Acta Math. Hung.
42 (3—4) (1983), 319—330.

ON THE STRONG LAW OF LARGE NUMBERS FOR PAIRWISE INDEPENDENT RANDOM VARIABLES

S. CSÖRGŐ, K. TANDORI, member of the Academy, and V. TOTIK (Szeged)

Results and discussion

The basic idea of Etemadi's clever "elementary proof of the strong law of large numbers" for identically distributed random variables in [2] consists in the observation that it is enough to establish the law for the positive and negative parts of the underlying variables separately. Then the monotonicity of the corresponding partial sums enables one to avoid the use of the Kolmogorov inequality, and the only property required from these positive or negative part variables is that the variance of a sum from some truncated versions of them be equal to the sum of the termwise variances. The simplest way to ensure the latter requirement of lack of pairwise correlation is to assume that the original sequence consists of pairwise independent variables. Thus Etemadi was able to show that the averages of pairwise independent identically distributed random variables converge almost surely to the common mean of the variables assumed to be finite, thereby relaxing the assumption of total independence in Kolmogorov's classic law.

The aim of the present note is to point out that Etemadi's idea works for nonidentically distributed random variables as well. We derive analogues of Kolmogorov's other classical strong law for nonidentically distributed variables assumed to be pairwise independent only. Our sufficient conditions raise a few problems and some of these are also solved and discussed below.

Setting $D^2(X)=E(X-EX)^2$ for the variance of the random variable X, the principal result is

THEOREM 1. *If the pairwise independent random variables $X_1, X_2, \ldots$ satisfy the conditions*

$$\sum_{m=1}^{\infty} \frac{D^2(X_m)}{m^2} < \infty \tag{1}$$

and

$$\frac{1}{n} \sum_{m=1}^{n} E|X_m - EX_m| = O(1), \tag{2}$$

then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} (X_m - EX_m) = 0 \tag{3}$$

almost surely.

For a sequence $a=\{a_n\}_1^\infty$ of positive numbers set

$$Y_n(a) = \begin{cases} Y_n, & \text{if } |Y_n| \leqq a_n \\ 0, & \text{otherwise,} \end{cases}$$

where $Y_n = X_n - EX_n$, $E|X_n| < \infty$, $n = 1, 2, \dots$. The "truncated" version of Theorem 1 is

THEOREM 2. *If the random variables* $X_1, X_2, \dots$ *are pairwise independent such that condition* (2) *is satisfied and the centered and truncated sequence satisfies*

$$\sum_{m=1}^{\infty} \frac{D^2(Y_m(a))}{m^2} < \infty \tag{4}$$

and

$$\sum_{m=1}^{\infty} P\{Y_m \neq Y_m(a)\} < \infty, \tag{5}$$

then for any sequence $\{b_n\}_1^\infty$ *of constants satisfying*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m-1}^{n} (b_m - EY_m(a)) = 0$$

we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} (Y_m - b_m) = 0 \tag{6}$$

almost surely.

In general, the choice $b_m = 0$, $m = 1, 2, \dots$ cannot be ensured under the given conditions. However, if, in addition, the variables $X_1, X_2, \dots$ are identically distributed in Theorem 2, then (4) and (5) are known to hold by the classical proof of Kolmogorov for the truncating sequence $a_n = n$, $n = 1, 2, \dots$ (pairwise independence suffices here!), and it is easy to see that $EY_n(a) - EX_n = EY_n(a) - EX_1 \to 0$, as $n \to \infty$, with this truncation. Therefore Etemadi's result follows from Theorem 1 in exactly the same classic way as Kolmogorov's theorem for identically distributed variables followed from his other theorem for non-identically distributed variables.

The latter Kolmogorov theorem tells us that condition (1) alone is sufficient for (3) to hold for totally independent random variables. It is well known that, in this case, this condition is best possible in the sense that if $\{\sigma_1^2\}_n^\infty$ is a sequence of positive numbers such that

$$\sum_{m=1}^{\infty} \frac{\sigma_m^2}{m^2} = \infty,$$

then there exists a sequence $X_1, X_2, \dots$ of totally independent random variables not obeying (3) but for which $D^2(X_m) = \sigma_m^2$, $m = 1, 2, \dots$ (cf. Révész [6]). On the other hand, according to the equally classical result of Menšov and Rademacher (cf. again [6]), the stronger condition

$$\sum_{m=2}^{\infty} \frac{D^2(X_m)(\log m)^2}{m^2} < \infty \tag{7}$$

already implies (3) for any sequence $X_1, X_2, \dots$ of pairwise uncorrelated (orthogonal) random variables. The latter condition is also best possible for uncorrelated variables in the above sense (Tandori [8]), provided that the corresponding σ_m^2/m^2 sequence is nonincreasing. We show that condition (1) alone is not enough to imply (3) when the variables are only pairwise independent.

THEOREM 3. *For every ε, $0<\varepsilon<1$, there exists a sequence $X_1, X_2, \dots$ of pairwise independent random variables such that*

$$\sum_{m=2}^{\infty} \frac{D^2(X_m)(\log\log m)^{1-\varepsilon}}{m^2} < \infty \tag{8}$$

and

$$P\left\{\overline{\lim_{n\to\infty}} \frac{1}{n}\left|\sum_{m=1}^{\infty}(X_m - EX_m)\right| = \infty\right\} > 0. \tag{9}$$

Therefore some auxiliary conditions are necessary to assume beside (1) to have (3). Theorem 1 provides an example for such a condition. The role of this side-condition (2) is not entirely clear in general because, as simpe counterexamples show, (2) is not necessary for (3) even if we assume

$$\sum_{m=2}^{\infty} \frac{D^2(X_m)(\log m)^2}{m^2} = \infty$$

beside (1).

The construction in Theorem 3 relies on a construction of divergent Walsh series by Tandori [7]. His construction was later refined by Bočkarev [1] and Nakata [5]. Using the results of these authors, one can in fact strengthen Theorem 3 to obtain the following version of it.

THEOREM 3*. *For every $\varepsilon, 0<\varepsilon<1$, there exists a sequence $X_1, X_2, \dots$ of pairwise independent random variables such that*

$$\sum_{m=2}^{\infty} \frac{D^2(X_m)(\log m)^{1-\varepsilon}}{m^2} < \infty \tag{10}$$

and (9) *still holds true.*

The proof of this result is not given in this note because it would require very large space as compared to that of Theorem 3. Condition (10) is now quite close to condition (7). In this respect we mention the following conjecture of ours, an affirmative proof of which would mean that pairwise independent and pairwise uncorrelated random variables do not really differ from each other from the point of view of the strong law of large numbers if one is looking at the growth of the variances only.

CONJECTURE. *For pairwise independent random variables the condition*

$$\sum_{m=2}^{\infty} \frac{D^2(X_m)(\log m)^{2-\varepsilon}}{m^2} < \infty, \quad 0 < \varepsilon < 1,$$

is not enough in general to ensure (3) *to hold almost surely.*

On the other hand, pairwise independent and pairwise uncorrelated random variables do differ from each other from the point of view of the strong law of large numbers below the domain of (10), or that of the Conjecture, i.e., when we have only (1) and, necessarily, an extra assumption such as (2). Indeed, pairwise independence cannot be relaxed to pairwise orthogonality in Theorem 1. This is the content of our last result in this paper.

THEOREM 4. *There exists a sequence* $X_1, X_2, \ldots$ *of pairwise uncorrelated random variables such that conditions* (1) *and* (2) *are satisfied, but*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \left| \sum_{m=1}^{n} (X_m - EX_m) \right| = \infty$$

almost surely.

In fact the random variables constructed in the proof of Theorem 4 satisfy the condition

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} E|X_m - EX_m| = 0.$$

Proofs

PROOF OF THEOREM 1. We may and do assume $EX_n = 0$, $n = 1, 2, \ldots$. Introduce the partial sums of the positive and negative parts of our variables $X_n = X_n^+ - X_n^-$:

$$S_n^+ = X_1^+ + \ldots + X_n^+, \quad S_n^- = X_1^- + \ldots + X_n^-, \quad n = 1, 2, \ldots .$$

By assumption (2) there is a constant A such that the inequality

$$0 \leqq \frac{1}{n} ES_n^+ \leqq A$$

is satisfied for each n. Let $\alpha > 1$, $\varepsilon > 0$ and $L = [A/\varepsilon]$, the integer part of A/ε. For each pair of integers m and s, $m \geqq 0$, $s = 0, \ldots, L$, put

$$k_s^-(m) = \inf \{k\colon \alpha^m \leqq k < \alpha^{m+1}, \frac{1}{k} ES_k^+ \in [s\varepsilon, (s+1)\varepsilon\},$$

$$k_s^+(m) = \sup \{k\colon \alpha^m \leqq k < \alpha^{m+1}, \frac{1}{k} ES_k^+ \in [s\varepsilon, (s+1)\varepsilon)\}$$

if the set on the right is not empty, and let $k_s^+(m) = k_s^-(m) = [\alpha^m]$ otherwise.

Since, obviously, $D^2(X_n^+) + D^2(X_n^-) \leqq D^2(X_n)$ for each n, we obtain, by the pairwise independence of $X_1^+, X_2^+, \ldots$, the estimate

$$\sum_{m=0}^{\infty} \frac{1}{(k_s^{\pm}(m))^2} E(S^+_{k_s^{\pm}(m)} - ES^+_{k_s^{\pm}(m)})^2 = \sum_{m=0}^{\infty} \frac{1}{(k_s^{\pm}(m))^2} \sum_{j=1}^{k_s^{\pm}(m)} D^2(X_j^+) \leqq \tag{11}$$

$$\leqq \sum_{m=0}^{\infty} \frac{1}{(k_s^{\pm}(m))^2} \sum_{j=1}^{k_s^{\pm}(m)} D^2(X_j) = \sum_{j=1}^{\infty} D^2(X_j) \sum_{\{m\colon k_s^{\pm}(m) \geqq j\}} \frac{1}{(k_s^{\pm}(m))^2} \leqq$$

$$\leqq \sum_{j=1}^{\infty} D^2(X_j) \sum_{\{m\colon \alpha^m \geqq j\}} \frac{1}{(\alpha^m)^2} \leqq \frac{\alpha^2}{\alpha^2 - 1} \sum_{j=1}^{\infty} \frac{D^2(X_j)}{j^2} < \infty$$

for any $s=0, \dots, L,$ by condition (1). Hence the monotone convergence theorem implies that

$$\lim_{m\to\infty} \frac{1}{k_s^{\pm}(m)} (S^+_{k_s^{\pm}} - ES^+_{k_s^{\pm}(m)}) = 0 \quad \text{a.s.} \tag{12}$$

for each $s = 0, \dots, L.$

Now for any natural number n there exists an $m=m(n)$ and an $s=s(n)$, $\lim\limits_{n\to\infty} m(n)=\infty$, $0 \leqq s(n) \leqq L$, such that $\alpha^m \leqq n < \alpha^{m+1}$ and $ES_n^+/n \in [s\varepsilon, (s+1)\varepsilon)$. By the definitions of $k_s^{\pm}(m)$ we have

$$k_s^-(m) \leqq n \leqq k_s^+(m), \qquad \left| \frac{1}{k_s^{\pm}(m)} ES^+_{k_s^{\pm}} - \frac{1}{n} ES_n^+ \right| \leqq \varepsilon,$$

and so

$$-\varepsilon - \left(1-\frac{1}{\alpha}\right) A + \frac{1}{\alpha} \frac{1}{k_s^-(m)} (S^+_{k_s^-(m)} - ES^+_{k_s^-(m)}) \leqq$$

$$\leqq -\varepsilon - \left(1-\frac{1}{\alpha}\right) \frac{1}{k_s^-(m)} ES^+_{k_s^-(m)} + \frac{1}{\alpha} \frac{1}{k_s^-(m)} (S^+_{k_s^-(m)} - ES^+_{k_s^-(m)}) \leqq$$

$$\leqq \frac{1}{n} S^+_{k_s^-(m)} - \frac{1}{n} ES_n^+ \leqq \frac{1}{n} (S_n^+ - ES_n^+) \leqq \frac{1}{n} S^+_{k_s^+(m)} - \frac{1}{k_s^+(m)} ES^+_{k_s^+(m)} + \varepsilon \leqq$$

$$\leqq \frac{\alpha}{k_s^+(m)} (S^+_{k_s^+(m)} - ES^+_{k_s^+(m)}) + (\alpha-1)A + \varepsilon.$$

By (12) these inequalities yield

$$-\varepsilon - \left(1-\frac{1}{\alpha}\right) A \leqq \varliminf_{n\to\infty} \frac{1}{n} (S_n^+ - ES_n^+) \leqq \varlimsup_{n\to\infty} \frac{1}{n} (S_n^+ - ES_n^+) \leqq (\alpha-1)A + \varepsilon$$

almost surely, and since this is true for any $\alpha > 1$ and $\varepsilon > 0$, we obtain that

$$\lim_{n\to\infty} \frac{1}{n} (S_n^+ - ES_n^+) = 0 \quad \text{a.s.}$$

It can be proved in the same way that

$$\lim_{n\to\infty} \frac{1}{n} (S_n^- - ES_n^-) = 0 \quad \text{a.s.},$$

and the theorem follows.

PROOF OF THEOREM 2. Clearly,

$$\frac{1}{n} \sum_{m=1}^{n} E|Y_m(a) - EY_m(a)| \leqq \frac{2}{n} \sum_{m=1}^{n} E|Y_m(a)| \leqq \frac{2}{n} \sum_{m=1}^{n} E|X_m - EX_m| = O(1),$$

and hence an application of Theorem 1 to $Y_1(a), Y_2(a), \dots$ yields

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} (Y_m(a) - EY_m(a)) = 0 \quad \text{a.s.}$$

Condition (5) and the Borel—Cantelly lemma then give (6) with $b_m=Y_m(a)$, $m=1, 2, \ldots$.

Proof of Theorem 3. Consider the probability space $(\Omega, \mathcal{A}, P)$, where $\Omega=(0, 1)$, $\mathcal{A}$ is the σ-algebra of the Lebesgue-measurable subsets of Ω and P is the Lebesgue measure on $\mathcal{A}$. Let $r_n(\omega)=\operatorname{sign}\sin(2^n\pi\omega)$, $n=1, 2, \ldots$; $\omega\in(0, 1)$, be the n-th Rademacher function. Consider the sequence $\{w_n(\omega)\}_0^\infty$ of Walsh functions, i.e., $w_0(\omega)\equiv 1$, and if n is a positive integer with the diadic representation $n=2^{\nu_1}+\ldots+2^{\nu_p}$, $\nu_1<\ldots<\nu_p$, then $w_n(\omega)=r_{\nu_1+1}(\omega)\ldots r_{\nu_p+1}(\omega)$, $\omega\in(0, 1)$. For any $n=1, 2, \ldots$ we have $Ew_n=0$ and $D(w_n)=1$. Set

$$\varphi(m)=2^{2^m} \quad\text{and}\quad \psi(m)=2^{2^{2^m}}, \quad m=1, 2, \ldots.$$

Tandori [7] proved that if $m>1$, then there is a rearrangement

$$w_{i_1(m)}(\omega), \ldots, w_{i_{\varphi(m)}(m)}(\omega)$$

of the functions $w_1(\omega), \ldots, w_{\varphi(m)}(\omega)$ and there can be given real numbers $a_1(m), \ldots$ $\ldots, a_{\varphi(m)}(m)$ such that

$$a_1^2(m)+\ldots+a_{\varphi(m)}^2(m)\leqq 5m$$

and if $\omega\in(0, 1/4)$ and is not diadic, then

$$\max_{1\leqq k\leqq\varphi(m)}\sum_{j=1}^{k}a_j(m)w_{i_j(m)}(\omega)\geqq\frac{m}{2}.$$

For any $m=1, 2, \ldots$, define the numbers

$$D^2_{\varphi(m)+n}=\frac{(\psi(m))^2a_n^2(2^m)}{2^{m(2-\varepsilon/2)}}, \quad n=1, \ldots, \psi(m),$$

where $0<\varepsilon<1$, and introduce the random variables

$$X_{\psi(m)+n}(\omega)=D_{\psi(m)+n}r_{\varphi(m)+1}(\omega)w_{i_n(2^m)}(\omega), \quad n=1, \ldots, \psi(m).$$

For any other index k, let $X_k(\omega)\equiv 0$, $\omega\in(0, 1)$. By the definition of the Walsh functions it is obvious that $X_1, X_2, \ldots$ are pairwise independent, $EX_n=0$, $n=1, 2, \ldots$, and the definition also ensures that condition (8) is satisfied (implying (1)). Exactly as we saw it in (11), condition (1) alone implies that

$$\lim_{m\to\infty}\frac{1}{\psi(m)}\sum_{j=1}^{\psi(m)}X_j=0$$

almost everywhere in $(0, 1)$. Nevertheless we show that

$$\varlimsup_{n\to\infty}\frac{1}{n}\left|\sum_{j=1}^{n}X_j(\omega)\right|=\infty \tag{13}$$

almost everywhere in $(0, 1/4)$.

For any integer $m\geqq 1$ let $\psi(m)<n\leqq 2\psi(m)$. Then

$$\frac{1}{n}\sum_{j=1}^{n}X_j(\omega)=\frac{1}{n}\sum_{j=\psi(m)+1}^{n}X_j(\omega)+\frac{\psi(m)}{n}\frac{1}{\psi(m)}\sum_{j=1}^{\psi(m)}X_j(\omega)=V_n(\omega)+Z_n(\omega),$$

say. We know that $\lim_{n\to\infty} Z_n = 0$ almost everywhere in $(0, 1)$. But

$$V_n(\omega) = \frac{\psi(m)}{n 2^{m(1-\varepsilon/4)}} r_{\varphi(m)+1}(\omega) \sum_{j=1}^{n-\psi(m)} a_j(2^m) w_{i_j(2^m)}(\omega),$$

and hence

$$\max_{\psi(m)<n\leqq 2\psi(m)} |V_n(\omega)| \geqq \frac{1}{2} \frac{1}{2^{m(1-\varepsilon/4)}} \max_{1\leqq k\leqq \psi(m)} \left| \sum_{j=1}^{k} a_j(2^m) w_{i_j(2^m)}(\omega) \right| \geqq \frac{1}{4} 2^{m(\varepsilon/4)}$$

almost everywhere in $(0, 1/4)$. It now follows that we have (13) indeed, almost everywhere in $(0, 1/4)$.

PROOF OF THEOREM 4. Our point of departure is a function system of Kaczmarcz [3] which, in essence, is a simplified version of a system considered originally by Menšov [4]. Let $p \geqq 2$ be an integer and for $l = 1, \ldots, 2p$, consider the functions

$$u_l(x) = \frac{1}{k-p-l-\dfrac{1}{2}}, \quad x \in \left[\frac{k-1}{p}, \frac{k}{p}\right), \quad k = 1, \ldots, 4p.$$

We have

$$\int_0^4 u_l^2(x)\,dx = \frac{1}{p} \sum_{k=1}^{4p} \frac{1}{\left(k-p-l-\dfrac{1}{2}\right)^2},$$

whence

$$\frac{C_1}{p} \leqq \int_0^4 u_l^2(x)\,dx \leqq \frac{C_2}{p}, \quad l = 1, \ldots, 2p, \tag{14}$$

where $C_1, C_2, \ldots$ will denote positive absolute constants. If $i > j$, then

$$\alpha_{ij} = \int_0^4 u_i(x) u_j(x)\,dx = \frac{1}{p} \sum_{k=1}^{4p} \frac{1}{\left(k-p-i-\dfrac{1}{2}\right)\left(k-p-j-\dfrac{1}{2}\right)} =$$

$$= \frac{1}{p} \frac{1}{i-j} \sum_{k=1}^{4p} \left\{ \frac{1}{k-p-i-\dfrac{1}{2}} - \frac{1}{k-p-j-\dfrac{1}{2}} \right\} =$$

$$= \frac{1}{p} \frac{1}{i-j} \left\{ \sum_{k=1-p-i}^{3p-i} \frac{1}{k-\dfrac{1}{2}} - \sum_{k=1-p-j}^{3p-j} \frac{1}{k-\dfrac{1}{2}} \right\} =$$

$$= \frac{1}{p} \frac{1}{i-j} \left\{ \sum_{k=1-p-i}^{-p-j} \frac{1}{k-\dfrac{1}{2}} - \sum_{k=3p-i+1}^{3p-j} \frac{1}{k-\dfrac{1}{2}} \right\},$$

and therefore

$$|\alpha_{ij}| \leqq \frac{1}{p}\frac{1}{i-j}\left\{\frac{i-j}{p+j+\frac{1}{2}} - \frac{i-j}{3p-i-\frac{1}{2}}\right\} \leqq \frac{C_3}{p^2}. \tag{15}$$

We extend the definition of these functions $u_l(x)$ to the interval $(4, 5]$ in such a way that they be orthogonal on the whole interval $[0, 5]$. Let us divide up the interval $(4, 5]$ into $N=2p(2p-1)$ pairwise disjoint subintervals I_{ij}, $i, j=1, \ldots, 2p$, $i\neq j$, of equal length and set

$$u_l(x) = \begin{cases} \sqrt{\frac{1}{2}N|\alpha_{lj}|}, & x\in I_{lj},\ j=1, \ldots, 2p,\ j\neq l, \\ -\sqrt{\frac{1}{2}N|\alpha_{lj}|}\operatorname{sign}\alpha_{lj}, & x\in I_{jl},\ j=1, \ldots, 2p,\ j\neq l, \\ 0, & \text{otherwise,} \end{cases}$$

for each $l=1, \ldots, 2p$. The functions $u_l(x)$ thus obtained are obviously simple (step) functions on $[0, 5]$ and they constitute an orthogonal system there. Furthermore

$$\int_0^5 u_l^2(x)\,dx = \int_0^4 u_l^2(x)\,dx + \frac{1}{2}\sum_{j=1}^{l-1}|\alpha_{lj}| + \frac{1}{2}\sum_{j=l+1}^{2p}|\alpha_{lj}|$$

and hence by (14) and (15),

$$\frac{C_4}{p} \leqq \int_0^5 u_l^2(x)\,dx \leqq \frac{C_5}{p}, \quad l=1, \ldots, 2p. \tag{16}$$

If $x\in[2, 3)$, then there is a non-negative integer $m(x)$ depending on x, $m(x)<p$, such that

$$x\in\left[\frac{2p+m(x)}{p}, \frac{2p+m(x)+1}{p}\right).$$

Then by definition, $u_1(x)\geqq 0, \ldots, u_{p+m(x)}(x)\geqq 0$, and

$$\sum_{l=1}^{p+m(x)} u_l(x) = \sum_{l=1}^{p+m(x)} \frac{1}{2p+m(x)+1+p-l-\frac{1}{2}} = \sum_{l=1}^{p+m(x)} \frac{1}{l-\frac{1}{2}},$$

and this implies that

$$\max_{1\leqq m\leqq 2p}\left|\sum_{l=1}^{m} u_l(x)\right| \geqq C_6\log p, \quad x\in[2, 3). \tag{17}$$

Moreover we have

$$\int_0^5 u_l^+(x)\,dx = \frac{1}{p}\sum_{k=p+l+1}^{4p}\frac{1}{k-p-l-\frac{1}{2}} + \frac{1}{\sqrt{2p(2p-1)}}\sum_{\substack{j=1\\ j\neq l}}^{2p}\sqrt{|\alpha_{lj}|},$$

$$\int_0^5 u_l^-(x)\,dx = \frac{1}{p}\sum_{k=1}^{p+l}\frac{1}{k-p-l-\frac{1}{2}} + \frac{1}{\sqrt{2p(2p-1)}}\sum_{\substack{j=1\\ j\neq l}}^{2p}\sqrt{|\alpha_{lj}|},$$

whence

$$\int_0^5 u_l^+(x)\,dx \leqq C_7 \frac{\log p}{p}, \quad \int_0^5 u_l^-(x)\,dx \leqq C_7 \frac{\log p}{p}. \tag{18}$$

Put

$$v_l(x) = u_l(x) \Big/ \sqrt{\int_0^5 u_l^2(x)\,dx}, \quad l = 1, \dots, 2p.$$

These are again simple functions on $[0, 5]$ and constitute an orthonormal system there, and by (16), (17) and (18) we have

$$\max_{1 \leqq m \leqq 2p} \left| \sum_{l=1}^m v_l(x) \right| \geqq C_8 \sqrt{p} \log p, \quad x \in [2, 3), \tag{19}$$

$$\int_0^5 v_l^+(x)\,dx \leqq C_9 \frac{\log p}{\sqrt{p}}, \quad \int_0^5 v_l^-(x)\,dx \leqq C_9 \frac{\log p}{\sqrt{p}}. \tag{20}$$

Let

$$g_l(x) = \begin{cases} \sqrt{5} v_l(5x), & x \in (0, 1), \\ 0, & \text{otherwise,} \end{cases}$$

and for $l = 1, \dots, 2p$, introduce

$$f_l(p; x) = \begin{cases} \frac{1}{\sqrt{2}} g_l(2x), & x \in \left(0, \frac{1}{2}\right), \\ -\frac{1}{\sqrt{2}} g_l\left(2\left(x - \frac{1}{2}\right)\right), & x \in \left(\frac{1}{2}, 1\right), \\ 0, & \text{otherwise.} \end{cases}$$

The latter functions are simple and they form an orthonormal system on $(0, 1)$ with

$$\int_0^1 f_l(p; x)\,dx = 0, \quad l = 1, \dots, 2p, \tag{21}$$

and by (20) we have for any $l = 1, \dots, 2p$ that

$$\int_0^1 f_l^+(p; x)\,dx \leqq C_{10} \frac{\log p}{\sqrt{p}}, \quad \int_0^1 f_l^-(p; x)\,dx \leqq C_{10} \frac{\log p}{\sqrt{p}}. \tag{22}$$

Furthermore, by (19) there is a simple set $A(p) \subseteqq (0, 1)$, i.e. $A(p)$ is the union of a finite number of intervals, such that

$$\operatorname{mes} A(p) \geqq \frac{1}{5} \tag{23}$$

and

$$\max_{1 \leqq m \leqq 2p} \left| \sum_{l=1}^m f_l(p; x) \right| \geqq C_{11} \sqrt{p} \log p, \quad x \in A(p). \tag{24}$$

Finally, for another parameter $a>1$, we introduce the functions

$$h_l(a, p; x) = \begin{cases} \sqrt{a}\, f_l(p; ax), & x\in\left(0, \dfrac{1}{a}\right), \\ 0, & \text{otherwise}, \end{cases}$$

$l=1, \ldots, 2p$. These functions are still simple on $(0, 1)$, they constitute an orthonormal system such that by (21) and (22) we have

$$\int_0^1 h_l(a, p; x)\, dx = 0, \quad l = 1, \ldots, 2p, \tag{25}$$

$$\int_0^1 h_l^+(a, p; x)\, dx \leqq C_{10}\frac{\log p}{\sqrt{ap}}, \quad \int_0^1 h_l^-(a, p; x)\, dx \leqq C_{10}\frac{\log p}{\sqrt{ap}}, \tag{26}$$

$l=1, \ldots, 2p$, and, by (23) and (24), there is a simple set $H(a, p)\subseteqq(0, 1)$ such that

$$\operatorname{mes} H(a, p) \geqq \frac{1}{5a}, \tag{27}$$

$$\max_{1\leqq m\leqq 2p}\left|\sum_{l=1}^{m} h_l(a, p; x)\right| \geqq C_{11}\sqrt{ap}\log p, \quad x\in H(a, p). \tag{28}$$

After so much preparation let $(\Omega, \mathcal{A}, P)$ be again the probability space in the proof of Theorem 4. On this space we now define a sequence $\{X_n\}_1^\infty$ of random variables and a sequence $\{E_m\}_2^\infty$ of sets in $\mathcal{A}$ such that the following conditions will be satisfied. The X_n will be simple functions and they will be pairwise uncorrelated. The events E_m will be simple sets and totally independent, and for any $m\geqq 2$ we shall have

$$P\{E_m\} \geqq \frac{1}{5(m-1)}. \tag{29}$$

Moreover, for any $m\geqq 2$, the following relations will be satisfied:

$$D^2(X_n) = \frac{2^m}{(m-1)^2}, \quad 2^{m+1} < n \leqq 2^{m+2}, \tag{30}$$

$$\max_{1\leqq k\leqq 2^{m+1}}\left|\frac{X_{2^{m+1}+1}(\omega)+\ldots+X_{2^{m+1}+k}(\omega)}{2^{m+1}+k}\right| \geqq C_{12}\sqrt{m-1}, \quad \omega\in E_m, \tag{31}$$

$$EX_l^+ \leqq C_{13}\frac{1}{\sqrt{m-1}}, \; EX_l^- \leqq C_{13}\frac{1}{\sqrt{m-1}}, \quad 2^{m+1} < l \leqq 2^{m+2}. \tag{32}$$

To begin the construction, set $X_n(\omega)=r_n(\omega)=\operatorname{sign}\sin(2^n\pi\omega)$, $\omega\in(0, 1)$, $n=1, \ldots, 2^3$. These are simple and uncorrelated. Let now m_0 be an integer, not less than two, and assume that the random variables $X_1, \ldots, X_{2^{m_0+1}}$ and the sets $E_1, \ldots, E_{m_0}$ are already defined such that the variables are simple and pairwise uncorrelated, the events are simple and independent, and relations (29)—(32) are satisfied for $m=2, \ldots, m_0$. Then a sequence $I_1, \ldots, I_r$ of pairwise disjoint intervals can be given such that $(0,1)=\bigcup_{k=1}^{r} I_k$, each of the variables X_j, $j=1,\dots,2^{m_0+1}$, is constant on any of the intervals $I_1,\dots,I_r$, and each of the sets $E_2,\dots,E_{m_0}$ appears as the union of some of the intervals $I_1,\dots,I_r$.

For a function $X(\omega)$ on $(0,1)$ and an interval $I=(a,b)\subseteq(0,1)$ let

$$X(I;\omega)=\begin{cases} X\left(\dfrac{\omega-a}{b-a}\right), & \omega\in(a,b),\\ 0, & \text{otherwise.}\end{cases}$$

Also, for a set $H\subseteq(0,1)$, let $H(I)$ be the set contained in I which is obtained from H by the application of the transformation $x\to(b-a)x+a$.

Consider now the above construction of the $h_l(a,p;x)$ functions and $H(a,p)$ sets for $p=2^{m_0}$, $a=m_0$. Let us break up the interval I_k into two disjoint intervals of equal length: $I_k=I_k'\cup I_k''$, $k=1,\dots,r$. We define the next block of variables as

$$X_{2^{m_0+1}+l}(\omega)=\sum_{k=1}^{r}\frac{\sqrt{2^{m_0}}}{m_0}h_l(m_0,2^{m_0};I_k';\omega)-\sum_{k=1}^{r}\frac{\sqrt{2^{m_0}}}{m_0}h_l(m_0,2^{m_0};I_k'';\omega),$$

$l=1,\dots,2^{m_0+1}$, and the next event as

$$E_{m_0+1}=\left(\bigcup_{k=1}^{r}H(m_0,2^{m_0};I_k')\right)\cup\left(\bigcup_{k=1}^{r}H(m_0,2^{m_0};I_k'')\right).$$

It is plain that the X_k, $k=2^{m_0+1},\dots,2^{m_0+2}$, are simple functions and E_{m_0+1} is a simple set. The events $E_2,\dots,E_{m_0+2}$ are obviously independent, and, by (27), inequality (29) is satisfied for $m=m_0+1$. Equation (25) implies that

$$EX_k=0, \tag{33}$$

$k=2^{m_0+1},\dots,2^{m_0+2}$, and hence it is evident by construction that $X_1,\dots,X_{2^{m_0+2}}$ are pairwise uncorrelated as well. Finally, (30), (31) and (32) also follow from (26) and (28) for $m=m_0+1$, on the basis of the construction.

Hence the required sequences $\{X_n\}_1^\infty$ and $\{E_m\}_2^\infty$ are obtained by induction, and (33) also holds for any $k\geqq1$. Now we claim that the constructed X_n's satisfy all the requirements of the theorem.

From (30) we get

$$\sum_{n=2^3+1}^{\infty}\frac{D^2(X_n)}{n^2}=\sum_{m=2}^{\infty}\sum_{n=2^{m+1}+1}^{2^{m+2}}\frac{D^2(X_n)}{n^2}\leqq$$

$$\leqq 2\sum_{m=2}^{\infty}\frac{1}{m^2}2^m\sum_{n=2^{m+1}+1}^{2^{m+2}}\frac{1}{n^2}\leqq 2\sum_{m=2}^{\infty}\frac{1}{m^2}<\infty.$$

This and (33) imply in the usual way that

$$\lim_{m\to\infty}\frac{X_1+\dots+X_{2^{m+1}}}{2^{m+1}}=0 \tag{34}$$

almost surely. It follows from (32) that

$$\lim_{l\to\infty} EX_l^+ = \lim_{l\to\infty} EX_l^- = 0,$$

and hence

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} E|X_n| = 0.$$

Through the second Borel—Cantelli lemma (29) gives that

$$P\{\overline{\lim_{m\to\infty}}\, E_m\} = 1.$$

Therefore it follows by (31) that the inequality

$$\max_{2^{m+1}<k\leqq 2^{m+2}} \left|\frac{X_1+\ldots+X_k}{k}\right| \geqq \max_{2^{m+1}<k\leqq 2^{m+2}} \left|\frac{X_{2^m+1}+\ldots+X_k}{k}\right| - \left|\frac{X_1+\ldots+X_{2^{m+1}}}{2^{m+1}}\right| \geqq$$

$$\geqq C_{12}\sqrt{m} - \left|\frac{X_1+\ldots+X_{2^{m+1}}}{2^{m+1}}\right|$$

holds almost surely for infinitely many m. This and (34) imply that, indeed, almost surely

$$\overline{\lim_{n\to\infty}} \left|\frac{X_1+\ldots+X_n}{n}\right| = \infty.$$

Acknowledgement

We thank Pál Révész for asking the question which is answered in Theorem 4.

References

[1] S. V. Bočkarev, Rearrangements of Fourier—Walsh series, *Izv. Akad. Nauk SSS, RSer. Mat.* **43** (1979), 1025—1041 (Russian).

[2] N. Etemadi, An elementary proof of the strong law of large numbers, *Z. Wahrscheinlichkeitstheorie verw. Gebiete,* **55** (1981), 119—122.

[3] S. Kaczmarcz, Notes on orthogonal series. II, *Studia Math.*, **5** (1943), 103—106.

[4] D. Menchoff, Sur les séries de fonctions orthogonales (Première partie), *Fundamenta Math.*, **4** (1923), 82—105.

[5] S. Nakata, On the unconditional convergence of Walsh series, *Anal. Math.*, **5** (1979), 201—205.

[6] P. Révész, *The strong laws of large numbers,* Academic Press (New York, 1967).

[7] K. Tandori, Über die Divergenz der Walshschen Reihen, *Acta Sci. Math Szeged,* **27** (1966), 261—263.

[8] K. Tandori, Bemerkung zum Gesetzt der großen Zahlen, *Periodica Math. Hung.*, **2** (1972), 33—39.

BOLYAI INSTITUTE
JÓZSEF ATTILA UNIVERSITY
ARADI VÉRTANÚK TERE 1.
H-6720 SZEGED, HUNGARY

(Received March 17, 1982; revised April 5, 1982)