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

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 $1 \le p < 2$ and show that a necessary and sufficient condition for the first alternative in the Kadec-Pelczynski 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

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

for every $n \ge 1$ and every $(a_1, \ldots, a_n) \in \mathbb{R}^n$. By a classical theorem of Kadec and Pełczynski [10], 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 $\{|x_n|^p, n \ge 1\}$ 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 [10].

The purpose of the present paper is to investigate the case $1 \le p < 2$ 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 $1 \le p < 2$ and let $(X_n)$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$; assume that $\{|X_n|^p, n \ge 1\}$ is uniformly integrable and $X_n \to 0$ weakly in $L^p$. (This is meant as $\lim_{n\to\infty} \mathbb{E}(X_n Y) = 0$ 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 P24302-N18.

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

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

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 in [13].) Since $\{|X_n|^p,\, n \ge 1\}$ is uniformly integrable, the sequence $(X_n)$ is tight and thus by an extension of the Helly-Bray theorem (see e.g. [5]), every tight sequence of r.v.'s 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

$$F_A(t) = \mathbb{E}_A(\mu(-\infty, t]), \tag{1.1}$$

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 $1 \le p < 2$ and let $(X_n)$ be a determining sequence of random variables such that $\|X_n\|_p = 1$ $(n = 1, 2, \ldots)$, $\{|X_n|^p,\, n \ge 1\}$ 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 $\ell^2$ if and only if*

$$\int_{-\infty}^{\infty} x^2 d\mu(x) \in L^{p/2}. \tag{1.2}$$

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 $P(A) \ge 1-\delta$ and a subsequence $(X_{n_k})$ such that

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

Thus the first alternative in the Kadec-Pełczynski 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 $k \ge 1$, any indices $j_1 > j_2 > \ldots j_k \ge K$, 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

$$(1-\varepsilon)\| \sum_{i=1}^k a_i X_{j_i} \|_p \le \| \sum_{i=1}^k a_i X_{\sigma(j_i)} \|_p \le (1+\varepsilon) \| \sum_{i=1}^k a_i X_{j_i} \|_p \,.$$

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 [8] implies the existence of a further subsequence $(X_{m_k})$ of $(X_{n_k})$ which is $(1+\varepsilon)$-equivalent to the to the unit vector basis of $\ell^2$. This implies immediately that $(X_n)$ has an almost symmetric subsequence. Note that this conclusion also follows from the proof of Theorem 1.1. Guerre and Raynaud [9] also showed that for any $1 \le p < q < 2$ 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

Let $(X_n)$ be a determining sequence of random variables on $(\Omega, \mathcal{F}, P)$ with limit random measure $\mu$. By a standard construction (see e.g. [1], p. 72) there exists, after suitably enlarging the probability space, a sequence $(Y_n)$ of random variables such that conditionally on $\mu$, the variables $Y_1, Y_2, \dots$ are independent with conditional distribution $\mu$. 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 $k \ge 1$ we have

$$(X_{n_{j_1}}, \dots, X_{n_{j_k}}) \xrightarrow{d} (Y_1, \dots, Y_k) \;\text{ if }\; j_1 < \dots < j_k \;,\; j_1 \to \infty. \tag{2.1}$$

Note that the existence of a subsequence $(X_{n_k})$ and exchangeable $(Y_k)$ satisfying (2.1) was first proved by Dacunha-Castelle and Krivine [6] via ultrafilter techniques. Relation (2.1) shows that the behavior of thin subsequences of $(X_n)$ resembles that of the exchangeable sequence $(Y_k)$. Aldous [1] proved that limit theorems valid for $(Y_k)$ remain valid for sufficiently thin $(X_{n_k})$, a profound result verifying the so called *subsequence principle*. As the results of the present paper show, the limit random measure is also a useful concept in studying the existence of Hilbertian subsequences of $(X_n)$.

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*

$$P\big(|X_{m_k} - Y_k| \ge \varepsilon_k\big) \le \varepsilon_k \quad k = 1, 2 \dots \tag{2.2}$$

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

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

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

$$P_B\big(|Y_j - Z_j^{(B)}| \ge \varepsilon_k\big) \le \varepsilon_k \quad j = k, k+1, \dots \tag{2.3}$$

*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 (2.2), 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, (2.2) 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 sets of a suitable partition of the probability space. This fact is an "effective" form of the fundamental subsequence principle of Aldous [1] and reduces the studied problem to the i.i.d. case which will be handled by the classical concentration technique of Lévy [11].

**Lemma 2.2** *Let $X_1, X_2, \dots, X_n$ 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*

$$P\left(\left|S_n\right| \le t\right) \le A \frac{t}{\sqrt{n}} \left[ \int\limits_{|x| \le t} x^2 dF(x) - 2 \left( \int\limits_{|x| \le t} x dF(x) \right)^2 \right]^{-1/2} \tag{2.4}$$

*provided the difference on the right-hand side is positive and $\int\limits_{|x|\le t} dF(x) \ge 1/2$. Here $A$ is an absolute constant.*

*Proof.* Let $F^*$ denote the distribution function obtained from $F$ by symmetrization. From a concentration function inequality of Esseen (see [7], formula (3.3)) it follows that the left-hand side of (2.4) cannot exceed

$$A_1 \frac{t}{\sqrt{n}} \left( \int\limits_{|x| \le 2t} x^2 dF^*(x) \right)^{-1/2}$$

where $A_1$ is an absolute constant. Hence to prove (2.4) it suffices to show that $\int\limits_{|x|\le t} dF(x) \ge 1/2$ implies

$$\int\limits_{|x| \le 2t} x^2 dF^*(x) \ge \int\limits_{|x| \le t} x^2 dF(x) - 2 \left( \int\limits_{|x| \le t} x dF(x) \right)^2. \tag{2.5}$$

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

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

Then

$$\int_{|x|\le 2t} x^2 dF^*(x) = \int_C (\xi-\eta)^2 dP \ge \int_D (\xi-\eta)^2 dP$$
$$= 2\int_{|\xi|\le t} \xi^2 dP \cdot P(|\eta|\le t) - 2\Big(\int_{|\xi|\le t} \xi dP\Big)^2 \ge \int_{|\xi|\le t} \xi^2 dP - 2\Big(\int_{|\xi|\le t} \xi dP\Big)^2$$

since $P(|\eta|\le t)\ge 1/2$. 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*

$$\mathbb{E}_A\Big(\int_{-\infty}^{+\infty} x^2 dF_\bullet(x)\Big) = \int_{-\infty}^{+\infty} x^2 dF_A(x) \tag{2.6}$$

*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 $F_A$.*

We use here the notation $F_\bullet$ to distinguish it from the ordinary limit distribution function $F$. Lemma 2.3 follows easily from (1.1) by integration by parts.

# 3 Proof of the theorem

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

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

Here

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

denotes the open $\varepsilon$-neighborhood of $A$. The sufficiency of the criterion (1.2) will be deduced from the following theorem proved in Berkes [2]:

**Theorem 3.1** *Let $p\ge 1$ and let $(X_n)$ be a determining sequence of r.v.'s such that $\{|X_n|^p,\ n\ge 1\}$ is uniformly integrable. Let $\mu$ and $(Y_n)$ denote the limit random measure and limit exchangeable sequence of $(X_n)$, respectively and assume that the measure $\mu$ is not concentrated at the point 0 a.s. Let*

$$\psi(a_1,\dots,a_n) = \Big\|\sum_{i=1}^n a_i Y_i\Big\|_p. \tag{3.1}$$

*Let $S$ be a Borel set in $(\mathcal{M}, \pi)$ such that*

$$P\{\mu \in S\} = 1 \tag{3.2}$$

*and assume that there exists a separable metric $d$ on $S$, generating the same Borel $\sigma$-algebra as the Prohorov metric $\pi$, such that $\mathbb{E}d(\mu, 0)^p < +\infty$ (0 denotes the zero distribution) and*

$$\left| \|t + \sum_{k=1}^n a_k \xi_k^{(\nu)}\|_p - \|t + \sum_{k=1}^n a_k \xi_k^{(\lambda)}\|_p \right| \le K d(\nu, \lambda) \psi(a_1, \dots, a_n) \tag{3.3}$$

*for some constant $K > 0$, every $n \ge 1$, $\nu, \lambda \in S$, real numbers $t, a_1, \dots, a_n$ and i.i.d. sequences $(\xi_n^{(\nu)}), (\xi_n^{(\lambda)})$ with respective distributions $\nu$ and $\lambda$. Then for any $\varepsilon > 0$ there exists a subsequence $(X_{n_k})$ which is $(1+\varepsilon)$-equivalent in $L^p$ to $(Y_k)$.*

We note that the main result of [2] states that under the conclusion of Theorem 3.1 the sequence $(X_n)$ has an almost symmetric subsequence, but the proof actually provides a subsequence $(1+\varepsilon)$-equivalent to $(Y_k)$, see relation (24) of [2]. We also note that in [2] the nondegeneracy assumption on $\mu$ is not assumed, but it is used implicitly to show that the constant $A$ in relation (14) of [2] is positive and thus $\psi(a_1, \dots, a_n) > 0$ except when $\mathbf{a} = (a_1, \dots, a_n)$ is the zero vector. Relation (3.3) means that the class of functions $\{f_{t,\mathbf{a},n}\}$ defined by

$$f_{t,\mathbf{a},n}(\nu) = \psi(\mathbf{a})^{-1} \left\| t + \sum_{k=1}^n a_k \xi_k^{(\nu)} \right\|_p \qquad \mathbf{a} = (a_1, \dots a_n) \ne \mathbf{0} \tag{3.4}$$

(where the variable is $\nu$ and $t, \mathbf{a}, n$ are parameters) is equicontinuous. In the context of unconditional convergence of lacunary series, the importance of such equicontinuity conditions in the uniform behavior of lacunary sequences was discovered by Aldous [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 [8]). The proof of our results is, however, purely probabilistic and no type theory will be used.

To prove the sufficiency of (1.2) in Theorem 1.1, we will use 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

$$C\|\xi\|_1 \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \le \left\| \sum_{i=1}^k a_i \xi_i \right\|_p \le \|\xi\|_2 \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \tag{3.5}$$

for any $1 \le p < 2$ and any $(a_1, \dots, a_n) \in \mathbb{R}^n$, where $C > 0$ is an absolute constant. Since the $L^p$ norm of $\sum_{i=1}^k a_i \xi_i$ in (3.5) cannot exceed the $L^2$ norm, the upper bound in (3.5) is obvious, while the lower bound is classical, see [12]. Since $\mathbb{E}\left|\sum_{i=1}^n a_i Y_i\right|^p$ can be obtained by integrating $\mathbb{E}\left|\sum_{i=1}^n a_i \xi_i^{(\omega)}\right|^p$ over $\Omega$ with respect to a probability measure where for each $\omega \in \Omega$ the $\xi_i^{(\omega)}$ are i.i.d. with distribution $\mu(\omega)$, relation (3.5) implies that

$$A \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \le \left\| \sum_{i=1}^k a_i Y_i \right\|_p \le B \left( \sum_{i=1}^k a_i^2 \right)^{1/2} \tag{3.6}$$

where

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

and $C$ is the constant in (3.5). By (1.2) and since $\mu$ is not concentrated at zero a.s., we have $0 < A \leq B < \infty$. Let

$$S = \left\{\nu \in \mathcal{M} : \int x d\nu(x) = 0\ ,\ \int x^2 d\nu(x) < +\infty\right\}. \tag{3.7}$$

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, (3.2) is satisfied. Following Aldous [1] we define a metric $d$ on $S$ by

$$d(\nu,\lambda) = \left(\int_0^1 \left(F_\nu^{-1}(x) - F_\lambda^{-1}(x)\right)^2 dx\right)^{1/2} \tag{3.8}$$

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

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

for any distribution function $F$. The right side of (3.8) equals $\|F_\nu^{-1}(\eta) - F_\lambda^{-1}(\eta)\|_2$ where $\eta$ is a random variable uniformly distributed in $(0,1)$. Since $F_\nu^{-1}(\eta)$ and $F_\lambda^{-1}(\eta)$ 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. [1], p. 80 and relation (5.15) on p. 74) that $d$ is separable and generates the same Borel $\sigma$-algebra as $\pi$. Let $(\eta_n)$ be a sequence of independent r.v.'s, uniformly distributed over $(0,1)$. Then $\xi_n^{(\nu)} = F_\nu^{-1}(\eta_n)$ and $\xi_n^{(\lambda)} = F_\lambda^{-1}(\eta_n)$ are i.i.d. sequences with distribution $\nu$ and $\lambda$, respectively. Using these sequences in (3.3), the left hand side is at most $\|\sum_{i=1}^n a_i(\xi_i^{(\nu)} - \xi_i^{(\lambda)})\|_p$ and since $\xi_n^{(\nu)} - \xi_n^{(\lambda)} = F_\nu^{-1}(\eta_n) - F_\lambda^{-1}(\eta_n)$ is also an i.i.d. sequence with mean 0 and variance $d(\nu,\lambda)^2$, using (3.5) and the first relation of (3.6) we get that the left hand side of (3.3) is at most $Kd(\nu,\lambda)\psi(a_1,\ldots,a_n)$ with some constant $K > 0$. On the other hand, by the definition of $d$ we have $\mathbb{E}d(\mu,0)^p = \mathbb{E}(\mathrm{Var}(\mu))^{p/2} = \mathbb{E}(\int_{-\infty}^{\infty} x^2 d\mu(x))^{p/2} < \infty$ by our assumption (1.2). Thus $\mu$ satisfies also $\mathbb{E}d(\mu,0)^p < +\infty$ and the sufficiency of (1.2) is proved.

We now turn to the proof of necessity of (1.2). 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 $P(A) > 0$,

$$P_A\left\{\left|\frac{1}{\sqrt{N}}\sum_{k=1}^N X_{m_k}\right| \geq T\right\} \leq 1/2 \qquad \text{for } T \geq T_0,\ N \geq 1 \tag{3.9}$$

where $T_0$ depends on $A$ and the sequence $(X_n)$. We show first that

$$\int_{-\infty}^{\infty} x^2 d\mu(x) < \infty \qquad \text{a.s.} \tag{3.10}$$

Let $F_\bullet(x)$ 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

$$\lim_{t\to\infty} \int_{|x|\le t} x^2 dF_\bullet(x) = +\infty \quad \text{on } B. \tag{3.11}$$

By Egorov's theorem there exists a set $B^* \subset B$ with $P(B^*) \ge P(B)/2$ such that on $B^*$ (3.11) holds uniformly, i.e. there exists a positive, nondecreasing, nonrandom function $K_t \to +\infty$ such that

$$\int_{|x|\le t} x^2 dF_\bullet(x) \ge K_t \;\; \text{on} \;\; B^*. \tag{3.12}$$

Also,

$$\int_{|x|\ge t} dF_\bullet(x) \longrightarrow 0 \quad \text{a.s. as } t \to \infty \tag{3.13}$$

and thus we can choose a set $B^{**} \subset B^*$ with $P(B^{**}) \ge P(B^*)/2$ such that on $B^{**}$ relation (3.13) holds uniformly, i.e. there exists a positive, nonincreasing, nonrandom function $\varepsilon_t \to 0$ such that

$$\int_{|x|\ge t} dF_\bullet(x) \le \varepsilon_t \;\; \text{on} \;\; B^{**}. \tag{3.14}$$

We show that there exists a subsequence $(X_{m_k})$ of $(X_n)$ such that (3.9) 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 (3.12), (3.14) hold on the whole probability space.

Let $C$ be an arbitrary set in the probability space with $P(C) > 0$. Integrating (3.12), (3.14) on $C$ and using (1.1) and Lemma 2.3 we get

$$\int_{|x|\le t} x^2 dF_C(x) \ge K_t, \qquad \int_{|x|\ge t} dF_C(x) \le \varepsilon_t, \quad t \in H \tag{3.15}$$

where $H$ denotes the set of continuity points of $F_C$. Since the integrals in (3.15) are monotone functions of $t$ and $H$ is dense, (3.15) remains valid with $K_{t/2}$ resp. $\varepsilon_{t/2}$ if we drop the assumption $t \in H$. Thus, keeping the original notation, in the sequel we can assume that (3.15) holds for all $t > 0$. Choose now $t_0$ so large that $\varepsilon_{t_0} \le 1/16$ and then choose $t_1 > t_0$ so large that

$$K_t^{1/2} \ge 4t_0 \;\; \text{for} \;\; t \ge t_1.$$

Then for $t \geq t_1$ we have, using the second relation of (3.15),

$$\Big|\int_{|x|\leq t} x dF_C(x)\Big| \leq t_0 + \int\limits_{t_0\leq |x|\leq t} |x| dF_C(x)$$

$$\leq t_0 + \Big(\int\limits_{|x|\geq t_0} dF_C(x)\Big)^{1/2}\Big(\int\limits_{|x|\leq t} x^2 dF_C(x)\Big)^{1/2}$$

$$\leq t_0 + \frac{1}{4}\Big(\int\limits_{|x|\leq t} x^2 dF_C(x)\Big)^{1/2} \leq \frac{1}{2}\Big(\int\limits_{|x|\leq t} x^2 dF_C(x)\Big)^{1/2}$$

and thus for any $C \subset \Omega$ with $P(C) > 0$ we have

$$\int_{|x|\leq t} x^2 dF_C(x) - 2\Big(\int_{|x|\leq t} x dF_C(x)\Big)^2 \geq \frac{1}{2}K_t, \qquad t \geq t_1. \tag{3.16}$$

Let further $(\varepsilon_n)$ tend to 0 so rapidly that

$$\sum_{j=a_k+1}^{\infty} \varepsilon_j \leq 2^{-k}. \tag{3.17}$$

Let $a_k = [\log k + 1]$ $(k = 1, 2, \ldots)$. By Lemma 2.1 there exists a subsequence $(X_{m_k})$ and a sequence $(Y_k)$ of discrete r.v.'s such that (2.2) 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

$$\sum_{B\in\Gamma_1} P(B) \leq \varepsilon_{a_k+1} \leq 2^{-k} \tag{3.18}$$

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

$$P_B\big(|Y_j - Z_j^{(B)}| \geq 2^{-k}\big) \leq 2^{-k} \quad (j = a_k + 1, \ldots, k). \tag{3.19}$$

Set

$$S_{a_k,k}^{(B)} = \sum_{j=a_k+1}^{k} Z_j^{(B)}, \qquad B \in \Gamma_2$$

$$\overline{S}_{a_k,k} = \sum_{B\in\Gamma_2} S_{a_k,k}^{(B)} I(B),$$

where $I(B)$ denotes the indicator function of $B$. By (3.19) and $k2^{-k} \leq 1$,

$$P_B\Big(\Big|\sum_{j=a_k+1}^{k} Y_j - \sum_{j=a_k+1}^{k} Z_j^{(B)}\Big| \geq 1\Big) \leq k2^{-k}, \qquad B \in \Gamma_2$$

and thus using (3.18) we get

$$P\Big(\Big|\sum_{j=a_k+1}^{k} Y_j - \overline{S}_{a_k,k}\Big| \geq 1\Big) \leq (k+1)2^{-k}. \tag{3.20}$$

Since $\|X_n\|_1 = O(1)$, by the Markov inequality we have

$$P\Big(\Big|\sum_{j=1}^{a_k} X_{m_j}\Big| \geq a_k k^{1/4}\Big) \leq a_k \max_{1\leq j\leq a_k} P\big(|X_{m_j}| \geq k^{1/4}\big)$$
$$\leq \text{const} \cdot (\log k + 1)k^{-1/4} =: \delta_k$$

which, together with (3.20), (2.2) and (3.17), yields

$$P\Big(\Big|\sum_{j=1}^{k} X_{m_j} - \overline{S}_{a_k,k}\Big| \geq 3a_k k^{1/4}\Big) \leq (k+2)2^{-k} + \delta_k. \tag{3.21}$$

Applying Lemma 2.2 to the i.i.d. sequence $\{Z_j^{(B)}, a_k + 1 \leq j \leq k\}$ and using (3.16) with $C = B$, $a_k \leq k/2$ and the monotonicity of $K_t$ we get for any $T \geq 2$,

$$P_B\Big(\Big|\frac{S_{a_k,k}^{(B)}}{\sqrt{k}}\Big| \leq T\Big) \leq P_B\Big(\frac{|S_{a_k,k}^{(B)}|}{\sqrt{k-a_k}} \leq 2T\Big) \leq \text{const} \cdot 2TK_{2T\sqrt{k-a_k}}^{-1/2}$$
$$\leq \text{const} \cdot TK_{2T\sqrt{k/2}}^{-1/2}$$

where the constants are absolute. Thus using (3.18) it follows that

$$P\Big(\Big|\frac{\overline{S}_{a_k,k}}{\sqrt{k}}\Big| \leq T\Big) \leq \text{const} \cdot TK_{\sqrt{k/2}}^{-1/2} + 2^{-k}. \tag{3.22}$$

Using (3.21), (3.22) and $a_k \leq \log k$ it follows that

$$P\left(\left|\frac{1}{\sqrt{k}}\sum_{j=1}^{k} X_{m_j}\right| \leq T\right) \leq P\Big(\Big|\frac{\overline{S}_{a_k,k}}{\sqrt{k}}\Big| \leq T + 3a_k k^{-1/4}\Big) + (k+2)2^{-k} + \delta_k$$
$$\leq \text{const} \cdot TK_{\sqrt{k/2}}^{-1/2} + (k+2)2^{-k} + \delta_k \longrightarrow 0 \qquad \text{as } k \to \infty$$

for any fixed $T \geq 2$ which clearly contradicts to (3.9) with $A = \Omega$. This completes the proof of (3.10).

Since $X_n \longrightarrow 0$ weakly in $L^p$, we have $\int_{-\infty}^{\infty} x d\mu(x) = 0$ a.s., and thus by Aldous'version of the subsequence principle (Theorem 6 of [1]), applied to the central limit theorem, implies that there exists a subsequence$(X_{n_k})$ such that

$$N^{-1/2}\sum_{k=1}^{N} X_{n_k} \xrightarrow{d} G \tag{3.23}$$

where $G$ is the randomized normal distribution $N(0, Y^2)$, i.e. the distribution of $Y\zeta$ 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. By Fatou's lemma we have

$$\|Y\zeta\|_p \leq \liminf_{N\to\infty} \|N^{-1/2}\sum_{k=1}^{N} X_{n_k}\|_p < \infty \tag{3.24}$$

where the second inequality follows from the equivalence of $(X_n)$ to the unit vector basis of $\ell^2$ in $L^p$. (We note that (3.23) holds only in distribution, but by a well known result of Skorokhod (see e.g. [4], 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 (3.23), $W$ has distribution $G$ and $W_N \longrightarrow W$ a.s., and thus Fatou's lemma applies). Since $Y$ and $\zeta$ are independent, (3.24) implies $E|Y|^p < \infty$, i.e. (1.2) holds, completing the proof of the converse part of Theorem 1.1.

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

[9] 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.

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

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

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

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