^{n+1}u^n/n!$.

Several authors [1–5] studied the convergence of some famous operators for functions of bounded variation. There are some results of convergence of G_n for functions of bounded variation on $[0,\infty)$ (cf. [6]).

In this paper, we consider the degree of simultaneous approximation by Gamma operator for functions in $B_r^{(\alpha)}$. We shall also prove that our estimates are essentially best possible. $B_r^{(\alpha)}=\{f\mid f^{(r-1)}\in C[0,\infty), f_\pm^{(r)}(x)$ exist everywhere and are bounded on every finite subinterval of $[0,\infty)$ and $f_\pm^{(r)}(x)=O(x^\alpha)(x\to\infty)$ for some $\alpha>0\}$, where $r=0,1,2,\dots$ and $f_\pm^{(0)}(x)$ means $f(x\pm)$.

1991 *Mathematics Subject Classification.* Primary 41A28; Secondary 41A36.

Key words and phrases. Operators, simultaneous approximation, derivatives, moduli of smoothness.

Project supported by Natural Science Foundation of Hebei Province.

2. Theorem

Let

$$h_r(t)=\begin{cases} f_+^{(r)}(t)-f_+^{(r)}(x), & x<t; \\ 0, & x=t; \\ f_-^{(r)}(t)-f_-^{(r)}(x), & 0\leqq t<x, \end{cases} \tag{2.1}$$

$$\omega_x(t)=\omega_x(h_r,t)=\sup\{|h_r(x+s)-h_r(x)|, |s|\leqq t\}.$$

It is clear that if $x\in[0,A](A>0)$, then

$$\omega_x(t)\leqq\omega_A(t), \tag{2.2}$$

where $\omega_A(t)=\sup\{\omega_x(t), x\in[0,A]\}$.

THEOREM. *If $f\in B_r^{(\alpha)}$, $(r\in\{0\}\cup N)$, $x>0$, then for n sufficiently large, we have*

$$\begin{aligned} &|G_n^{(r)}(f,x)-\frac{1}{2}(f_+^{(r)}(x)+f_-^{(r)}(x))| \\ &\leqq C_r\frac{33}{n}\sum_{k=1}^{n}\omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right)+|f_+^{(r)}(x)-f_-^{(r)}(x)|\,O\left(\frac{r}{\sqrt{n}}\right) \\ &+O\left(\frac{(2x)^p2^r}{\sqrt{n-r}}\left(\frac{\sqrt{e}}{2}\right)^n\right)+\frac{2^{r-1}r}{n}\left|f_+^{(r)}(x)+f_-^{(r)}(x)\right|, \end{aligned} \tag{2.3}$$

where $\omega_x(t)=\omega_x(h_r,t)$ is the pointwise modulus of continuity of h_r at x and h_r is defined by (2.1), $\Delta(x)=\max\{1,x^2\}$, $C_r=\dfrac{n^r}{n(n-1)\cdots(n+1-r)}$, p is an integer so that $\alpha\leqq p$.

3. Lemmas and preliminaries

In order to prove the above theorem we need the following lemmas.

LEMMA 1 ([7]). *If $f^{(r)}\in L_1$, then we have*

$$G_n^{(r)}(f,x)=C_r\int_0^\infty g_{n-r}(x,u)f^{(r)}\left(\frac{n}{u}\right)du, \tag{3.1}$$

where $C_r=\dfrac{n^r}{n(n-1)\cdots(n+1-r)}$.

LEMMA 2. *If p is a positive integer, then*

$$\int_{\frac{2x}{n}}^{\infty}(nt)^pg_{n-r}\left(x,\frac{1}{t}\right)\frac{dt}{t^2}=O(1)\frac{(2x)^p2^r}{\sqrt{n-r}}\left(\frac{\sqrt{e}}{2}\right)^n. \tag{3.2}$$

PROOF. Let

$$I = \frac{1}{(n-r)!} \int_{\frac{2x}{n}}^{\infty} (nt)^p \left(\frac{x}{t}\right)^{n-r+1} e^{-\frac{x}{t}} \frac{dt}{t}.$$

Then I is the left part of (3.2). By replacement of the variable in the integral, we have

$$\begin{aligned} I &= \frac{1}{(n-r)!} \int_{2x}^{\infty} u^p \left(\frac{xn}{u}\right)^{n-r+1} e^{-\frac{nx}{u}} \frac{du}{u} \\ &= \frac{1}{(n-r)!} \int_{2}^{\infty} (xv)^p \left(\frac{n}{v}\right)^{n-r+1} e^{-\frac{n}{v}} \frac{dv}{v} \qquad (3.3) \\ &= \frac{n^{n-r+1}}{(n-r)!} x^p \int_{2}^{\infty} e^{-\frac{n}{t}} t^{-(n-r+2-p)} dt. \end{aligned}$$

Using integration by part again and again for the term on the right-hand side of (3.3), we get

$$\begin{aligned} \int_{2}^{\infty} e^{-\frac{n}{t}} t^{-(n-r+2-p)} dt &= \int_{0}^{\frac{1}{2}} e^{-nt} t^{n-r-p} dt \\ &= \frac{e^{-\frac{n}{2}}}{(n-r-p+1)2^{n-r-p+1}} + \frac{n}{n-r-p+1} \int_{0}^{\frac{1}{2}} e^{-nt} t^{n-r-p+1} dt \qquad (3.4) \\ &= \sum_{k=0}^{2n} \frac{n^k}{(n-r-p+1)(n-r-p+2)\cdots(n-r-p+k+1)2^{n-r-p+k+1}} e^{-\frac{n}{2}} \\ &\quad + \frac{n^{2n+1}}{(n-r-p+1)\cdots(3n-r-p+2)} \int_{0}^{\frac{1}{2}} e^{-nt} t^{3n-r-p+2} dt. \end{aligned}$$

When n is sufficiently large, the right-hand side of (3.4)

$$\leqq 2e^{-\frac{n}{2}} \sum_{k=0}^{\infty} \frac{1}{n2^{n-r-p+k+1}} + 2\int_0^{\frac{1}{2}} t^{3n-r-p+2} dt$$
$$= 2e^{-\frac{n}{2}} \frac{1}{n2^{n-r-p}} + \frac{2}{3n-r-p+3} \left(\frac{1}{2}\right)^{3n-r-p+3}.$$

Since $\left(\frac{1}{2}\right)^{2n} = \left(\frac{1}{4}\right)^{n} < \left(\frac{1}{e}\right)^{\frac{n}{2}}$, hence for n sufficiently large, we have

$$\int_2^{\infty} e^{-\frac{n}{t}} t^{-(n-r+2-p)} dt \leqq 3\frac{e^{-\frac{n}{2}}}{n} \left(\frac{1}{2}\right)^{n-r} 2^p. \tag{3.5}$$

By (3.3), (3.5) and Stirling's formula

$$n! = \sqrt{2n\pi}\, n^n e^{-n} H_n,$$
$$H_n = e^{\frac{\theta_n}{12n}}, \qquad (0 < \theta_n < 1)$$

we get

$$I \leqq \frac{3e^{-\frac{n}{2}} n^{n-r} (2x)^p}{(n-r)!2^{n-r}}$$
$$= \frac{3e^{-\frac{n}{2}} n^{n-r} (2x)^p}{\sqrt{2\pi(n-r)}(n-r)^{n-r} e^{-(n-r)} 2^{n-r} H_{n-r}}$$
$$= \frac{O(1)(2x)^p 2^r}{\sqrt{n-r}} \left(\frac{\sqrt{e}}{2}\right)^n.$$

LEMMA 3. *If n is sufficiently large, then:*
(1) *For $0 \leqq y \leqq \frac{x}{n}$, $n \geqq 2r^2 + 2$, we have*

$$\int_0^y g_{n-r}\left(x, \frac{1}{u}\right) \frac{du}{u^2} \leqq \frac{8x^2}{n(x-ny)^2}. \tag{3.6}$$

(2) *For $\frac{x}{n} < z < \infty$, $n \geqq 2r^2 + 2$, we have*

$$\int_z^{\infty} g_{n-r}\left(x, \frac{1}{u}\right) \frac{du}{u^2} \leqq \frac{8x^2}{n(x-nz)^2}. \tag{3.7}$$

PROOF. By simple computation we have

$$\int_0^\infty g_{n-r}(x,u)du = 1,$$

$$\int_0^\infty g_{n-r}(x,u)\frac{n}{u}du = \frac{nx}{n-r},$$

$$\int_0^\infty g_{n-r}(x,u)\left(\frac{n}{u}\right)^2 du = \frac{n^2x^2}{(n-r)(n-r-1)}.$$

Hence for $n \geqq 2r^2+2$, we have

$$\begin{aligned} C_r \int_0^\infty g_{n-r}(x,u)\left(x-\frac{n}{u}\right)^2 du &= \frac{(n+r^2+r)n^{r-1}x^2}{(n-1)\cdots(n-r-1)} \\ \leqq \frac{2n^r x^2}{(n-1)\cdots(n-r-1)} &= \frac{2x^2}{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{r}{n}\right)(n-r-1)} \\ \leqq \frac{4x^2}{n(1-\frac{r}{n})^r} &\leqq \frac{8x^2}{n}. \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^y g_{n-r}(x,\frac{1}{u})\frac{du}{u^2} &\leqq \int_0^y \left(\frac{x-nu}{x-ny}\right)^2 g_{n-r}\left(x,\frac{1}{u}\right)\frac{du}{u^2} \\ &\leqq \frac{8x^2}{n(x-ny)^2}. \end{aligned}$$

So (3.6) is proved. The proof of (3.7) is similar.

LEMMA 4.

(3.8) $$\int_0^{\frac{n}{x}} g_{n-r}(x,u)du = \frac{1}{2} + O\left(\frac{r}{\sqrt{n}}\right),$$

(3.9) $$\int_{\frac{n}{x}}^\infty g_{n-r}(x,u)du = \frac{1}{2} + O\left(\frac{r}{\sqrt{n}}\right).$$

PROOF. Let

$$A_n(x) = \int_0^{\frac{n}{x}} g_{n-r}(x,u)du = \int_0^{\frac{n}{x}} \frac{x^{n-r+1}e^{-xu}u^{n-r}}{(n-r)!}du$$
$$= \int_0^n \frac{t^{n-r}e^{-t}}{(n-r)!}dt.$$

Using integration by part, we have

$$A_n(x) = \frac{1}{(n-r)!}\left[-e^{-t}(t^{n-r} + (n-r)t^{n-r-1} + \cdots + (n-r)!)|_0^n\right]$$
$$= 1 - e^{-n}\left(\frac{n^{n-r}}{(n-r)!} + \frac{n^{n-r-1}}{(n-r-1)!} + \cdots + n + 1\right).$$

From the relation ([6] Lemma 1), if n is a positive integer, then

$$1 + \frac{n}{1!} + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!}\theta(n) = \frac{e^n}{2},$$

where $\theta(n)$ lies between $\frac{1}{2}$ and $\frac{1}{3}$. We can obtain

$$A_n(x) = 1 - e^{-n}\left(\frac{1}{2}e^n - \frac{n^{n-r+1}}{(n-r+1)!} - \cdots - \frac{n^n}{n!}\theta(n)\right)$$
$$= \frac{1}{2} + e^{-n}\left(\frac{n^{n-r+1}}{(n-r+1)!} + \cdots + \frac{n^{n-1}}{(n-1)!} + \frac{n^n}{n!}\theta(n)\right)$$
$$(3.10) \qquad = \frac{1}{2} + e^{-n}\left(\frac{n^n}{n!}\left(\left(1-\frac{r-2}{n}\right)\left(1-\frac{r-3}{n}\right)\cdots\left(1-\frac{1}{n}\right)\right.\right.$$
$$\left.\left. + \cdots + \left(1-\frac{1}{n}\right)+1\right) + \frac{n^n}{n!}\theta(n)\right)$$
$$\leqq \frac{1}{2} + e^{-n}\left(\frac{n^n}{n!}(r+1) + \frac{n^n}{n!}\theta(n)\right).$$

By Stirling's formula, we get

$$e^{-n}\left(\frac{n^n}{n!}(r+1) + \frac{n^n}{n!}\theta(n)\right)$$
$$= \frac{(r+1)e^{-n}n^n}{\sqrt{2n\pi}n^ne^{-n}H_n} + \frac{e^{-n}n^n\theta(n)}{\sqrt{2n\pi}n^ne^{-n}H_n}$$

$$= O\left(\frac{r}{\sqrt{n}}\right).$$

Hence

$$A_n(x) = \frac{1}{2} + O\left(\frac{r}{\sqrt{n}}\right).$$

(3.9) can be obtained from $A_n(x) + B_n(x) = 1$, where $B_n(x)$ is the left side of (3.9).

4. Proof of the Theorem

Let $\delta = \sqrt{\frac{\Delta(x)}{n}}$, $\Delta(x) = \max\{1, x^2\}$. From

$$f^{(r)}(t) = \frac{1}{2}\left(f_+^{(r)}(x) + f_-^{(r)}(x)\right) + h_r(t) + \frac{1}{2}\left(f_+^{(r)}(x) - f_-^{(r)}(x)\right) \operatorname{sign}(t - x),$$

we have

$$\begin{aligned} &G_n^{(r)}(f, x) - \frac{1}{2}\left(f_+^{(r)}(x) + f_-^{(r)}(x)\right) \\ &= C_r \int_0^\infty h_r\left(\frac{n}{t}\right) g_{n-r}(x, t)dt + \frac{C_r(f_+^{(r)}(x) - f_-^{(r)}(x))}{2} \\ &\quad \times \left(\int_{\frac{n}{x}}^\infty - \int_0^{\frac{n}{x}}\right) g_{n-r}(x, t)dt + \frac{C_r - 1}{2}(f_+^{(r)}(x) + f_-^{(r)}(x)) \\ &:= \Sigma_1 + \Sigma_2 + \Sigma_3. \end{aligned} \tag{4.1}$$

Obviously, for $n > 2r$, we have $C_r - 1 \leqq 2^{r-1}r/n$. So

$$|\Sigma_3| \leqq \frac{2^{r-1}r}{n}\left|f_+^{(r)}(x) + f_-^{(r)}(x)\right|. \tag{4.2}$$

From (3.8) and (3.9), we have

$$|\Sigma_2| = |f_+^{(r)}(x) - f_-^{(r)}(x)| O\left(\frac{r}{\sqrt{n}}\right). \tag{4.3}$$

Next, we evaluate Σ_1, replacing the variable $\frac{1}{t}$ by t, we obtain

$$\Sigma_1 = C_r \int_0^\infty h_r(nt) g_{n-r}\left(x, \frac{1}{t}\right) \frac{dt}{t^2}$$

$$(4.4) \qquad = C_r \left(\int_0^{\frac{x-\delta}{n}} + \int_{\frac{x-\delta}{n}}^{\frac{x+\delta}{n}} + \int_{\frac{x+\delta}{n}}^{\frac{2x}{n}} + \int_{\frac{2x}{n}}^{\infty} \right) h_r(nt) \frac{g_{n-r}\left(x, \frac{1}{t}\right)}{t^2} dt$$

$$= C_r(I_1 + I_2 + I_3 + I_4).$$

Let

$$K_n(x,t) = \int_0^t \frac{g_{n-r}\left(x, \frac{1}{u}\right)}{u^2} du = \int_{1/t}^\infty g_{n-r}(x,u)du.$$

Obviously, if $[a,b] \subset [0,\infty)$, we have

$$\int_a^b dK_n(x,t) \leqq \int_0^\infty dK_n(x,t) = 1.$$

$$I_1 = \int_0^{\frac{x-\delta}{n}} [h_r(nt) - h_r(x)] g_{n-r}\left(x, \frac{1}{t}\right) \frac{dt}{t^2},$$

$$|I_1| \leqq \int_0^{\frac{x-\delta}{n}} \omega_x(x-nt) g_{n-r}\left(x, \frac{1}{t}\right) \frac{dt}{t^2}$$

$$= \int_0^{\frac{x-\delta}{n}} \omega_x(x-nt) d_t K_n(x,t).$$

Using integration by parts, from (3.6) (3.7), we have

$$|I_1| \leqq \omega_x(x-nt) K_n(x,t)\Big|_0^{\frac{x-\delta}{n}} - \int_0^{\frac{x-\delta}{n}} K_n(x,t) d_t \omega_x(x-nt)$$

$$\leqq \frac{8x^2}{n\delta^2}\omega_x(\delta) + \frac{8x^2}{n} \int_0^{\frac{x-\delta}{n}} \frac{d_t(-\omega_x(x-nt))}{(x-nt)^2},$$

$$\int_0^{\frac{x-\delta}{n}} \frac{d_t(-\omega_x(x-nt))}{(x-nt)^2} = -\frac{\omega_x(\delta)}{\delta^2} + \frac{\omega_x(x)}{x^2} + 2n \int_0^{\frac{x-\delta}{n}} \frac{\omega_x(x-nt)}{(x-nt)^3} dt.$$

Hence

$$\begin{aligned} |I_1| &\leqq \frac{8}{n}\omega_x(x) + 16x^2 \int_0^{\frac{x-\delta}{n}} \frac{\omega_x(x-nt)}{(x-nt)^3} dt \\ &= \frac{8}{n}\omega_x(x) + \frac{16x^2}{n} \int_0^{x-\delta} \frac{\omega_x(x-u)}{(x-u)^3} du. \end{aligned}$$

Replacing the variable u in the last integral by $x - \sqrt{\frac{\Delta(x)}{v}}$, we find that

$$\begin{aligned} \int_0^{x-\delta} \frac{\omega_x(x-u)}{(x-u)^3} du &= \frac{1}{2\Delta(x)} \int_{\frac{\Delta(x)}{x^2}}^{n} \omega_x(\sqrt{\frac{\Delta(x)}{v}}) dv \\ &\leqq \frac{1}{2\Delta(x)} \sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right). \end{aligned}$$

Hence

$$|I_1| \leqq \frac{8}{n}\omega_x(x) + \frac{8}{n}\sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right) \leqq \frac{16}{n}\sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right). \tag{4.5}$$

Now let us turn to the estimation of I_2:

$$\begin{aligned} |I_2| &\leqq \int_{\frac{x-\delta}{n}}^{\frac{x+\delta}{n}} |h_r(nt) - h_r(x)| d_t K_n(x,t) \leqq \omega_x(\delta) \int_{\frac{x-\delta}{n}}^{\frac{x+\delta}{n}} d_t K_n(x,t) \\ &\leqq \frac{1}{n}\sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right). \end{aligned} \tag{4.6}$$

Similarly, using integration by parts and (3.7) (3.8), replacing the variable u in the estimation by $x + \sqrt{\frac{\Delta(x)}{t}}$, we have

$$|I_3| \leqq \int_{\frac{x+\delta}{n}}^{\frac{2x}{n}} \omega_x(nt - x) d_t(-\overline{K}_n(x,t)),$$

where $\overline{K}_n(x,z) = \int_z^{\frac{2x}{n}} g_{n-r}\left(x, \frac{1}{t}\right) \frac{dt}{t^2}$.

$$|I_3| \leqq \frac{8x^2}{n\delta^2}\omega_x(\delta) + \frac{8x^2}{n} \int_{\frac{x+\delta}{n}}^{\frac{2x}{n}} \frac{d_t(\omega_x(nt-x))}{(nt-x)^2}$$

$$\leqq \frac{8\omega_x(x)}{n} + 16x^2 \int_{\frac{x+\delta}{n}}^{\frac{2x}{n}} \frac{\omega_x(nt-x)}{(nt-x)^3} dt,$$

$$\int_{\frac{x+\delta}{n}}^{\frac{2x}{n}} \frac{\omega_x(nt-x)}{(nt-x)^3} dt = \frac{1}{n} \int_{x+\delta}^{2x} \frac{\omega_x(u-x)}{(u-x)^3} du$$

$$= \frac{1}{2n\Delta(x)} \int_{\frac{\Delta(x)}{x^2}}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{v}}\right) dv$$

$$\leqq \frac{1}{2n\Delta(x)} \int_1^n \omega_x\left(\sqrt{\frac{\Delta(x)}{v}}\right) dv$$

$$\leqq \frac{1}{2n\Delta(x)} \sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right).$$

Hence

$$|I_3| \leqq \frac{16}{n} \sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right). \tag{4.7}$$

Applying Lemma 2, we have

$$\begin{aligned} |I_4| &\leqq \int_{\frac{2x}{n}}^{\infty} |h_r(nt)| g_{n-r}\left(x, \frac{1}{t}\right) \frac{dt}{t^2} \\ &= O(1) \frac{(2x)^p 2^r}{\sqrt{n-r}} \left(\frac{\sqrt{e}}{2}\right)^n. \end{aligned} \tag{4.8}$$

Collecting (4.1)-(4.8), we can obtain (2.3).

5. Remarks

(1) If $f^{(r)}$ is not constant in any neighbourhood of x, since $\frac{\sqrt{e}}{2} < 1$, for n sufficiently large, we have

$$O(1)\frac{(2x)^p 2^r}{\sqrt{n-r}}\left(\frac{\sqrt{e}}{2}\right)^n n \leqq \sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right),$$

hence

$$|I_4| \leqq \frac{1}{n}\sum_{k=1}^{n} \omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right).$$

So in that case (2.3) becomes

$$\begin{aligned} &|\,G_n^{(r)}(f,x) - \frac{1}{2}(f_+^{(r)}(x) + f_-^{(r)}(x))\,| \\ &\leqq C_r \frac{34}{n}\sum_{k=1}^{n}\omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right) + |\,f_+^{(r)}(x) - f_-^{(r)}(x)\,|\,O\left(\frac{r}{\sqrt{n}}\right) \qquad (5.1)\\ &+ \frac{2^{r-1}r}{n}\left|f_+^{(r)}(x) + f_-^{(r)}(x)\right|. \end{aligned}$$

(2) If $f^{(r)}$ is continuous at x, then (5.1) can be further simplified to

$$|G_n^{(r)}(f,x) - f^{(r)}(x)| \leqq C_r\frac{34}{n}\sum_{k=1}^{n}\omega_x\left(\sqrt{\frac{\Delta(x)}{k}}\right) + \frac{2^r r}{n}|f^{(r)}(x)|. \qquad (5.2)$$

(3) As far as the precision of the above estimates is concerned, we can prove that (5.2) cannot be asymptotically improved. Consider the function

$$f(t) = \underbrace{\int_0^t \int_0^{t_{r-1}} \cdots \int_0^{t_1}}_{r\ factors} |u-x|\,du\,dt_1 \cdots dt_{r-1} \qquad (x>0)$$

at $t=x$. From (5.2) we have

$$|G_n^{(r)}(f,x) - f^{(r)}(x)| \leqq C_r\frac{34x}{\sqrt{n}}.$$

On the other hand, we can prove (similarly to [6], here we omit the details)

$$|G_n^{(r)}(f,x) - f^{(r)}(x)| \geqq \frac{x}{\sqrt{2\pi n}}. \qquad (5.3)$$

Hence

$$\frac{x}{\sqrt{2\pi n}} \leqq |G_n^{(r)}(f,x) - f^{(r)}(x)| \leqq C_r \frac{34x}{\sqrt{n}}.$$

Therefore (5.2) cannot be asymptotically improved.

REFERENCES

[1] Bojanic, R. and Vuilleumier, M., On the rate of convergence of Fourier–Legendre series of functions of bounded variation, *J. Approx. Theory* **31** (1981), 67–79. *MR* **82j**:42018

[2] Cheng, F. H., On the rate of convergence of Bernstein polynomials of functions of bounded variation, *J. Approx. Theory* **39** (1983), 259–274. *MR* **85g**:41014

[3] Cheng, F., On the rate of convergence of the Szász–Mirakyan operator for functions of bounded variation, *J. Approx. Theory* **40** (1984), 226–241. *MR* **85i**:41014

[4] Guo, S. S., On the rate of convergence of the Durrmeyer operator for functions of bounded variation, *J. Approx. Theory* **51** (1987), 183–192. *MR* **89h**:41041

[5] Guo, S. S., Approximation by Feller operators for functions having discontinuity points of the first kind, *Acta Math. Appl. Sinica* **14** (1991), no. 1, 57–65 (in Chinese). *MR* **96b**:41028

[6] Cheng, W. Z. and Guo, S. S., On the rate of convergence of the Gamma operator for functions of bounded variation, *Approx. Theory Appl.* **1** (1985), no. 5, 85–96. *MR* **87m**:41023

[7] Hermann, T., On the derivatives of positive linear operators, *Constructive theory of functions* (Varna, 1987), Bulgar. Acad. Sci., Sofia, 1988, 214–221. *MR* **90h**:41023

[8] Totik, V., Uniform approximation by positive operators on infinite intervals, *Anal. Math.* **10** (1984), 163–182. *MR* **86e**:41043

(Received January 17, 1996)

DEPARTMENT OF MATHEMATICS
HEBEI NORMAL UNIVERSITY
SHIJIAZHUANG 050016
PEOPLE'S REPUBLIC OF CHINA
lhzhong@public.sj.hj.cn

Studia Scientiarum Mathematicarum Hungarica

Volume 34

1828

Akadémiai Kiadó, Budapest

1998

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 34

CONTENTS

Typeset by TypoTEX Ltd., Budapest
PRINTED IN HUNGARY
Akadémiai Nyomda, Martonvásár

RECENTLY ACCEPTED PAPERS

VOLKMER, H., Eigenvalue problems for Bessel's equation and zero-pairs of Bessel functions

TÖLKE, J., Ein konkretes Beispiel zu den symmetrischen Möbius-Zwangläufen

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