The Kadec-Pełczynski theorem in $L^p$, $1 \leq p < 2$

I. Berkes* and R. Tichy†

Abstract

By a classical result of Kadec and Pełczynski (1962), every normalized weakly null sequence in $L^p$, $p > 2$ contains a subsequence equivalent to the unit vector basis of $\ell^2$ or to the unit vector basis of $\ell^p$. In this paper we investigate the case $1 \leq p < 2$ and show that a necessary and sufficient condition for the first alternative in the Kadec-Pełczynski theorem is that the limit random measure $\mu$ of the sequence satisfies $\int_{\mathbb{R}} x^2 d\mu(x) \in L^{p/2}$.

1 Introduction

Call two sequences $(x_n)$ and $(y_n)$ in a Banach space $(B, \| \cdot \|)$ equivalent if there exists a constant $K > 0$ such that

$$K^{-1} \sum_{i=1}^{n} a_i x_i \leq \sum_{i=1}^{n} a_i y_i \leq K \sum_{i=1}^{n} a_i x_i$$

for every $n \geq 1$ and every $(a_1, \ldots, a_n) \in \mathbb{R}^n$. By a classical theorem of Kadec and Pełczynski [11], any normalized weakly null sequence $(x_n)$ in $L^p(0,1)$, $p > 2$ has a subsequence equivalent to the unit vector basis of $\ell^2$ or to the unit vector basis of $\ell^p$. In the case when $\{|x_n|^p, n \geq 1\}$ is uniformly integrable, the first alternative holds, while if the functions $(x_n)$ have disjoint support, the second alternative holds trivially. The general case follows via a subsequence splitting argument as in [11].

The purpose of the present paper is to investigate the case $1 \leq p < 2$ and to give a necessary and sufficient condition for the first alternative in the Kadec-Pełczynski theorem. To formulate our result, we use probabilistic terminology. Let $1 \leq p < 2$ and let $(X_n)$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$: assume that $\{|X_n|^p, n \geq 1\}$ is uniformly integrable and $X_n \to 0$ weakly in $L^p$. (This is meant as $\lim_{n \to \infty} E(X_n Y) = 0$ for all $Y \in L^q$ where $1/p + 1/q = 1$.

To avoid confusion with weak convergence of probability measures and distributions,

*Graz University of Technology, Institute of Statistics, Kopernikusgasse 24, 8010 Graz, Austria. e-mail: berkes@tugraz.at. Research supported by FWF grant P24302-N18 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 grant SFB F5510.
the latter will be called convergence in distribution and denoted by $\xrightarrow{D}$. Using the terminology of [5], 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

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

for all continuity points $t$ of $F_A$. Here $P(\cdot \mid A)$ denotes conditional probability given $A$. (This concept is the same as that of stable convergence, introduced in [15].) Since $\{\|X_n\|^p, n \geq 1\}$ is uniformly integrable, the sequence $(X_n)$ is tight and thus by an extension of the Helly-Bray theorem (see e.g. [5]), it contains a determining subsequence. Hence in the sequel we can assume, without loss of generality, that the sequence $(X_n)$ itself is determining. As is shown in [1, 5], for any determining sequence $(X_n)$ there exists a random measure $\mu$ (i.e. a measurable map from $(\Omega, \mathcal{F}, P)$ to $(\mathcal{M}, \pi)$, where $\mathcal{M}$ is the set of probability measures on $\mathbb{R}$ and $\pi$ is the Prohorov distance, see Section 3) such that for any $A$ with $P(A) > 0$ and any continuity point $t$ of $F_A$ we have

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

(1.1)

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

**Theorem 1.1** Let $1 \leq p < 2$ and let $(X_n)$ be a determining sequence of random variables such that $\|X_n\|_p = 1$ ($n = 1, 2, \ldots$), $\{\|X_n\|^p, n \geq 1\}$ is uniformly integrable and $X_n \to 0$ weakly in $L^p$. Let $\mu$ be the limit random measure of $(X_n)$. Then there exists a subsequence $(X_{n_k})$ equivalent to the unit vector basis of $l^2$ if and only if

$$\int_{-\infty}^{\infty} x^2 d\mu(x) \in L^{p/2}.$$  

(1.2)

By assuming the uniform integrability of $\|X_n\|^p$, we exclude "spike" situations leading to a subsequence equivalent to the unit vector basis of $\ell^p$ as in the Kadec-Pelczynski theorem. It is easily seen that (1.2) (and in fact $\int_{-\infty}^{\infty} x^2 d\mu(x) < \infty$ a.s.) imply that for any $\delta > 0$ there exists a set $A \subset \Omega$ with $P(A) \geq 1 - \delta$ and a subsequence $(X_{n_k})$ such that

$$\sup_{k \geq 1} \int_A |X_{n_k}|^2 dP < \infty.$$  

Thus the first alternative in the Kadec-Pelczynski theorem 'almost' implies bounded $L^2$ norms.

Call a sequence $(X_n)$ of random variables in $L^p$ almost symmetric if for any $\varepsilon > 0$ there exists a $K = K(\varepsilon)$ such that for any $k \geq 1$, any indices $j_1 > j_2 > \ldots j_k \geq K$, any permutation $(\sigma(j_1), \ldots, \sigma(j_k))$ of $(j_1, \ldots, j_k)$ and any $(a_1, \ldots, a_k) \in \mathbb{R}^k$ we have

$$(1 - \varepsilon)\| \sum_{i=1}^{k} a_i X_{j_i} \|_p \leq \| \sum_{i=1}^{k} a_i X_{\sigma(j_i)} \|_p \leq (1 + \varepsilon)\| \sum_{i=1}^{k} a_i X_{j_i} \|_p.$$
Once in Theorem 1.1 we found a subsequence \((X_{n_k})\) equivalent to the unit vector basis of \(\ell^2\), a result of Guerre [9] implies the existence of a further subsequence \((X_{m_k})\) of \((X_{n_k})\) which is almost symmetric. Note that this conclusion also follows from the proof of Theorem 1.1. Guerre and Raynaud [10] also showed that for any \(1 \leq p < q < 2\) there exists a sequence \((X_n)\) in \(L^p\), equivalent to the unit vector basis of \(\ell^q\), but not having an almost symmetric subsequence. No characterization for the existence of almost symmetric subsequences of \((X_n)\) in terms of the limit random measure of \((X_n)\) or related quantities is known.

2 Some lemmas

The necessity of the proof of Theorem 1.1 depends on a general structure theorem for lacunary sequences proved in [3] (see Theorem 2 of [3] and the definition preceding it); for the convenience of the reader we state it here as a lemma.

Lemma 2.1. Let \((X_n)\) be a determining sequence of r.v.’s and \((\varepsilon_n)\) a positive numerical sequence tending to 0. Then, if the underlying probability space is rich enough, there exists a subsequence \((X_{m_k})\) and a sequence \((Y_k)\) of discrete r.v.’s such that

\[
P(|X_{m_k} - Y_k| \geq \varepsilon_k) \leq \varepsilon_k \quad k = 1, 2, \ldots \tag{2.1}
\]

and for each \(k > 1\) the atoms of the \(\sigma\)-field \(\sigma\{Y_1, \ldots, Y_{k-1}\}\) can be divided into two classes \(\Gamma_1\) and \(\Gamma_2\) such that

(i) \(\sum_{B \in \Gamma_1} P(B) \leq \varepsilon_k;\)

(ii) For any \(B \in \Gamma_2\) there exist \(P_B\)-independent r.v.’s \(\{Z_j^{(B)}\}, j = k, k+1, \ldots\) defined on \(B\) with common distribution function \(F_B\) such that

\[
P_B(|Y_j - Z_j^{(B)}| \geq \varepsilon_k) \leq \varepsilon_k \quad j = k, k+1, \ldots \tag{2.2}
\]

Here \(F_B\) denotes the limit distribution of \((X_n)\) relative to \(B\) and \(P_B\) denotes conditional probability given \(B\).

Note that, instead of (2.1), in Theorem 2 of [3] the conclusion is \(\sum_{k=1}^{\infty} |X_{m_k} - Y_k| < \infty\) a.s., but after a further thinning, (2.1) will also hold. The phrase ”the underlying probability space is rich enough” is meant in Lemma 2.1 in the sense that on the underlying space there exists a sequence of independent r.v.’s, uniformly distributed over \((0, 1)\) and also independent of the sequence \((X_n)\). Clearly, this condition can be guaranteed by a suitable enlargement of the probability space not changing the distribution of \((X_n)\) and \(\mu\) and thus this assumption can be assumed without loss of generality.

Lemma 2.1 means that every tight sequence of r.v.’s has a subsequence which can be closely approximated by an exchangeable sequence having a very simple structure, namely which is i.i.d. on each set of a suitable partition of the probability space. This fact is an ”effective” form of the general subsequence principle of Aldous [1] (for a related result see Berkes and Rosenthal [5]) and reduces the studied problem to the i.i.d. case which will be handled by the classical concentration technique of Lévy [12], as improved by Esseen [7].
Lemma 2.2 Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with distribution function $F$ and put $S_n = X_1 + \cdots + X_n$. Then for any $t > 0$ we have

$$P\left( \left| S_n \right| \leq t \right) \leq A \frac{t}{\sqrt{n}} \left[ \int_{|x| < t} x^2 dF(x) - 2 \left( \int_{|x| < t} x dF(x) \right)^2 \right]^{-1/2}$$

(2.3)

provided the difference on the right-hand side is positive and $\int_{|x| < t} dF(x) \geq 1/2$. Here $A$ is an absolute constant.

Proof. Let $F^*$ denote the distribution function obtained from $F$ by symmetrization. The left hand side of (2.3) is clearly bounded by $Q_{S_n}(2t)$, where $Q_{S_n}(\lambda) = \sup_x P(x \leq S_n \leq x + \lambda)$ is the concentration function of $S_n$. By a well known concentration function inequality of Esseen (see [7], formula (3.3)) we have

$$Q_{S_n}(\lambda) \leq A n^{-1/2} \left( \int_{|x| < \lambda} x^2 dF^*(x) + \int_{|x| \geq \lambda} dF^*(x) \right)^{-1/2}$$

(2.4)

for any $\lambda > 0$, where $A$ is an absolute constant. Thus the left hand side of (2.3) is bounded by the last expression in (2.4) with $\lambda = 2t$ and thus to prove (2.3) it suffices to show that

$$\int_{|x| < t} x^2 dF^*(x) \geq \int_{|x| < t} x^2 dF(x) - 2 \left( \int_{|x| < t} x dF(x) \right)^2.$$  

(2.5)

Let $\xi$ and $\eta$ be independent r.v.’s with distribution function $F$, set

$$C = \{ |\xi - \eta| < 2t \}, \quad D = \{ |\xi| < t, |\eta| < t \}.$$

Then

$$\int_{|\xi| < 2t} x^2 dF^*(x) = \int_C (\xi - \eta)^2 dP \geq \int_D (\xi - \eta)^2 dP$$

$$= 2 \int_{|\xi| < t} \xi^2 dP \cdot P(|\eta| < t) - 2 \left( \int_{|\xi| < t} \xi dP \right)^2 \geq \int_{|\xi| < t} \xi^2 dP - 2 \left( \int_{|\xi| < t} \xi dP \right)^2$$

since $P(|\eta| < t) \geq 1/2$. Thus (2.5) is valid.

Lemma 2.3 Let $(X_n)$ be a determining sequence of r.v.’s with limit random distribution function $F_\bullet$. Then for any set $A \subset \Omega$ with $P(A) > 0$ we have

$$\mathbb{E}_A \left( \int_{-\infty}^{+\infty} x^2 dF_\bullet(x) \right) = \int_{-\infty}^{+\infty} x^2 dF_A(x)$$

(2.6)
in the sense that if one side is finite then the other side is also finite and the two
sides are equal. The statement remains valid if in (2.6) we replace the intervals of
integration by \((-t, t)\), provided \(t\) and \(-t\) are continuity points of \(F_{\Omega}\).

We used here the notation \(F_{\bullet}\) to distinguish it from the ordinary limit distribution
function of \((X_n)\).

Proof. Assume that \(t\) and \(-t\) are continuity points of \(F_{\Omega}\). As observed in [5, p.
482], \(t\) and \(-t\) are continuity points of \(F_{\bullet}\) with probability 1 (and hence also for \(F_{A}\)
for any \(A \subset \Omega\) with \(P(A) > 0\)) and thus almost surely

\[
\int_{|x| < t} x^2 dF_{\bullet}(x) = -[x^2 (1 - F_{\bullet}(x) + F_{\bullet}(-x))]_t^t + \int_{|x| < t} (1 - F_{\bullet}(x) + F_{\bullet}(-x)) 2x dx
\]
as it is seen by splitting the integral on the left hand side into subintegrals over
\((-t, 0)\) and \((0, t)\) (the integral over \(\{0\}\) clearly equals 0) and using integration by
parts. The same formula holds with \(F_{\bullet}\) replaced by \(F_{A}\). Integrating the last relation
over \(A \subset \Omega\) and using (1.1) and Fubini’s theorem, we get the validity of (2.6) over
\((-t, t)\). Letting \(t \to \infty\) we get (2.6) over \((-\infty, \infty)\).

For the following lemma (which is the key tool for the proof of the sufficiency
part of Theorem 1.1) we need some definitions. Given probability measures \(\nu_n, \nu\) on
the Borel sets of a separable metric space \((S, d)\) we say that \(\nu_n \overset{D}{\to} \nu\) if

\[
\int_S f(x)\nu_n(dx) \to \int_S f(x)\nu(dx) \text{ as } n \to \infty
\] (2.7)

for every bounded, real valued continuous function \(f\) on \(S\). (For equivalent definitions
and properties of this convergence see [4].) (2.7) is clearly equivalent to

\[
Ef(Z_n) \to Ef(Z)
\] (2.8)

where \(Z_n, Z\) are r.v.’s valued in \((S, d)\) (i.e., measurable maps from some probability
space to \((S, d)\)) with distribution \(\nu_n, \nu\). A class \(G\) of real valued functions on \(S\) is
called locally equicontinuous if for every \(\varepsilon > 0\) and \(x \in S\) there is a \(\delta = \delta(\varepsilon, x) > 0\)
such that \(y \in S, d(x, y) \leq \delta\) imply \(|f(x) - f(y)| \leq \varepsilon\) for every \(f \in G\).

**Lemma 2.4** (Ranga Rao [14]) Let \((S, d)\) be a separable metric space and \(\nu, \nu_n\) \((n =
1, 2, \ldots)\) probability measures on the Borel sets of \((S, d)\) such that \(\nu_n \overset{D}{\to} \nu\). Let \(G\) be
a class of real valued functions on \((S, d)\) such that

(a) \(G\) is locally equicontinuous

(b) There exists a continuous function \(g \geq 0\) on \(S\) such that \(|f(x)| \leq g(x)\) for all
\(f \in G\) and \(x \in S\) and

\[
\int_S g(x)\nu_n(dx) \to \int_S g(x)\nu(dx) (< \infty) \text{ as } n \to \infty.
\] (2.9)

Then

\[
\int_S f(x)\nu_n(dx) \to \int_S f(x)\nu(dx) \text{ as } n \to \infty
\] (2.10)

uniformly in \(f \in G\).
3 Proof of Theorem 1.1

Let \((\Omega, \mathcal{F}, P)\) be the probability space of the \(X_n\)'s and \(X = (X_1, X_2, \ldots)\); let further \(\mu\) be the limit random measure of \((X_n)\). Let \((Y_n)\) be a sequence of r.v.'s on \((\Omega, \mathcal{F}, P)\) such that, given \(X\) and \(\mu\), the r.v.'s \(Y_1, Y_2, \ldots\) are conditionally i.i.d. with distribution \(\mu\), i.e.,

\[
P(Y_1 \in A_1, \ldots, Y_k \in A_k | X, \mu) = \prod_{i=1}^{k} P(Y_i \in A_i | X, \mu) \quad \text{a.s.} \quad (3.1)
\]

\[
P(Y_j \in A | X, \mu) = \mu(A) \quad \text{a.s.} \quad (3.2)
\]

for any \(j, k\) and Borel sets \(A, A_1, \ldots, A_k\) on the real line. Such a sequence \((Y_n)\) always exists after a suitable enlargement of the probability space (in fact \((Y_n)\) exists on \((\Omega, \mathcal{F}, P)\) if \((\Omega, \mathcal{F}, P)\) is atomless over \(\sigma(X, \mu)\), see the vector-valued version of Theorem (1.5) of \([5]\); see also the remark preceding Theorem (1.3) in \([5, p. 479]\)) or, alternatively, the sequence \((X_n)\) can be redefined, without changing its distribution, on a standard sequence space over which \((Y_n)\) can be defined, see \([1, p. 72]\). Clearly, \((Y_n)\) is an exchangeable sequence; we call it the limit exchangeable sequence of \((X_n)\).

It is not hard to see (cf. \([1, 5]\)) that there exists a subsequence \((X_{n_k})\) such that for every \(k \geq 1\) we have

\[
(X_{n_{j_1}}, \ldots, X_{n_{j_k}}) \xrightarrow{D} (Y_1, \ldots, Y_k) \quad \text{if} \quad j_1 < \cdots < j_k \quad \text{and} \quad j_1 \to \infty. \quad (3.3)
\]

Note that the existence of a subsequence \((X_{n_k})\) and exchangeable \((Y_k)\) satisfying \(3.3\) was first proved by Dacunha-Castelle and Krivine \([6]\) via ultrafilter techniques. The limit exchangeable sequence, as defined above, also has the following simple property, proved in \([1, \text{Lemma 12}]\).

**Lemma 3.1** For every \(\sigma(X)-\text{measurable}\) r.v. \(Z\) and any \(j \geq 1\) we have

\[
(X_n, Z) \xrightarrow{D} (Y_j, Z)
\]

As before, let \(\mathcal{M}\) denote the set of all probability measures on \(\mathbb{R}\) and let \(\pi\) be the Prohorov metric on \(\mathcal{M}\) defined by

\[
\pi(\nu, \lambda) = \inf \{ \varepsilon > 0 : \nu(A) \leq \lambda(A^\varepsilon) + \varepsilon \quad \text{and} \quad \lambda(A) \leq \nu(A^\varepsilon) + \varepsilon \quad \text{for all Borel sets} \quad A \subset \mathbb{R} \},
\]

Here

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

denotes the open \(\varepsilon\)-neighborhood of \(A\). Let

\[
S = \left\{ \nu \in \mathcal{M} : \int x d\nu(x) = 0, \quad \int x^2 d\nu(x) < +\infty \right\}. \quad (3.4)
\]

Since \(\int_{-\infty}^{\infty} x^2 d\mu(x) < \infty\) a.s. (which follows from \((1.2)\)) and \(\int_{-\infty}^{\infty} x d\mu(x) = 0\) a.s. by \(X_n \to 0\) weakly, we have

\[
P\{ \mu \in S \} = 1. \quad (3.5)
\]
Following Aldous \cite{1} we define another metric $d$ on $S$ by

$$d(\nu, \lambda) = \left( \int_0^1 (F^{-1}_\nu(x) - F^{-1}_\lambda(x))^2 \, dx \right)^{1/2} \tag{3.6}$$

where $F_\nu$ and $F_\lambda$ are the distribution functions of $\nu$ and $\lambda$, respectively, and $F^{-1}$ is defined by

$$F^{-1}(x) = \inf \{ t : F(t) \geq x \}, \quad 0 < x < 1$$

for any distribution function $F$. The right side of (3.6) equals $\| F^{-1}_\nu(\eta) - F^{-1}_\lambda(\eta) \|_2$, where $\eta$ is a random variable uniformly distributed in $(0, 1)$. Since $F^{-1}_\nu(\eta)$ and $F^{-1}_\lambda(\eta)$ are r.v.’s with distribution $\nu$ and $\lambda$, respectively (and thus square integrable), it follows that $d$ is a metric on $S$. It is easily seen (cf. \cite{1} p. 80 and relation (5.15) in \cite{1} p. 74) that $d$ is separable and generates the same Borel $\sigma$-field as $\pi$. By the definition of $d$ we have, letting 0 denote the zero distribution,

$$E d(\mu, 0)^p = E(\text{Var}(\mu))^{p/2} = E \left( \int_{-\infty}^{\infty} x^2 d\mu(x) \right)^{p/2} < \infty \quad \tag{3.7}$$

by our assumption (12). The following lemma expresses the crucial equicontinuity property of $d$.

**Lemma 3.2** Let

$$\psi(a_1, \ldots, a_n) = \left\| \sum_{i=1}^n a_i Y_i \right\|_p. \tag{3.8}$$

Then we have

$$\left\| t + \sum_{k=1}^n a_k \xi_k^{(\nu)} \right\|_p - \left\| t + \sum_{k=1}^n a_k \xi_k^{(\lambda)} \right\|_p \leq K d(\nu, \lambda) \psi(a_1, \ldots, a_n) \tag{3.9}$$

for some constant $K > 0$, every $n \geq 1$, $\nu, \lambda \in S$, real numbers $t, a_1, \ldots, a_n$ and i.i.d. sequences $(\xi_k^{(\nu)}), (\xi_k^{(\lambda)})$ with respective distributions $\nu$ and $\lambda$.

Relation (3.9) means that the class of functions $\{f_{t,a,n}\}$ defined by

$$f_{t,a,n}(\nu) = \psi(a)^{-1} \left\| t + \sum_{k=1}^n a_k \xi_k^{(\nu)} \right\|_p, \quad a = (a_1, \ldots, a_n) \neq 0 \tag{3.10}$$

(where the variable is $\nu$ and $t, a, n$ are parameters) is equicontinuous. In the context of unconditional convergence of lacunary series, the importance of such equicontinuity conditions was discovered by Aldous \cite{1}. A similar condition in terms of the compactness of the 1-conic class belonging to the type determined by $(X_n)$ was given by Krivine and Maurey (see Proposition 3 in Guerre \cite{9}). The proof of our results is, however, purely probabilistic and we will not use types.

**Proof of Lemma** We start with recalling the well known fact that if $(\xi_n)$ is an i.i.d. sequence with $E \xi_n = 0$, $E\xi_n^2 < +\infty$ then

$$C \| \xi \|_1 \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \leq \left\| \sum_{i=1}^k a_i \xi_i \right\|_p \leq \| \xi \|_2 \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \tag{3.11}$$

7
for any $1 \leq p < 2$ and any $(a_1, \ldots, a_n) \in \mathbb{R}^n$, where $C > 0$ is an absolute constant.

Since the $L^p$ norm of $\sum_{i=1}^k a_i \xi_i$ in (3.11) cannot exceed the $L^2$ norm, the upper bound in (3.11) is obvious, while the lower bound is classical, see [13]. Since $\mathbb{E} |\sum_{i=1}^n a_i Y_i|^p$ can be obtained by integrating $\mathbb{E} \left| \sum_{i=1}^n a_i \xi_i^{(\omega)} \right|^p$ over $\Omega$ with respect to $dP(\omega)$ where for each $\omega \in \Omega$ the $\xi_i^{(\omega)}$ are i.i.d. with distribution $\mu(\omega)$, relation (3.11) implies that

$$A \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \leq \left\| \sum_{i=1}^k a_i Y_i \right\|_p \leq B \left( \sum_{i=1}^k a_i^2 \right)^{1/2}$$

(3.12)

where

$$A = C \left[ \mathbb{E} \left( \int_{-\infty}^{\infty} |x| d\mu(x) \right)^p \right]^{1/p}, \quad B = \left[ \mathbb{E} \left( \int_{-\infty}^{\infty} x^2 d\mu(x) \right)^{p/2} \right]^{1/p}.$$

By (1.2) and since the assumptions of Theorem 1.1 imply that $\mu$ is not concentrated at zero a.s., we have $0 < A \leq B < \infty$.

Turning to the proof of (3.11), note that the $L_p$ norms on the left hand side depend on the choice of these i.i.d. sequences and thus it suffices to verify (3.11) for any specific construction. Let $(\eta_n)$ be a sequence of independent r.v.’s, uniformly distributed over $(0, 1)$. Then $\xi^{(\nu)}_n = F^{-1}_\nu(\eta_n)$ and $\xi^{(\lambda)}_n = F^{-1}_\lambda(\eta_n)$ are i.i.d. sequences with distribution $\nu$ and $\lambda$, respectively. Using these sequences in (3.11), the left hand side is at most $\| \sum_{i=1}^n a_i (\xi^{(\nu)}_i - \xi^{(\lambda)}_i) \|_p$ and since $\xi^{(\nu)}_i - \xi^{(\lambda)}_i = F^{-1}_\nu(\eta_i) - F^{-1}_\lambda(\eta_i)$ is also an i.i.d. sequence with mean 0 and variance $d(\nu, \lambda)^2$, using (3.11) and the first relation of (3.12) we get that the left hand side of (3.11) is at most $K d(\nu, \lambda) \psi(a_1, \ldots, a_n)$ with some constant $K > 0$. This completes the proof of Lemma 3.2.

With the equicontinuity statement of Lemma 3.2 at hand, we can prove the sufficiency part of Theorem 1.1 with a selection procedure similar to [2]. Assume that $(X_n)$ satisfies (1.2) and fix $0 < \varepsilon \leq 1/2$. We shall construct a sequence $n_1 < n_2 < \cdots$ of integers such that

$$(1 - \varepsilon) \psi(a_1, \ldots, a_k) \leq \left\| \sum_{i=1}^k a_i X_{n_i} \right\|_p \leq (1 + \varepsilon) \psi(a_1, \ldots, a_k)$$

(3.13)

for every $k \geq 1$ and $(a_1, \ldots, a_k) \in \mathbb{R}^k$. In view of (3.12), this will imply that $(X_{n_k})$ is equivalent to the unit vector base of $\ell^2$, but it actually shows more, namely that under the assumptions of Theorem 1.1 there is a subsequence $(1 + \varepsilon)$-equivalent to the limit exchangeable sequence and hence $(1 + \varepsilon)$-symmetric.

To construct $n_1$ we set

$$Q(a, n, \ell) = |a_1 X_n + a_2 Y_2 + \cdots + a_\ell Y_\ell|^p$$
$$R(a, \ell) = |a_1 Y_1 + a_2 Y_2 + \cdots + a_\ell Y_\ell|^p$$

8
for every $n \geq 1$, $\ell \geq 2$ and $a = (a_1, \ldots, a_\ell) \in \mathbb{R}^\ell$. We show that
\[
E \left\{ \frac{Q(a, n, \ell)}{\psi(a)^p} \right\} \to E \left\{ \frac{R(a, \ell)}{\psi(a)^p} \right\} \quad \text{as } n \to \infty \quad \text{uniformly in } a, \ell. \quad (3.14)
\]
(The right side of (3.14) equals 1.) To this end we recall that, given $X$ and $\mu$, the r.v.'s $Y_1, Y_2, \ldots$ are conditionally i.i.d. with common conditional distribution $\mu$ and thus, given $X, \mu$ and $Y_1$, the r.v.'s $Y_2, Y_3, \ldots$ are conditionally i.i.d. with distribution $\mu$. Thus
\[
E\left(Q(a, n, \ell)|X, \mu\right) = g^{a, \ell}(X_n, \mu) \quad (3.15)
\]
and
\[
E\left(R(a, \ell)|X, \mu, Y_1\right) = g^{a, \ell}(Y_1, \mu) \quad (3.16)
\]
where $g^{a, \ell}(t, \nu) = E\left|a_1 t + \sum_{i=2}^{\ell} a_i \xi_i^{(\nu)}\right|^p \quad (t \in \mathbb{R}^1, \nu \in S)$
and $(\xi_n^{(\nu)})$ is an i.i.d. sequence with distribution $\nu$. Integrating (3.15) and (3.16) we get
\[
E\left(Q(a, n, \ell)\right) = Eg^{a, \ell}(X_n, \mu) \quad (3.17)
\]
and
\[
E\left(R(a, \ell)\right) = Eg^{a, \ell}(Y_1, \mu) \quad (3.18)
\]
and thus (3.14) is equivalent to
\[
E\frac{g^{a, \ell}(X_n, \mu)}{\psi(a)^p} \to E\frac{g^{a, \ell}(Y_1, \mu)}{\psi(a)^p} \quad \text{as } n \to \infty , \quad \text{uniformly in } a, \ell. \quad (3.19)
\]
We shall derive (3.19) from Lemma 2.4 and Lemma 3.1. As we have seen above, there exists a separable metric $d$ on $S$, generating the same $\sigma$-field as the Prohorov metric $\pi$, such that (3.9) holds. But then the limit random measure $\mu$, which is a random variable taking values in $(S, \pi)$ (i.e., a measurable map from the underlying probability space to $(S, B_\pi)$ where $B_\pi$ denotes the Borel $\sigma$-field in $S$ generated by $\pi$) can be also regarded as a random variable taking values in $(S, d)$. Also, $\mu$ is clearly $\sigma(X)$ measurable and thus $(X_n, \mu) \stackrel{D}{\to} (Y_1, \mu)$ by Lemma 3.1. Hence, (3.19) will follow from Lemma 2.4 (note the equivalence of (2.7) and (2.8)) if we show that the class of functions
\[
\left\{ \frac{g^{a, \ell}(t, \nu)}{\psi(a)^p} \right\} \quad (3.20)
\]
defined on the product metric space $(\mathbb{R}^1 \times S, \lambda^1 \times d)$ ($\lambda^1$ denotes the ordinary distance on $\mathbb{R}^1$) satisfies conditions (a),(b) of Lemma 2.4. Observe now that
\[
\psi(a_1, \ldots, a_n) \geq \psi(a_1^*, \ldots, a_n^*) \quad (3.21)
\]
where $a_i^*$ equals either $a_i$ or 0. (In case $(Y_n)$ is an i.i.d. sequence with mean 0, (3.21) follows from Jensen’s inequality (see e.g. [8, p. 153]) and the fact that, for any $H \subset \{1, 2, \ldots, n\}$, the conditional expectation of $\sum_{i=1}^{n} a_i Y_i$ given $\sigma\{Y_j, \ j \in H\}$ is
\[ \sum_{i \in I} a_iY_i. \] Since \((Y_n)\) is a mixture of i.i.d. sequences with mean 0, \((3.21)\) holds in general. In particular,

\[ \psi(a_1, \ldots, a_n) \geq \psi(0, a_2, \ldots, a_n) \quad (3.22) \]

and

\[ \psi(a_1, \ldots, a_n) \geq \psi(a_1, 0, \ldots, 0) = \text{const} \cdot |a_1| \quad (3.23) \]

and thus using \((3.9)\) we get for any \(\nu \in S, t \in \mathbb{R}^1\) and \(a = (a_1, \ldots, a_\ell) \in \mathbb{R}^\ell,\)

\[
\left\| a_1t + \sum_{i=2}^\ell a_i\xi_i^{(\nu)} \right\|_p \leq \left\| a_1t \right\|_p + \left\| \sum_{i=2}^\ell a_i\xi_i^{(\nu)} \right\|_p \leq |a_1| + \psi(a_2, \ldots, a_\ell)d(\nu, 0) \leq \text{const} \cdot \psi(a) \cdot |t| + \psi(a)d(\nu, 0) \leq \text{const} \psi(a)(|t| + d(\nu, 0)).
\]

Hence using \((3.9), (3.22), (3.24)\) and the inequality \(|x^p - y^p| \leq |x - y| \cdot p \cdot (x^{p-1} + y^{p-1})\) \((x > 0, y > 0)\) we get for any real \(t, t'\) and \(\nu, \nu' \in S\)

\[
|g^{a,t}(t, \nu) - g^{a,t'}(t', \nu')| = \left| \left\| a_1t + \sum_{i=2}^\ell a_i\xi_i^{(\nu)} \right\|_p - \left\| a_1t' + \sum_{i=2}^\ell a_i\xi_i^{(\nu')} \right\|_p \right| \\
\leq \left| \left\| a_1t \right\|_p + \sum_{i=2}^\ell a_i\xi_i^{(\nu)} \right\|_p - \left\| a_1t' + \sum_{i=2}^\ell a_i\xi_i^{(\nu')} \right\|_p \right| \cdot \text{const} \psi(a)^p(\left| t \right| + \left| t' \right| + d(\nu, 0) + d(\nu', 0))^{-1} \\
\leq \text{const} \psi(a)^p(\left| t - t' \right| + d(\nu, 0) + d(\nu', 0))^{-1}.
\]

Given \(t, \nu\) and \(\varepsilon > 0\), there exists a \(\delta = \delta(t, \nu, \varepsilon) > 0\) such that the last expression is \(\leq \varepsilon \psi(a)^p\) provided \(|t - t'| + d(\nu, \nu') \leq \delta\) and thus the class \((3.20)\) is locally equicontinuous on the product metric space \((\mathbb{R}^1 \times S, \lambda^1 \times d)\). On the other hand, \((3.24)\) shows that the function in \((3.20)\) is bounded by \(\text{const} \left| t \right| + d(\nu, 0)^p \leq \text{const} \left| t \right|^p + d(\nu, 0)^p\). Now, using \((X_n, \mu) \overset{D}{\longrightarrow} (Y_1, \mu)\), the uniform integrability of \(|X_n|^p\) and \(Ed(\mu, 0)^p < +\infty\) (see \((3.7)\)) we get

\[
E(|X_n|^p + d(\mu, 0)^p) \longrightarrow E(|Y_1|^p + d(\mu, 0)^p).
\]

Thus the class \((3.20)\) satisfies also condition (b) of Lemma 2.4. We thus proved relation \((3.19)\) and thus also \((3.14)\) whence it follows (note again that the right side of \((3.14)\) equals 1) that

\[
\psi(a)^{-1} \left\| a_1X_1 + a_2Y_2 + \cdots + a_\ell Y_\ell \right\|_p \\
\longrightarrow \psi(a)^{-1} \left\| a_1Y_1 + a_2Y_2 + \cdots + a_\ell Y_\ell \right\|_p \quad \text{as} \quad n \to \infty
\]

uniformly in \(\ell, a\). Hence we can choose \(n_1\) so large that

\[
\left\| a_1X_{n_1} + a_2Y_2 + \cdots + a_\ell Y_\ell \right\|_p - \left\| a_1Y_1 + a_2Y_2 + \cdots + a_\ell Y_\ell \right\|_p \leq \frac{\varepsilon}{2} \psi(a_1, \ldots, a_\ell)
\]
for every $\ell, a$. This completes the first induction step.

Assume now that $n_1, \ldots, n_{k-1}$ have already been chosen. Exactly in the same way as we proved (3.25), it follows that for $\ell > k$
\[
\psi(a)^{-1} \|a_1X_{n_1} + \cdots + a_kX_{n_{k-1}} + a_kX_n + a_{k+1}Y_{k+1} + \cdots + a_\ell Y_\ell\|_p
\]
\[
\rightarrow \psi(a)^{-1} \|a_1X_{n_1} + \cdots + a_kX_{n_{k-1}} + a_kY_k + \cdots + a_\ell Y_\ell\|_p
\]
as $n \to \infty$ uniformly in $a$ and $\ell$. Hence we can choose $n_k$ so that $n_k > n_{k-1}$ and
\[
\|a_1X_{n_1} + \cdots + a_kX_{n_{k-1}} + a_kY_k + \cdots + a_\ell Y_\ell\|_p
\]
\[
- \|a_1X_{n_1} + \cdots + a_kX_{n_{k-1}} + a_kY_k + \cdots + a_\ell Y_\ell\|_p \leq \frac{\varepsilon}{2} |\psi(a_1, \ldots, a_\ell)|
\]
for every $(a_1, \ldots, a_\ell) \in \mathbb{R}^\ell$ and $\ell > k$. This completes the $k$-th induction step; the so constructed sequence $(n_k)$ obviously satisfies
\[
\|a_1X_{n_1} + \cdots + a_\ell Y_\ell\|_p - \|a_1Y_1 + \cdots + a_\ell Y_\ell\|_p \leq \varepsilon |\psi(a_1, \ldots, a_\ell)|
\]
for every $\ell \geq 1$ and $(a_1, \ldots, a_\ell) \in \mathbb{R}^\ell$. The last relation is equivalent to (3.13) and thus the sufficiency of (1.2) in Theorem 1.1 is proved.

We now turn to the proof of necessity of (1.2) in Theorem 1.1. Assume that $(X_n)$ is equivalent to the unit vector basis of $\ell^2$; then for any increasing sequence $(m_k)$ of integers we have
\[
\left\| \frac{1}{\sqrt{N}} \sum_{k=1}^{N} X_{m_k} \right\|_p = O(1)
\]
and thus by the Markov inequality we have for any $A \subset \Omega$ with $P(A) > 0$,
\[
P_A \left\{ \left\| \frac{1}{\sqrt{N}} \sum_{k=1}^{N} X_{m_k} \right\|_p \geq T \right\} \leq 1/2 \quad \text{for } T \geq T_0, \ N \geq 1 \quad (3.26)
\]
where $T_0$ depends on $A$ and the sequence $(X_n)$. We show first that
\[
\int_{-\infty}^{\infty} x^2 d\mu(x) < \infty \quad \text{a.s.} \quad (3.27)
\]
Let $F_\mu(x)$ denote the random distribution function corresponding to $\mu$ and assume indirectly that there exists a set $B \subset \Omega$ with $P(B) > 0$ such that
\[
\lim_{t \to \infty} \int_{|x| < t} x^2 dF_\mu(x) = +\infty \quad \text{on } B. \quad (3.28)
\]
By Egorov’s theorem there exists a set $B^* \subset B$ with $P(B^*) \geq P(B)/2$ such that on $B^*$ (3.28) holds uniformly, i.e. there exists a positive, nondecreasing, nonrandom function $K_t \to +\infty$ such that
\[
\int_{|x| < t} x^2 dF_\mu(x) \geq K_t \quad \text{on } B^*. \quad (3.29)
\]
Also,
\[ \int_{|x| \geq t} dF_{\bullet}(x) \rightarrow 0 \quad \text{a.s. as } t \to \infty \]  
(3.30)
and thus we can choose a set \( B^{**} \subset B^* \) with \( P(B^{**}) \geq P(B^*)/2 \) such that on \( B^{**} \) relation (3.30) holds uniformly, i.e. there exists a positive, nonincreasing, nonrandom function \( \tilde{\varepsilon}_t \to 0 \) such that
\[ \int_{|x| \geq t} dF \cdot (x) \leq \tilde{\varepsilon}_t \quad \text{on } B^{**}. \]  
(3.31)

We show that there exists a subsequence \((X_{m_k})\) of \((X_n)\) such that (3.26) fails for \( A = B^{**} \). Since our argument will involve the sequence \((X_n)\) only on the set \( B^{**} \) and on \( B^{**} \) \((X_n)\) satisfies the assumptions of Theorem 1.1 with the same \( \mu \) and with \( \|X_n\|_p = 1 \) replaced by \( \|X_n\|_p = O(1) \) (which is all we need for the rest of the proof), in the sequel we can assume, without loss of generality, that \( B^{**} = \Omega \). That is, we may assume that (3.29), (3.31) hold on the whole probability space.

Let \( C \) be an arbitrary set in the probability space with \( P(C) > 0 \). Integrating (3.29), (3.31) on \( C \) and using (1.1) and Lemma 2.3 we get
\[
\int_{|x| < t} x^2 dF_C(x) \geq K_t, \quad \int_{|x| \geq t} dF(x) \leq \tilde{\varepsilon}_t \quad \text{for } t, -t \in H
\]  
(3.32)
where \( H \) denotes the set of continuity points of \( F_C \). Since the integrals in (3.32) are monotone functions of \( t \) and \( \mathbb{R} \setminus H \) is countable, (3.32) remains valid with \( K_{t/2} \) resp. \( \tilde{\varepsilon}_{t/2} \) if we drop the assumption \( t, -t \in H \). Thus, keeping the original notation, in the sequel we can assume that (3.32) holds for all \( t > 0 \). Choose now \( t_0 \) so large that \( \tilde{\varepsilon}_{t_0} \leq 1/16 \) and then choose \( t_1 > t_0 \) so large that
\[ K_t^{1/2} \geq 4t_0 \quad \text{for } t \geq t_1. \]

Then for \( t \geq t_1 \) we have, using (3.32),
\[
\left| \int_{|x| < t} x dF_C(x) \right| \leq t_0 + \int_{t_0 \leq |x| < t} |x| dF_C(x) \\
\leq t_0 + \left( \int_{|x| \geq t_0} dF_C(x) \right)^{1/2} \left( \int_{|x| < t} x^2 dF_C(x) \right)^{1/2} \\
\leq t_0 + \frac{1}{4} \left( \int_{|x| < t} x^2 dF_C(x) \right)^{1/2} \leq \frac{1}{2} \left( \int_{|x| < t} x^2 dF_C(x) \right)^{1/2}
\]
and thus for any \( C \subset \Omega \) with \( P(C) > 0 \) we have
\[
\int_{|x| < t} x^2 dF_C(x) - 2 \left( \int_{|x| < t} x dF_C(x) \right)^2 \geq \frac{1}{2} K_t, \quad t \geq t_1. \]  
(3.33)
Let now \((\varepsilon_n)\) tend to 0 so rapidly that
\[
\sum_{j=a_k+1}^\infty \varepsilon_j \leq 2^{-k}. \tag{3.34}
\]
Let \(a_k = \lfloor \log k + 1 \rfloor \) \((k = 1, 2, \ldots)\). By Lemma 2.1 there exists a subsequence \((X_{m_k})\) and a sequence \((Y_k)\) of discrete r.v.’s such that (2.1) holds and for each \(k \geq 1\) the atoms of the finite \(\sigma\)-field \(\sigma\{Y_1, \ldots, Y_{a_k}\}\) can be divided into two classes \(\Gamma_1\) and \(\Gamma_2\) such that
\[
\sum_{B \in \Gamma_1} P(B) \leq \varepsilon_{a_k+1} \leq 2^{-k} \tag{3.35}
\]
and for each \(B \in \Gamma_2\) there exist \(P_B\)-independent r.v.'s \(Z_{a_k+1}^{(B)}, \ldots, Z_k^{(B)}\) defined on \(B\) with common distribution \(F_B\) such that
\[
P_B(|Y_j - Z_j^{(B)}| \geq 2^{-k}) \leq 2^{-k} \quad (j = a_k + 1, \ldots, k). \tag{3.36}
\]
Set
\[
S_{a_k,k}^{(B)} = \sum_{j=a_k+1}^k Z_j^{(B)}, \quad B \in \Gamma_2
\]
\[
\mathcal{S}_{a_k,k} = \sum_{B \in \Gamma_2} S_{a_k,k}^{(B)} I(B),
\]
where \(I(B)\) denotes the indicator function of \(B\). By (3.36) and \(k2^{-k} \leq 1\),
\[
P_B \left( \left| \sum_{j=a_k+1}^k Y_j - \sum_{j=a_k+1}^k Z_j^{(B)} \right| \geq 1 \right) \leq k2^{-k}, \quad B \in \Gamma_2
\]
and thus using (3.35) we get
\[
P \left( \left| \sum_{j=a_k+1}^k Y_j - \mathcal{S}_{a_k,k} \right| \geq 1 \right) \leq (k+1)2^{-k}. \tag{3.37}
\]
Since \(\|X_n\|_1 = O(1)\), by the triangular inequality and the Markov inequality we have
\[
P \left( \left| \sum_{j=1}^{a_k} X_{m_j} \right| \geq a_k k^{1/4} \right) \leq \sum_{j=1}^{a_k} P(|X_{m_j}| \geq k^{1/4}) \leq \text{const} (\log k + 1)k^{-1/4} =: \delta_k
\]
which, together with (3.37), (2.1) and (3.34), yields
\[
P \left( \left| \sum_{j=1}^k X_{m_j} - \mathcal{S}_{a_k,k} \right| \geq 3a_k k^{1/4} \right) \leq P \left( \left| \sum_{j=1}^{a_k} X_{m_j} \right| \geq a_k k^{1/4} \right) + P \left( \left| \sum_{j=a_k+1}^k X_{m_j} \right| \geq a_k k^{1/4} \right) + P \left( \left| \sum_{j=a_k+1}^k |X_{m_j} - Y_j| \right| \geq 1 \right) + P \left( \left| \sum_{j=a_k+1}^k Y_j - \mathcal{S}_{a_k,k} \right| \geq 1 \right). \]
\[ \leq \delta_k + (k + 2)2^{-k}. \] (3.38)

Applying Lemma 2.2 to the i.i.d. sequence \( \{Z_j^{(B)}, a_k + 1 \leq j \leq k\} \) and using (3.33) with \( C = B, a_k \leq k/2 \) and the monotonicity of \( K_t \) we get for any \( T \geq 2, \)

\[ P_B \left( \left| \frac{S_{a_k,k}^{(B)}}{\sqrt{k}} \right| \leq T \right) \leq P_B \left( \left| \frac{S_{a_k,k}^{(B)}}{\sqrt{k}} \right| \leq 2T \right) \leq \text{const} \cdot 2TK^{-1/2} \]

where the constants are absolute. Thus using (3.33) it follows that

\[ P \left( \left| \frac{S_{a_k,k}}{\sqrt{k}} \right| \leq T \right) \leq \text{const} \cdot TK^{-1/2} + 2^{-k}. \] (3.39)

Using (3.38), (3.39) and \( a_k \leq \log k + 1 \) it follows that

\[ P \left( \left| \frac{1}{\sqrt{k}} \sum_{j=1}^{k} X_{n_j} \right| \leq T \right) \leq P \left( \frac{S_{a_k,k}}{\sqrt{k}} \leq T + 3a_kk^{-1/4} \right) + (k + 2)2^{-k} + \delta_k \]

\[ \leq \text{const} \left( T + 3a_kk^{-1/4} \right) K^{-1/2} + (k + 2)2^{-k} + \delta_k \rightarrow 0 \quad \text{as } k \rightarrow \infty \]

for any fixed \( T \geq 2 \) which clearly contradicts to (3.20) with \( A = \Omega \). This completes the proof of (3.27).

Since \( X_n \rightarrow 0 \) weakly in \( L^p \), we have \( \int_{-\infty}^{\infty} x^2 d\mu(x) = 0 \) a.s., and thus the random measure \( \mu \) has mean zero and finite variance with probability one. Thus the subsequence principle, specialized to the central limit theorem (see e.g. [3, Theorem 3] and the remark following it) there exists a subsequence \( (X_{n_k}) \) such that

\[ N^{-1/2} \sum_{k=1}^{N} X_{n_k} \overset{D}{\rightarrow} Y \zeta \] (3.40)

where \( Y = \left( \int_{-\infty}^{\infty} x^2 d\mu(x) \right)^{1/2} \), \( \zeta \) is a standard normal variable and \( Y \) and \( \zeta \) are independent. Note that (3.40) holds in distribution, but by a well known result of Skorokhod (see e.g. [11, p. 70]) there exist r.v.’s \( W_N, W (N = 1, 2, \ldots) \) such that \( W_N \) has the same distribution as \( N^{-1/2} \sum_{k=1}^{N} X_{n_k} \) in (3.40), \( W \) has the same distribution as \( Y \zeta \) and \( W_N \rightarrow W \) a.s. Thus (3.40) and Fatou’s lemma imply

\[ \| Y \zeta \|_p \leq \liminf_{N \rightarrow \infty} \| N^{-1/2} \sum_{k=1}^{N} X_{n_k} \|_p < \infty \] (3.41)

where the second inequality follows from the equivalence of \( (X_n) \) to the unit vector basis of \( \ell^2 \), assumed at the beginning of the proof. Since \( Y \) and \( \zeta \) are independent, (3.41) implies \( E|Y|^p < \infty \), i.e. (1.2) holds, completing the proof of the converse part of Theorem 1.1.

**Acknowledgement.** The authors are indebted to the referee for his/her valuable remarks and suggestions leading to a considerable improvement of the presentation.
References

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

[2] I. Berkes, On almost symmetric sequences in $L_p$. *Acta Math. Hung.* **54** (1989) 269-278.

[3] I. Berkes and E. Péter, Exchangeable r.v.s and the subsequence principle. *Probability Theory and Rel. Fields* **73** (1986) 395–413.

[4] P. Billingsley, *Convergence of Probability Measures*. Wiley, New York 1999.

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

[6] D. Dacunha-Castelle. Indiscernability and exchangeability in $L^p$ spaces. Proc. Aarhus Seminar on random series, convex sets and geometry of Banach spaces. Aarhus Universitet, various publications **25** (1975), 50-56.

[7] C. G. Esseen. On the concentration function of a sum of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1968), 290–308.

[8] W. Feller. *An Introduction to Probability Theory and its Applications*, Vol. II, 2nd Edition, Wiley, New York 1970.

[9] S. Guerre. Types and suites symétriques dans $L^p$, $1 \leq p < +\infty$, $p \neq 2$, *Israel J. Math.* **53** (1986), 191–208.

[10] S. Guerre and Y. Raynaud. On sequences with no almost symmetric subsequence. Texas Functional Analysis Seminar 1985–1986 (Austin, TX, 19851986), 83–93, Longhorn Notes, Univ. Texas, Austin, TX, 1986.

[11] M.I. Kadec and W. Peczynski. Bases, lacunary sequences and complemented subspaces in the spaces $L_p$. *Studia Math.* **21** 1961/1962, 161–176.

[12] P. Lévy. Théorie de l’addition des variables aléatoires. Gauthier-Villars, 1937.

[13] J. Marcinkiewicz and A. Zygmund. Quelques théorèmes sur les fonctions indépendantes. *Studia Math.* **7** (1938), 104–120.

[14] R. Ranga Rao. Relations between weak and uniform convergence of measures with applications. *Ann. of Math. Statist.* **33** (1962), 659–680.

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