# On permutation-invariance of limit theorems

I. Berkes* and R. Tichy†

**Abstract**

By a classical principle of probability theory, sufficiently thin subsequences of general sequences of random variables behave like i.i.d. sequences. This observation not only explains the remarkable properties of lacunary trigonometric series, but also provides a powerful tool in many areas of analysis, such the theory of orthogonal series and Banach space theory. In contrast to i.i.d. sequences, however, the probabilistic structure of lacunary sequences is not permutation-invariant and the analytic properties of such sequences can change after rearrangement. In a previous paper we showed that permutation-invariance of subsequences of the trigonometric system and related function systems is connected with Diophantine properties of the index sequence. In this paper we will study permutation-invariance of subsequences of general r.v. sequences.

**AMS 2000 Subject classification**. Primary 42A55, 42A61, 60F05, 60G09.

**Key words and phrases:** lacunary series, limit theorems, permutation-invariance, subsequence principle, exchangeable sequences

*Graz University of Technology, Institute of Statistics, Kopernikusgasse 24, 8010 Graz, Austria. e-mail: berkes@tugraz.at. Research supported by FWF grants P24302-N18, W1230 and OTKA grant K 108615.

†Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: tichy@tugraz.at. Research supported by FWF grants P24302-N18 and W1230.

# 1 Introduction

It is known that sufficiently thin subsequences of general r.v. sequences behave like i.i.d. sequences. For example, Révész [23] showed that if a sequence $(X_n)$ of r.v.'s satisfies $\sup_n EX_n^2 < \infty$, then one can find a subsequence $(X_{n_k})$ and a r.v. $X \in L^2$ such that $\sum_{k=1}^\infty c_k(X_{n_k} - X)$ converges a.s. provided $\sum_{k=1}^\infty c_k^2 < \infty$. Under the same condition, Gaposhkin [13], [14] and Chatterji [9], [10] proved that there exists a subsequence $(X_{n_k})$ and r.v.'s $X \in L^2$, $Y \in L^1$, $Y \geq 0$ such that

$$\frac{1}{\sqrt{N}} \sum_{k \leq N} (X_{n_k} - X) \xrightarrow{d} N(0, Y) \tag{1.1}$$

and

$$\limsup_{N\to\infty} \frac{1}{\sqrt{2N \log\log N}} \sum_{k \leq N} (X_{n_k} - X) = Y^{1/2} \qquad \text{a.s..} \tag{1.2}$$

Here $N(0,Y)$ denotes the distribution of the r.v. $Y^{1/2}\zeta$, where $\zeta$ is a standard normal r.v. independent of $Y$. Komlós [18] showed that if $\sup_n E|X_n| < \infty$, then there exists a subsequence $(X_{n_k})$ and a r.v. $X \in L^1$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N X_{n_k} = X \qquad \text{a.s..}$$

Chatterji [8] showed that if $\sup_n E|X_n|^p < \infty$ where $0 < p < 2$, then the conclusion of the previous theorem can be changed to

$$\lim_{N\to\infty} \frac{1}{N^{1/p}} \sum_{k=1}^N (X_{n_k} - X) = 0 \qquad \text{a.s.}$$

for some $X \in L^p$. Note the randomization in all these examples: the role of the mean and variance of the subsequence $(X_{n_k})$ is played by random variables $X$, $Y$. For further limit theorems for subsequences of general r.v. sequences and for the history of the topic until 1966, see Gaposhkin [13].

Since the asymptotic properties of an i.i.d. sequence do not change if we permute its terms, it is natural to expect that limit theorems for lacunary subsequences of general r.v. sequences remain valid after any permutation of their terms. This is, however, not the case. By classical results of Salem and Zygmund [24], [25] and Erdős and Gál [12], under the Hadamard gap condition

$$n_{k+1}/n_k \geq q > 1 \qquad k = 1, 2, \ldots \tag{1.3}$$

the sequence $(\sin 2\pi n_k x)$ satisfies

$$\frac{1}{\sqrt{N/2}} \sum_{k=1}^N \sin 2\pi n_k x \xrightarrow{d} N(0,1) \tag{1.4}$$

and

$$\limsup_{N\to\infty} \frac{1}{\sqrt{N \log\log N}} \sum_{k=1}^N \sin 2\pi n_k x = 1 \qquad \text{a.s.} \tag{1.5}$$

with respect to the probability space $((0,1), \mathcal{B}, \mu)$, where $\mu$ denotes the Lebesgue measure. Erdős [11] and Takahashi [27] proved that (1.4), (1.5) remain valid under the weaker gap condition

$$n_{k+1}/n_k \geq 1 + ck^{-\alpha}, \qquad k = 1, 2, \ldots \tag{1.6}$$

for $0 < \alpha < 1/2$ and that for $\alpha = 1/2$ this becomes false. As it was shown in [2], [3], under the Hadamard gap condition (1.3) the CLT (1.4) and the LIL (1.5) are permutation-invariant, i.e. they remain valid after any permutation of the sequence $(n_k)$, but this generally fails under the gap condition (1.6). Similar results hold for lacunary sequences $f(n_k x)$, where $f$ is a measurable function satisfying

$$f(x+1) = f(x), \qquad \int_0^1 f(x)\,dx = 0, \qquad \int_0^1 f^2(x)\,dx < \infty. \tag{1.7}$$

In this case, assuming the Hadamard gap condition (1.3), the validity of the CLT

$$\frac{1}{\sqrt{N}} \sum_{k=1}^{N} f(n_k x) \xrightarrow{d} N(0, \sigma^2) \tag{1.8}$$

and of its permuted version depend on the number of solutions of the Diophantine equation

$$an_k + bn_\ell = c, \qquad 1 \leq k, \ell \leq N. \tag{1.9}$$

As shown in [1], [2], [3], a sharp condition for the CLT is that the number of solutions of (1.9) is $o(N)$ for any fixed nonzero $a, b, c$, while the permuted CLT requires the stronger bound $O(1)$ for the number of solutions.

Permutation-invariance of limit theorems becomes a particularly difficult problem for parametric limit theorems, e.g. for limit theorems containing arbitrary coefficients. By a classical result of Menshov [20], from every orthonormal system $(f_n)$ one can select a subsequence $(f_{n_k})$ which is a convergence system, i.e. the series $\sum_{k=1}^\infty c_k f_{n_k}$ converges almost everywhere provided $\sum_{k=1}^\infty c_k^2 < \infty$. The question of whether a subsequence $(f_{n_k})$ exists such that this property remains valid after any permutation of $(f_{n_k})$ (i.e., by the standard terminology, $(f_{n_k})$ is an unconditional convergence system) remained open for nearly 40 years until it was answered in the affirmative by Komlós [19]. For another proof see Aldous [4]. The problem of whether every orthonormal system can be rearranged to become a convergence system is still open; for a partial result see Garsia [15]. Kolmogorov showed (see [17]) that there exists an $f \in L^2(0,1)$ whose Fourier series, suitably permuted, diverges a.e. But even though the Rademacher-Menshov convergence theorem yields a sharp a.e. convergence criterion for orthonormal series, there is no similar complete result for rearranged trigonometric series.

The previous results show that permutation-invariance of limit theorems lies substantially deeper than that of the original theorems and raise the question of which limit theorems hold in a permutation-invariant form for lacunary sequences. In this paper we will prove the surprising fact that, in a sense to be made precise, *all* nonparametric distributional limit theorems for i.i.d. random variables hold for lacunary subsequences $(f_{n_k})$ of general r.v. sequences in a permutation-invariant form provided that the subsequence is sufficiently thin, i.e. the gaps of the sequence (depending on the limit theorem) grow sufficiently rapidly. We will deduce this result from a general structure theorem for lacunary sequences proved in [6] stating that sufficiently thin subsequences of any tight sequence of random variables are nearly exchangeable. While this idea is simple and elementary, formulating our results is somewhat technical and requires some preparations in Section 2. The proof of our theorem will be given in Section 3.

# 2 Main result

We start with a formal definition of the concept "weak limit theorem". Let $\mathcal{M}$ denote the set of all probability measures on $\mathbb{R}$ and $\varrho$ the Prohorov metric on $\mathcal{M}$ defined by

$$\varrho(\nu, \lambda) = \inf\big\{\varepsilon > 0 : \nu(A) \le \lambda(A^\varepsilon) + \varepsilon \ \text{ and } \\ \lambda(A) \le \nu(A^\varepsilon) + \varepsilon \ \text{ for all Borel sets } \ A \subset \mathbb{R}\big\}.$$

Here

$$A^\varepsilon = \{x \in \mathbb{R} : |x - y| < \varepsilon \text{ for some } y \in A\}$$

denotes the open $\varepsilon$-neighborhood of $A$. A random measure is a measurable map from a probability space to $\mathcal{M}$. The following definition is due to Aldous [4].

**Definition**. *A weak limit theorem of i.i.d. random variables is a system*

$$T = (f_1, f_2, \ldots, S, \{G_\mu, \mu \in S\})$$

*where*

*(a) $S$ is a Borel subset of $\mathcal{M}$;*

*(b) For each $k \ge 1$, $f_k = f_k(x_1, x_2, \ldots, \mu)$ is a continuous function on $\mathbb{R}^\infty \times \mathcal{M}$, satisfying the Lipschitz condition*

$$|f_k(x_1, x_2, \ldots, \mu) - f_k(x'_1, x'_2, \ldots, \mu)| \le \sum_{i=1}^{\infty} c_{k,i} |x_i - x'_i|$$

*where $0 \le c_{k,i} \le 1$ and $\lim_{k\to\infty} c_{k,i} = 0$ for all $i$;*

*(c) For each $\mu \in S$, $G_\mu$ is a probability distribution on $\mathbb{R}$ such that the function $\mu \to G_\mu$ is measurable (with respect to the Borel $\sigma$-fields in $S$ and $\mathcal{M}$);*

*and*

*(d) If $\mu \in S$ and $X_1, X_2, \ldots$ are independent r.v.'s with common distribution $\mu$ then*

$$f_k(X_1, X_2, \ldots, \mu) \xrightarrow{d} G_\mu \quad \text{as } k \to \infty. \tag{2.1}$$

For example, the central limit theorem corresponds to

$$S = \{\mu \in \mathcal{M} : \int x^2 d\mu(x) < \infty\}, \qquad G_\mu = N(0, \operatorname{Var} \mu),$$

$$f_k(x_1, x_2, \dots, \mu) = (x_1 + \dots + x_k - k \cdot E\mu)/\sqrt{k}, \qquad c_{k,i} = k^{-1/2} I_{\{i \le k\}}.$$

The theorem itself is expressed by (2.1).

Using the terminology of [7], we call a sequence $(X_n)$ of random variables *determining* if it has a limit distribution relative to any set $A$ in the probability space with $P(A) > 0$, i.e. for any $A \subset \Omega$ with $P(A) > 0$ there exists a distribution function $F_A$ such that

$$\lim_{n \to \infty} P(X_n \le t \mid A) = F_A(t)$$

for all continuity points $t$ of $F_A$. Here $P(\cdot|A)$ denotes conditional probability given $A$. (This concept is the same as that of stable convergence, introduced by Rényi [22]; our terminology follows that of functional analysis.) By an extension of the Helly-Bray theorem (see [7]), every tight sequence of r.v.'s contains a determining subsequence. As is shown in [4], [7], for any determining sequence $(X_n)$ there exists a random measure $\tilde{\mu}$ (i.e. a measurable map from the underlying probability space $(\Omega, \mathcal{F}, \mathcal{P})$ to $\mathcal{M}$) such that for any $A$ with $P(A) > 0$ and any continuity point $t$ of $F_A$ we have

$$F_A(t) = \mathbb{E}_A(\tilde{\mu}(-\infty, t]) \tag{2.2}$$

where $\mathbb{E}_A$ denotes conditional expectation given $A$. We call $\tilde{\mu}$ the *limit random measure* of $(X_n)$.

The following result is Aldous' celebrated subsequence theorem [4].

**Theorem 2.1** *Let $(X_n)$ be a determining sequence with limit random measure $\tilde{\mu}$. Let $T = (f_1, f_2, \dots, S, \{G_\mu, \mu \in S\})$ be a weak limit theorem and assume $P(\tilde{\mu} \in S) = 1$. Then there exists a subsequence $(X_{n_k})$ such that*

$$f_k(X_{n_1}, X_{n_2}, \dots, \tilde{\mu}) \xrightarrow{d} \int G_{\tilde{\mu}} dP. \tag{2.3}$$

In case of the CLT formalized above, assuming $\sup_n \mathbb{E}X_n^2 < +\infty$ implies easily that $\tilde{\mu}$ has finite variance almost surely and thus denoting its mean and variance by $X$ and $Y$, respectively, we see that the integral in (2.3) is the distribution $N(0, Y)$. Hence (2.3) states in the present case that

$$\frac{1}{\sqrt{N}} \sum_{k=1}^{N} (X_{n_k} - X) \xrightarrow{d} N(0, Y)$$

which is exactly the CLT of Chatterji [9] and Gaposhkin [14] formulated in the Introduction. Theorem 2.1 shows that a similar subsequence theorem holds for any weak limit theorem of i.i.d. random variables. For a version of this result for strong (a.s.) limit theorems, we refer to Aldous [4].

In what follows, we change the technical conditions on $f_k$ in the definition of weak limit theorems slightly, leading to a class more convenient for our purposes.

**Definition.** *The limit theorem $T = (f_1, f_2, \dots, S, \{G_\mu, \mu \in S\})$ is called regular if there exist two sequences $p_k \le q_k$ of positive integers tending to $+\infty$ and a sequence $\omega_k \to +\infty$ such that*

*(i)* $f_k(x_1, x_2, \ldots, \mu)$ *depends only on* $x_{p_k}, \ldots, x_{q_k}, \mu$
*(ii)* $f_k$ *satisfies the Lipschitz condition*

$$|f_k(x_{p_k}, \ldots, x_{q_k}, \mu) - f_k(x'_{p_k}, \ldots, x'_{q_k}, \mu')| \leq \frac{1}{\omega_k} \sum_{i=p_k}^{q_k} |x_i - x'_i|^\alpha + \varrho^*(\mu, \mu') \tag{2.4}$$

*for some* $0 < \alpha \leq 1$ *where* $\varrho^*$ *is a metric on* $S$ *generating the same topology as the Prohorov metric* $\varrho$.

Thus in this case the function $f_k$ depends only on a finite segment $x_{p_k}, \ldots x_{q_k}$ of the variables $x_1, x_2, \ldots$ . On the role of $\varrho^*$ see [4]. The above definition brings out clearly the crucial feature of limit theorems, namely the fact that the validity of the theorem does not depend on finitely many terms of $(X_n)$, while the original definition assumes only that the dependence of $f_k(X_1, X_2, \ldots)$ on any fixed variable $X_j$ of the sequence is weak if $k$ is large. However, there is very little difference between these assumptions. For example, the central limit theorem can be formalized by either of the functions

$$f_k(x_1, \ldots, x_k, \mu) = (x_1 + \ldots + x_k - k \cdot E\mu)/\sqrt{k}$$

and

$$f_k^*(x_{[k^{1/4}]}, \ldots, x_k, \mu) = (x_{[k^{1/4}]} + \ldots + x_k - k \cdot E\mu)/\sqrt{k}$$

of which the second leads to a regular limit theorem with the Wasserstein metric

$$\varrho^*(\mu, \mu') = \left( \int_0^1 |F_\mu^{-1}(x) - F_{\mu'}^{-1}(x)|^2 dx \right)^{1/2},$$

where $F_\mu, F_{\mu'}$ denote the distribution function of $\mu$ and $\mu'$, respectively. Under bounded second moments, the contribution of the first $k^{1/4}$ terms in the normed sum defining $f_k$ are irrelevant and thus we can always switch from $f_k$ to $f_k^*$ and back again. The same procedure applies in the general case.

We are now in a position to formulate the main result of our paper.

**Theorem 2.2** *Let* $(X_n)$ *be a determining sequence with limit random measure* $\tilde{\mu}$. *Let* $T = (f_1, f_2, \ldots, S, \{G_\mu, \mu \in S\})$ *be a regular weak limit theorem and assume that* $P(\tilde{\mu} \in S) = 1$. *Then there exists a subsequence* $(X_{n_k}) = (Y_k)$ *such that for any permutation* $(Y'_k)$ *of* $(Y_k)$ *we have*

$$f_k(Y'_1, Y'_2, \ldots, \tilde{\mu}) \xrightarrow{d} \int G_{\tilde{\mu}} dP. \tag{2.5}$$

Note that we assumed the regularity of the limit theorem, but as we pointed out before, this is no restriction of generality.

The limit theorem $T$ in Theorem 2.2 is nonparametric, i.e. the function $f_k$ depends on $x_1, x_2, \ldots$ and $\mu$, but on no additional parameters. A simple example of a parametric distributional limit theorem is the weighted CLT, where

$$f_k = A_k^{-1} \sum_{j=1}^k a_j (x_j - E\mu), \qquad A_k = \left( \sum_{j=1}^k a_j^2 \right)^{1/2}.$$

For any fixed coefficient sequence $(a_k)$ this defines a nonparametric limit theorem $T$ and Theorem 2.2 applies, but the selected subsequence $(X_{n_k})$ depends on $(a_k)$. As the discussion above shows, in the case of a parametric limit theorem $T$ deciding whether a universal subsequence $(X_{n_k})$ working for all parameters is generally a very difficult problem; an example of a limit theorem where such a choice is impossible is given in [16]. For this reason, in the present paper we deal only with nonparametric limit theorems.

In Aldous [4] a formalization of strong limit theorems is also given and the analogue of Theorem 2.1 is proved. Using a reformulation of strong limit theorems as a sequence of probability inequalities as given in [5], [6], a version of our Theorem 2.2 can be given for a subclass of limit theorems considered in [4]. We also mention that for a more limited class of weak limit theorems Theorem 2.2 was proved in [21].

# 3 Proof of Theorem 2.2.

To simplify the formulas, let $f_k(\mu)$ denote, for any $\mu \in S$, the distribution of the random variable $f_k(\xi_1, \xi_2, \ldots, \mu)$ where $\xi_1, \xi_2, \ldots$ are independent r.v.'s with common distribution $\mu$. The following statements are easy to verify:

(A) If $\varrho(\mu, \nu) \le \varepsilon$ then $\varrho(f_k(\mu), f_k(\nu)) \le \varepsilon^\alpha q_k + \varrho^*(\mu, \nu)$ where $\alpha, q_k$ and $\varrho^*$ are the quantities appearing in (2.4).

(B) Let $\mu_1, \ldots, \mu_r$ and $\nu_1, \ldots \nu_r$ be probability distributions, further let $c_1, \ldots, c_r$ be nonnegative numbers with $\sum_{i=1}^r c_i = 1$. Assume that the sum of those $c_i$'s such that $\varrho(\mu_i, \nu_i) \ge \varepsilon$ is at most $\varepsilon$. Then the Prohorov distance between $\sum\limits_{i=1}^r c_i \mu_i$ and $\sum\limits_{i=1}^r c_i \nu_i$ is at most $2\varepsilon$.

(C) Let $\tilde{\mu}$ and $\tilde{\nu}$ be random measures (i.e. measurable maps from a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ to $\mathcal{M}$) such that $P(\varrho(\tilde{\mu}, \tilde{\nu}) \ge \varepsilon) \le \varepsilon$. Then the Prohorov distance between $\int \tilde{\mu} dP$ and $\int \tilde{\nu} dP$ is $\le 2\varepsilon$.

To prove statement (A) note that if $\varrho(\mu, \nu) \le \varepsilon$ then by a theorem of Strassen [26] there exist, on some probability space, r.v.'s $\xi$ and $\eta$ with distribution $\mu$ and $\nu$ such that $P(|\xi - \eta| \ge \varepsilon) \le \varepsilon$. On a larger probability space, let $(\xi_n, \eta_n)$ $(n = 1, 2, \ldots)$ be independent random vectors distributed as $(\xi, \eta)$. Clearly $P(|\xi_i - \eta_i| \ge \varepsilon) \le \varepsilon$ $(i = 1, 2, \ldots)$ and thus using (2.4) we see that

$$\left| f_k(\xi_{p_k}, \ldots, \xi_{q_k}, \mu) - f_k(\eta_{p_k}, \ldots, \eta_{q_k}, \nu) \right| \le \varepsilon^\alpha q_k + \varrho^*(\mu, \nu)$$

except on a set with probability $\le \varepsilon q_k \le \varepsilon^\alpha q_k$, proving (A). (Clearly we can assume $0 < \varepsilon \le 1$ and that in the definition of regular limit theorems we have $\omega_k \ge 1$ for all $k$.) Statements (B) and (C) are almost evident, (B) is a special case of (C).

To prove our theorem, let $(X_n)$ be a determining sequence of r.v.'s with limit random measure $\tilde{\mu}$. Then $(X_n)$ is tight, i.e. $\sup_j P(|X_j| \ge t) \to 0$ as $t \to \infty$. As $\omega_k \to +\infty$, we can choose a nondecreasing sequence $(r_k)$ of integers tending to $+\infty$ so slowly that

$$r_k \le \min(p_k - 1, \omega_k^{1/4}) \tag{3.1}$$

and

$$\sup_j P\Big(|X_j| \geq \frac{1}{2}\omega_k^{1/(4\alpha)}\Big) \leq \frac{1}{2} r_k^{-2} \qquad (k \geq 1). \tag{3.2}$$

Let $(\varepsilon_k)$ tend to 0 monotonically and so rapidly that

$$\varepsilon_{r_k}^{\alpha} q_k \leq k^{-1}. \tag{3.3}$$

Using the structure theorem [6, Theorem 2], it follows that there exists a subsequence $(X_{n_k})$ and a sequence $(X_k')$ of r.v.'s such that

$$|X_{n_k} - X_k'| = O(2^{-k}) \quad \text{a.s.} \tag{3.4}$$

and $X_k'$ has the following properties:

($A_1$) Each $X_k'$ takes only finitely many values

($B_1$) $\sigma\{X_1'\} \subset \sigma\{X_2'\} \subset \dots$

($C_1$) For each $k \geq 1$ the atoms of the finite $\sigma$-field $\sigma\{X_{r_k}'\}$ can be divided into two classes $\Gamma_1$ and $\Gamma_2$ so that

$$\sum_{A \in \Gamma_1} P(A) \leq \varepsilon_{r_k} \tag{3.5}$$

and for any $A \in \Gamma_2$ there exist i.i.d.r.v.'s $\{Z_j^{(A)}, j = r_k + 1, r_k + 2, \dots\}$ defined on $A$ with distribution function $F_A$ such that

$$P_A\big(|X_j' - Z_j^{(A)}| \geq \varepsilon_{r_k}\big) \leq \varepsilon_{r_k} \quad j = r_k + 1, r_k + 2, \dots. \tag{3.6}$$

Here $F_A$ denotes the limit distribution of $(X_n)$ on the set $A$ (which exists since $(X_n)$ is determining) and $P_A$ denotes conditional probability with respect to $A$. Let $\tilde{\mu}_n$ denote the random measure defined by $\tilde{\mu}_n(B) = \mathbb{E}(\tilde{\mu}(B) \mid X_n')$. By Lemma 7 of [6] we have $\tilde{\mu}_n \xrightarrow{d} \tilde{\mu}$ a.s. and thus by passing to a further subsequence of $(X_{n_k})$ we can also assume that

$$P\{\varrho(\tilde{\mu}_n, \tilde{\mu}) \geq \varepsilon_n\} \leq \varepsilon_n \tag{3.7}$$

$$P\{\varrho^*(\tilde{\mu}_n, \tilde{\mu}) \geq \varepsilon_n\} \leq \varepsilon_n. \tag{3.8}$$

We show that the last obtained subsequence $(X_{n_k})$ satisfies the conclusion of the theorem. In view of (2.4) and (3.4), $X_{n_k}$ and $X_k'$ are interchangeable in the statement of the theorem and thus it suffices to prove that if $(X_k')$ satisfies statements ($A_1$), ($B_1$), ($C_1$) above then for any permutation $(Y_k')$ of $(X_k')$ we have (2.5). To verify this, note that by (2.4) and (3.6) we have

$$\begin{aligned} P_A\big\{\big|f_k(X_{i_1}', \dots, X_{i_\ell}', \mu_A) - f_k(Z_{i_1}^{(A)}, \dots, Z_{i_\ell}^{(A)}, \mu_A)\big| \geq \varepsilon_{r_k}^{\alpha} q_k\big\} \\ \leq \varepsilon_{r_k}^{\alpha} q_k \qquad A \in \Gamma_2 \end{aligned} \tag{3.9}$$

where $\ell = q_k - p_k + 1$, $i_1, \dots, i_\ell$ are different integers $> r_k$ and $\mu_A$ is the probability measure corresponding to $F_A$. (Note that we do not assume here $i_1 < \dots < i_\ell$ ; the vectors $(X_{i_1}' \dots, X_{i_\ell}')$ and $(Z_{i_1}^{(A)}, \dots, Z_{i_\ell}^{(A)})$ are close to each other coordinatewise, i.e. for any order of $i_1, \dots, i_\ell$. Since the $Z_j^{(A)}$ are i.i.d., the distribution of the vector $(Z_{i_1}^{(A)}, \dots, Z_{i_\ell}^{(A)})$ is permutation-invariant, providing an explanation for the phenomenon described in Theorem 2.2.) Since (3.9) is valid for all $A \in \Gamma_2$ and $\mu_A$ in (3.9) is identical to $\tilde{\mu}_{r_k}$ on $A$ (see Lemma 6 of [5]), using (3.5), (3.9) and statement (B) at the beginning of the proof we get

$$\varrho\big(f_k(X'_{i_1}, \dots, X'_{i_\ell}, \tilde{\mu}_{r_k}), \sum_A f_k(\mu_A)P(A)\big) \le 2\varepsilon_{r_k}^\alpha q_k \tag{3.10}$$

where the sum is extended for all atoms $A$ of $\sigma\{X'_{r_k}\}$ and a r.v. in a Prohorov distance is meant as its distribution. Next we show that (3.10) remains valid, with the right hand side increased by $r_k^{-1}$, if $i_1, \dots, i_\ell$, $\ell = q_k - p_k + 1$, are arbitrary different positive integers (not necessarily $> r_k$). Indeed, remove from $X'_{i_1}, \dots, X'_{i_\ell}$ those whose index is $\le r_k$ and replace them with (different) $X'_j$'s with $j > \max(r_k, i_1, \dots, i_\ell)$. This means that we change $f_k(X'_{i_1}, \dots, X'_{i_\ell}, \tilde{\mu}_{r_k})$ at most at $r_k$ locations and at each such position we replace an $X'_\mu$ by an $X'_\nu$ where $\mu \le r_k$ and $\nu > r_k$. By (2.4), $f_k$ changes at most by

$$\frac{1}{\omega_k} \sum |X'_\mu - X'_\nu|^\alpha =: W$$

where the sum has $\le r_k$ terms. Using (3.1), (3.2) we get

$$P(|W| \ge r_k^{-1}) \le P(|W| \ge \omega_k^{-1/2}) = P\left(\sum |X'_\mu - X'_\nu|^\alpha \ge \omega_k^{1/2}\right)$$

$$\le \sum P\left(|X'_\mu - X'_\nu| \ge \left(\frac{\omega_k^{1/2}}{r_k}\right)^{1/\alpha}\right) \le 2r_k \cdot \sup_j P\left(|X'_j| \ge \frac{1}{2}\omega_k^{1/(4\alpha)}\right) \le r_k^{-1}$$

and thus the above changes increase the left hand side of (3.10) by at most $r_k^{-1}$, i.e.

$$\varrho\Big(f_k(X'_{i_1}, \dots, X'_{i_\ell}, \tilde{\mu}_{r_k}), \sum_A f_k(\mu_A)P(A)\Big) \le 2\varepsilon_{r_k}^\alpha q_k + r_k^{-1} \tag{3.11}$$

for any different positive integers $i_1, \dots, i_\ell$, $\ell = q_k - p_k + 1$. Changing $\tilde{\mu}_{r_k}$ into $\tilde{\mu}$ will change $f_k(X'_{i_1}, \dots, X'_{i_\ell}, \tilde{\mu}_{r_k})$ on the left hand side of (3.11) by at most $\varepsilon_{r_k}$, except on a set of probability $\le \varepsilon_{r_k}$ (see (3.8) and (2.4)) and thus the left hand side of (3.11) changes by at most $\varepsilon_{r_k}$. Thus observing that the sum $\sum_A f_k(\mu_A)P(A)$ in (3.11) equals $\int f_k(\tilde{\mu}_{r_k})dP$, we proved the following

**Proposition**. Let $(X_k^*)$ be any permutation of $(X'_k)$. Then

$$\varrho\big(f_k(X_{p_k}^*, \dots, X_{q_k}^*, \tilde{\mu}), \int f_k(\tilde{\mu}_{r_k})dP\big) \le 3\varepsilon_{r_k}^\alpha q_k + r_k^{-1}.$$

To complete the proof of our theorem it suffices to show that the Prohorov distance of any two of the distributions

$$\int f_k(\tilde{\mu}_{r_k})dP \qquad \int f_k(\tilde{\mu})dP \qquad \int G_{\tilde{\mu}}dP \tag{3.12}$$

tends to zero as $k \to \infty$. To verify this observe first that (3.7), (3.8) and statement (A) at the beginning of the proof imply that the Prohorov distance of $f_k(\tilde{\mu}_{r_k})$ and

$f_k(\tilde{\mu})$ is $\le \varepsilon_{r_k}^{\alpha} q_k + \varepsilon_{r_k}$, except on a set with probability $\le \varepsilon_{r_k}^{\alpha} q_k + \varepsilon_{r_k}$ and thus by statement (C) and (2.2) the Prohorov distance of the first two distributions in (3.12) is $\le 2(\varepsilon_{r_k}^{\alpha} q_k + \varepsilon_{r_k}) \le 4k^{-1}$. On the other hand, the validity of $f_k(\mu) \xrightarrow{d} G_\mu$ for any $\mu \in S$ (which is a part of the definition of a weak limit theorem) and $P(\tilde{\mu} \in S) = 1$ imply $\varrho(f_k(\tilde{\mu}), G_{\tilde{\mu}}) \to 0$ a.s. and thus there exists a numerical sequence $\delta_k \downarrow 0$ such that

$$P\{\varrho(f_k(\tilde{\mu}), G_{\tilde{\mu}}) \ge \delta_k\} \le \delta_k \qquad (k = 1, 2, \ldots).$$

Thus by statement (C) above we get that the Prohorov distance of the second and third distribution in (3.12) is $\le 2\delta_k$. This completes the proof of Theorem 2.2.

**Acknowledment.** We would like to thank two anonymous referees for their remarks leading to a substantial improvement of the presentation.

# References

[1] C. Aistleitner and I. Berkes. On the central limit theorem for $f(n_k x)$. Probab. Theory Rel. Fields 146 (2010), 267–289.

[2] C. Aistleitner, I. Berkes and R. Tichy. On the law of the iterated logarithm for permuted lacunary sequences. *Proc. Steklov Inst. Math.* **276** (2012), 3–20.

[3] C. Aistleitner, I. Berkes and R. Tichy. On permutations of lacunary series. *RIMS Kôkyûroku Bessatsu* **B34** (2012), 1–25.

[4] D. J. Aldous. Limit theorems for subsequences of arbitrarily-dependent sequences of random variables, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **40** (1977), 59–82.

[5] I. Berkes. An extension of the Komlós subsequence theorem. *Acta Math. Hung.* **55** (1990) 103-110.

[6] I. Berkes and E. Péter. Exchangeable random variables and the subsequence principle, *Prob. Theory Rel. Fields* **73** (1986), 395–413.

[7] I. Berkes and H. P. Rosenthal. Almost exchangeable sequences of random variables, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **70** (1985), 473–507.

[8] S. D. Chatterji. A general strong law. *Invent. Math.* **9** 1969/1970 235–245.

[9] S. D. Chatterji. A principle of subsequences in probability theory: The central limit theorem. *Adv. Math.* **13** (1974), 31–54.

[10] S. D. Chatterji. A subsequence principle in probability theory II. The law of the iterated logarithm. *Invent. Math.* **25** (1974), 241–251.

pp. 72–89. Springer, 1972.

[11] P. Erdős. On trigonometric sums with gaps. *Magyar Tud. Akad. Mat. Kut. Int. Közl.* **7** (1962), 37–42.

[12] P. Erdős and I.S. Gál. On the law of the iterated logarithm. *Proc. Nederl. Akad. Wetensch.* Ser A **58**, 65-84, 1955.

[13] V. F. Gaposhkin. Lacunary series and independent functions. *Russian Math. Surveys* **21** (1966), 3-82.

[14] V. F. Gaposhkin. Convergence and limit theorems for subsequences of random variables. (Russian) *Teor. Verojatnost. i Primenen.* **17** (1972), 401–423.

[15] A. M. Garsia. Existence of almost everywhere convergent rearrangements for Fourier series of $L_2$ functions. *Ann. of Math.* 79 (1964), 623–629.

[16] S. Guerre and Y. Raynaud. On sequences with no almost symmetric subsequence. *Texas Functional Analysis Seminar 1985–1986 Longhorn Notes, Univ. of Texas* pp. 83–93. Austin, 1986.

[17] A. N. Kolmogorov and D. Menshov. Sur la convergence des series de fonctions orthogonales. *Math. Z.* 26 (1927), 432–441.

[18] J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* **18**, 217–229, (1967)

[19] J. Komlós. Every sequence converging to 0 weakly in $L_2$ contains an unconditional convergence sequence. *Ark. Mat.* 12 (1974), 41-49.

[20] D. E. Menshov. Sur la convergence et la sommation des se'ries de fonctions orthogonales. *Bull. Soc. Math. France* 60 (1936), 147–170.

[21] E. Péter, An extension of the subsequence principle. *Studia Sci. Math. Hung.* **36** (2000), 335–345.

[22] A. Rényi, On stable sequences of events. Sankhya Ser. A **25** (1963), 293–302.

[23] P. Révész, On a problem of Steinhaus. *Acta Math. Acad. Sci. Hung.* **16** (1965), 311–318.

[24] R. Salem and A. Zygmund. On lacunary trigonometric series, *Proc. Nat. Acad. Sci. USA* **33** (1947), 333–338.

[25] R. Salem and A. Zygmund, La loi du logarithme itéré pour les séries trigonométriques lacunaires. *Bull. Sci. Math.* **74**, 209–224, 1950.

[26] V. Strassen. The existence of probability measures with given marginals, *Ann. Math. Statist.* **36** (1965), 423–439.

[27] S. Takahashi. On the law of the iterated logarithm for lacunary trigonometric series. *Tohoku Math. J.* **24** (1972), 319–329.