UNIVERSITATIS IAGELLONICAE ACTA MATHEMATICA, FASCICULUS XLIII
2005

STRONG LAW OF LARGE NUMBERS FOR OPTIMAL POINTS

by Anna Denkowska

Dedicated to my husband Maciej.

Abstract. This paper was inspired by the work of B. Beauzamy and S. Guerre [**3**], who gave a new version of the strong law of large numbers taking a generalization of Cesaro averages and then considering independent random variables with values in L_p spaces. We first investigate analogues of this theorem with Cesaro-type averages given by Orlicz functions and then we modify the random variables so as to place ourselves in a modular space.

1. Introduction. In [**3**] B. Beauzamy and S. Guerre introduced a summation process generalizing the Cesaro averages, which permitted them to obtain new versions of the strong law of large numbers, also for random variables with values in L_p spaces.

Our aim is to investigate under what kind of hypothesis one can obtain a strong law of large numbers with Cesaro-type averages given by an Orlicz function or a sequence of Orlicz functions. Then we turn to considering random variables defining functions in a uniformly convex Banach space of measurable functions. Finally, with reference to [**9**], we consider the problem in modular spaces.

For more information about geometrical properties of Musielak–Orlicz spaces see e.g. [**4**], [**5**], [**7**], [**8**]. One may found notions related to probability theory in Banach spaces in [**11**].

1991 *Mathematics Subject Classification.* 46A80, 46E30, 60F15.

Key words and phrases. Strong law of large numbers, Orlicz functions, Luxemburg norm, modular spaces, Musielak–Orlicz spaces.

Throughout this paper (Ω, A, P) denotes a probability space and $X_1(\omega)$, $X_2(\omega), \ldots$ are *independent and identically distributed* (iid, for short) random variables. Unless stated otherwise, they are supposed to take values in $\mathbb{R}$.

Any function $\Phi\colon \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ which is strictly convex, differentiable and such that $\Phi(0) = 0$ will be called an *Orlicz function*.

Given such a function, for $t \in \mathbb{R}$ and $\omega \in \Omega$, we can define a *Cesaro-type average* of the form

$$\varphi_n(t, \omega) := \frac{1}{n} \sum_{k=1}^{n} \Phi(\mid t - X_k(\omega) \mid).$$

Most of the time, we omit the variable ω and write simply $\varphi_n(t)$ instead of $\varphi_n(t, \omega)$ as long as it does not lead to confusion. This function $\varphi_n(t)$ may be regarded as a kind of distance from the point $(X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}^n$ to the diagonal.

Observe that whenever ω is fixed, $\varphi_n(t)$ is a strictly convex function such that there is the unique point $S_n(\omega)$ in which $\varphi_n(t)$ attains its minimum. This obviously defines a new random variable. It has analogous properties to these described in [**3**].

Remark now that if we assume that the expectation

$$E(\Phi'(\mid X_k(\omega) \mid) \operatorname{sgn}(X_k(\omega))) = 0,$$

then

$$(*) \qquad \frac{\partial}{\partial t} \varphi_n(t, \omega)\big|_{t=0} = \varphi_n'(0, \omega) = \frac{1}{n} \sum_{k=1}^{n} \Phi'(\mid X_k(\omega) \mid) \operatorname{sgn}(X_k(\omega)) \xrightarrow{\;1\;} 0,$$

applying the standard (Khintchine's) strong law of large numbers (here the arrow $\xrightarrow{\;1\;}$ denotes convergence with probability 1, i.e. almost surely). Indeed, in the strong law of large numbers there is $\frac{P_1 + \ldots + P_n}{n} \xrightarrow{\;1\;} 0$ for iid random variables P_n such that $E(P_n) = 0$. It is clear that $\Phi'(|X_n(\omega)|)\operatorname{sgn}(X_n(\omega))$ satisfy these assumptions.

On the other hand, from the convexity of $\varphi_n(t)$ we obtain

$$(**) \qquad 0 \leq \varphi_n(0, \omega) - \varphi_n(S_n(\omega), \omega) \leq \mid S_n(\omega) \mid \cdot \mid \varphi_n'(0, \omega) \mid,$$

(since the graph of a convex function is contained in the upper half plane delimited by any of its supporting lines).

We finally define

$$\delta_M(\varepsilon) := \inf\left\{ 1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x) + \Phi(y)};\ x, y \in \mathbb{R}_+,\ x, y \leq M,\ \mid x - y \mid \geq \varepsilon \right\}.$$

Since Φ is strictly convex, there is $\Phi(\frac{x+y}{2}) < \frac{\Phi(x)+\Phi(y)}{2}$, and so $\frac{2\Phi(\frac{x+y}{2})}{\Phi(x)+\Phi(y)} < 1$. Thus, $\delta_M(\varepsilon) > 0$, since the infimum is taken on a compact set.

The inequality $1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x)+\Phi(y)} \geq \delta_M(\varepsilon)$ is obvious and it is equivalent to

$$(***) \qquad \Phi\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(1-\delta_M(\varepsilon))(\Phi(x)+\Phi(y)),$$

for $x, y \in \mathbb{R}_+$ such that $x, y \leq M$ and $|x-y| \geq \varepsilon$.

2. Strong law of large numbers for Orlicz functions and for modulars. We maintain the notations introduced in the first section and we begin with the following easy lemma:

Lemma 2.1. *let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function such that $\Phi(0) = 0$. Then for all $t \geq 0$ and $\lambda \geq 1$ there is $\Phi(\lambda t) \geq \lambda\Phi(t)$.*

Proof. By the convexity of Φ,

$$\Phi(t) = \Phi\left(\frac{1}{\lambda}(\lambda t)\right) \leq \left(1-\frac{1}{\lambda}\right)\Phi(0) + \frac{1}{\lambda}\Phi(\lambda t),$$

which gives the result. □

Now we turn to proving the following lemma.

Lemma 2.2. *Suppose that the variables $X_1(\omega), X_2(\omega), \ldots$ are pointwise bounded. Then*

$$2\delta_M(|\, S_n(\omega)\,|)\Phi\left(\left|\frac{S_n(\omega)}{2}\right|\right) \leq \varphi_n(0) - \varphi_n(S_n(\omega)) \leq |\, \varphi_n'(0)\,||\, S_n(\omega)\,|,$$

for a well-chosen $M = M(\omega) > 0$.

Proof. Fix $\omega \in \Omega$, $n \in \mathbb{N}$ and put $(m_1, \ldots, m_n) := (X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}^n$. Let s denote the minimum point of $\varphi_n(t) = \frac{1}{n}\sum_{j=1}^n \Phi(|t-m_j|)$ and suppose that

$$M > \max\{|s-m_1|, \ldots, |s-m_n|, |m_1|, \ldots, |m_n|\}.$$

Then by $(***)$ the following holds for all t small enough:

$$\Phi\left(\left|\frac{s+t}{2} - m_j\right|\right) \leq \frac{1}{2}(1-\delta_M(|\, t-s\,|))(\Phi(|s-m_j|)+\Phi(|t-m_j|)).$$

Since the latter is equal to

$$\frac{1}{2}(\Phi(|s-m_j|)+\Phi(|t-m_j|)) - \frac{1}{2}\delta_M(|\, t-s\,|)(\Phi(|s-m_j|)+\Phi(|t-m_j|)),$$

by the convexity of Φ we obtain

$$\Phi\left(\left|\frac{s+t}{2} - m_j\right|\right) \leq \frac{\Phi(|s-m_j|)+\Phi(|t-m_j|)}{2} - \delta_M(|\, t-s\,|)\Phi\left(\left|\frac{t-s}{2}\right|\right).$$

Thence (remember that s is the minimum point)

$$\sum_{j=1}^{n} \Phi(|s-m_j|) \le \sum_{j=1}^{n} \Phi\left(\left|\frac{s+t}{2} - m_j\right|\right)$$
$$\le \sum_{j=1}^{n} \frac{\Phi(|s-m_j|)}{2} + \sum_{j=1}^{n} \frac{\Phi(|t-m_j|)}{2} - n\delta_M(|\, t-s \,|)\Phi\left(\left|\frac{t-s}{2}\right|\right).$$

Dividing by n we get

$$\frac{1}{2}\varphi_n(s) \le \frac{1}{2}\varphi_n(t) - \delta_M(|\, t-s \,|)\Phi\left(\left|\frac{t-s}{2}\right|\right),$$

whence

$$\varphi_n(t) - \varphi_n(s) \ge 2\delta_M(|\, t-s \,|)\Phi\left(\left|\frac{t-s}{2}\right|\right).$$

Since $s = S_n(\omega)$, for $t := 0$ there is

$$\varphi_n(0) - \varphi_n(S_n(\omega)) \ge 2\delta_M(|\, S_n(\omega) \,|)\Phi\left(\left|\frac{S_n(\omega)}{2}\right|\right),$$

which combined with $(*)$ ends the proof. $\square$

This lemma yields the following counterpart of one of the Beauzamy–Guerre results:

THEOREM 2.3. *If $\Phi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is an Orlicz function and*

$$E(\Phi'(|\, X_k(\omega) \,|)\,\mathrm{sgn}(X_k(\omega))) = 0$$

for $k=1,2,\dots$, and if the iid variables $X_1(\omega), X_2(\omega), \dots$ are pointwise bounded, then for the minimum point $S_n(\omega)$ of $\varphi_n(t,\omega) = \frac{1}{n}\sum_{k=1}^{n}\Phi(|\, t - X_k(\omega) \,|)$, there is $S_n(\omega) \xrightarrow{1} 0$.

PROOF. The statement follows directly from the inequalities which we have just obtained:

$$2\delta_M(|\, S_n(\omega) \,|)\Phi\left(\left|\frac{S_n(\omega)}{2}\right|\right) \le \varphi_n(0) - \varphi_n(S_n(\omega)) \le |\, \varphi_n'(0) \,||\, S_n(\omega) \,| \,.$$

Indeed,

$$\frac{2\delta_M(|\, S_n(\omega) \,|)\Phi\left(\left|\frac{S_n(\omega)}{2}\right|\right)}{|\, S_n(\omega) \,|} \le |\, \varphi_n'(0,\omega) \,| \xrightarrow{1} 0,$$

in view of $(*)$.

Suppose $S_n(\omega)$ does not converge to 0 with probability one and set

$$D := \{\omega \in \Omega;\ S_n(\omega) \not\to 0\}.$$

Then D is of a positive measure and for each $\omega \in D$ we can find $d = d(\omega) > 0$ and a subsequence $\{S_{n_k}(\omega)\}$ (with $k = k(\omega)$) such that $\mid S_{n_k}(\omega) \mid \geq d > 0$. Then by the convexity of Φ (cf. Lemma 2.1)

$$\begin{aligned} |\varphi'_{n_k}(0,\omega)| &\geq \frac{2\delta_M(\mid S_{n_k}(\omega) \mid)\Phi\big(\frac{|S_{n_k}(\omega)|}{2}\big)}{\mid S_{n_k}(\omega) \mid} \\ &= \frac{2\delta_M(\mid S_{n_k}(\omega) \mid)\Phi\big(\frac{|S_{n_k}(\omega)|}{d}\frac{d}{2}\big)}{\mid S_{n_k}(\omega) \mid} \\ &\geq \frac{2\delta_M(\mid S_{n_k}(\omega) \mid)\frac{|S_{n_k}(\omega)|}{d}\Phi(\frac{d}{2})}{\mid S_{n_k}(\omega) \mid} \\ &= \frac{2\delta_M(\mid S_{n_k}(\omega) \mid)\Phi(\frac{d}{2})}{d} \\ &\geq \frac{2\delta_M(d)\Phi(\frac{d}{2})}{d} > 0. \end{aligned}$$

Hence $\varphi_n(0,\omega) \not\to 0$ on D, which leads to a contradiction. □

If we drop the boundedness condition in the last theorem, we have to assume that the numbers

$$\delta(\varepsilon) := \inf\left\{1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x)+\Phi(y)};\ \mid x-y \mid \geq \varepsilon,\ x,y \in \mathbb{R}_+\right\}$$

are strictly positive for all ε small enough (hence for all ε), which is true for uniformly convex functions. Then the following theorem holds.

Theorem 2.4. *Let Φ be an Orlicz function and $X_1(\omega), X_2(\omega), \ldots$ a sequence of iid random variables such that*

$$E(\Phi'(\mid X_k(\omega) \mid)\operatorname{sgn}(X_k(\omega))) = 0.$$

If $\delta(\varepsilon) > 0$ holds for any $\varepsilon > 0$ and $S_n(\omega)$ are the minimum points of the function

$$\varphi_n(t,\omega) = \frac{1}{n}\sum_{k=1}^{n}\Phi(\mid t - X_k(\omega) \mid),$$

then $S_n(\omega) \xrightarrow{1} 0$.

Proof. It is analogous to the previous one and so we omit it here. □

Example 2.5. For some kind of Orlicz functions the boundedness of the random variables is not a necessary condition and we are automatically in the setting of the first theorem. Consider $\Phi(t) = t^p$ for $t \in \mathbb{R}_+,\ p > 1$. We compute

$$1 - \frac{2(\frac{x+y}{2})^p}{x^p + y^p} = 1 - \frac{\left(\frac{x+y}{2(x+y)}\right)^p}{\left(\frac{x}{x+y}\right)^p + \left(\frac{y}{x+y}\right)^p} = 1 - \frac{2\left(\frac{x}{2(x+y)} + \frac{y}{2(x+y)}\right)^p}{\left(\frac{x}{x+y}\right)^p + \left(\frac{y}{x+y}\right)^p}.$$

So there is

$$\begin{aligned} &\inf\left\{1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x) + \Phi(y)};\ x, y \in \mathbb{R}_+\ |x - y| \geq \varepsilon\right\} \\ &= \inf\left\{1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x) + \Phi(y)};\ x, y \in \mathbb{R}_+\ |x - y| \geq \varepsilon,\ x \leq 1, y \leq 1\right\}. \end{aligned}$$

Example 2.6. Among Orlicz functions such that

$$\inf\left\{1 - \frac{2\Phi(\frac{x+y}{2})}{\Phi(x) + \Phi(y)};\ x, y \in \mathbb{R}_+\ |x - y| \geq \varepsilon\right\} = 0,$$

there are functions Φ having an oblique asymptote, e.g.

$$\Phi(x) := \frac{ax^n}{bx^{n-1} + c}, \quad a \neq 0,\ n \in \mathbb{N}.$$

Indeed, if one takes $x = 0$, $y > 0$, then

$$\frac{2\Phi(\frac{x+y}{2})}{\Phi(x) + \Phi(y)} = \frac{2\Phi(\frac{y}{2})}{\Phi(y)} = \frac{2a(\frac{y}{2})^n[by^{n-1} + c]}{[b(\frac{y}{2})^{n-1} + c]ay^n} = 1.$$

Example 2.7. There exist Orlicz functions without oblique asymptotes but for which $\delta(\varepsilon) = 0$. One can easily construct an example of such a function starting from the function t^p with $p > 1$. The idea is first to take a sequence of disjoint intervals. Then to cut out the graph of t^p above such an interval, replacing it by a curve 'close' to a segment, doing this in such a way that the obtained function Φ is still differentiable. Then, obviously, $\delta(\varepsilon) = 0$.

Example 2.8. Any Orlicz function Φ which is uniformly convex gives $\delta(\varepsilon) > 0$ directly from the definition of uniform convexity, which precisely says that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that for any two points satisfying $|x - y| \geq \varepsilon$, there is

$$\Phi\left(\frac{x + y}{2}\right) \leq (1 - \delta)\left(\frac{\Phi(x) + \Phi(y)}{2}\right).$$

We now turn to considering Musielak–Orlicz modulars.

Let now $\Phi := \{\Phi_i\}_{i=1}^{\infty}$ be a Musielak–Orlicz function (i.e. all Φ_i are Orlicz functions) and put

$$\rho_{\Phi}^{n}(x) := \sum_{i=1}^{n} \Phi_i(\mid x_i \mid)$$

for a finite sequence $x = \{x_i\}_{i=1}^{n}$ of real numbers.

Set

$$\delta^{i}(\varepsilon) := \inf\left\{1 - \frac{2\Phi_i(\frac{x+y}{2})}{\Phi_i(x) + \Phi_i(y)};\ x, y \in \mathbb{R}_{+}\ \ \mid x - y \mid \geq \varepsilon,\ i = 1, \ldots, n\right\},$$

and

$$\delta^{\Phi}(\varepsilon) := \inf\{\delta^{i}(\varepsilon),\ i = 1, 2, \ldots\}.$$

Analogously, for any $M > 0$, we define $\delta_{M}^{\Phi}(\varepsilon)$.

Finally, if

$$\varphi_n(t, \omega) := \frac{1}{n}\rho_{\Phi}^{n}((t, \ldots, t) - (X_1(\omega), \ldots, X_n(\omega)),$$

then it is a strictly convex function with a (unique) minimum point (it follows from the fact, that a strictly convex non-decreasing function composed with a convex one is still strictly convex; and if f_1, f_2 are strictly convex functions both having a minimum point, then $f_1 + f_2$ is strictly convex and has a minimum point, automatically unique).

Obviously

$$\varphi_n'(0, \omega) = \frac{1}{n}\sum_{i=1}^{n} \Phi_i'(|X_i(\omega)|)\,\mathrm{sgn}(X_i(\omega)).$$

Hence, if the considered variables are independent and such that the variables $\Phi_i'(|X_i(\omega)|)\mathrm{sgn}(X_i(\omega))$ are identically distributed and have expectation zero, then by the strong law of large numbers $\varphi_n'(0) \xrightarrow{1} 0$ (compare with $(*)$). Thus the following theorem is true.

THEOREM 2.9. *Let $\{\Phi_i\}_{i=1}^{\infty}$ be a sequence of Orlicz functions and let $S_n(\omega)$ denote the minimum point of the strictly convex function*

$$\varphi_n(t, \omega) = \frac{1}{n}\sum_{i=1}^{n} \Phi_i(|t - X_i(\omega)|),$$

where the random variables $X_i(\omega)$ are independent and such that the variables $\Phi_i'(|X_i(\omega)|)\mathrm{sgn}(X_i(\omega))$ are identically distributed. Assume that

$$E(\Phi_i'(|X_i(\omega)|)\mathrm{sgn}(X_i(\omega))) = 0, \quad i = 1, 2, \ldots.$$

If moreover one of the following conditions is fulfilled:

(i) $\delta^{\Phi}(\varepsilon) > 0$ *for any* $\varepsilon > 0$ *and the function* $\widetilde{\Phi}(x) := \inf\{\Phi_i(x);\ i = 1, 2, \dots\}$ *is strictly positive for all* $x > 0$*;*

(ii) *the random variables* $X_1(\omega), X_2(\omega), \dots$ *have a common pointwise bound and* $\delta^{\Phi}_M(\varepsilon) > 0$ *(which is automatically verified if* Φ *consists of a finite number of different functions* Φ_i*),*

then $S_n(\omega) \xrightarrow{1} 0$*.*

Proof. Fix $\omega \in \Omega$. By the convexity of φ, as earlier we obtain

$$0 \leq \varphi_n(0) - \varphi_n(S_n(\omega)) \leq \mid \varphi'_n(0) \mid\mid S_n(\omega) \mid .$$

Executing similar computations as in Lemma 2.2, we get

$$\frac{2\delta^{\Phi}(\mid S_n(\omega) \mid) \sum_{i=1}^{n} \Phi_i\big(\big|\frac{S_n(\omega)}{2}\big|\big)}{n \mid S_n(\omega) \mid} \leq \mid \varphi'_n(0) \mid .$$

Put $D := \{\omega \in \Omega;\ S_n(\omega) \not\to 0\}$ and suppose that $P(D) > 0$. Then for each $\omega \in D$ we can find $d = d(\omega) > 0$ and a subsequence $S_{n_k}(\omega) \geq d$, $k = 1, 2, \dots$

Thus, by Lemma 2.1, for $\omega \in D$, there is

$$\begin{aligned}\frac{2\delta^{\Phi}(\mid S_{n_k}(\omega) \mid) \sum_{i=1}^{n_k} \Phi_i\big(\big|\frac{S_{n_k}(\omega)}{2}\big|\big)}{n_k \mid S_{n_k}(\omega) \mid} &\geq \frac{2\delta^{\Phi}(\mid S_{n_k}(\omega) \mid) \frac{|S_{n_k}(\omega)|}{d(\omega)} \sum_{i=1}^{n_k} \Phi_i\big(\frac{d(\omega)}{2}\big)}{n_k \mid S_{n_k}(\omega) \mid} \\ &\geq \frac{2\delta^{\Phi}(d(\omega)) n_k \widetilde{\Phi}\big(\frac{d(\omega)}{2}\big)}{n_k d(\omega)} > 0.\end{aligned}$$

Hence for $\omega \in D$ we have found a subsequence $\mid \varphi'_{n_k}(0, \omega) \mid \geq c(\omega) > 0$. That means that $\mid \varphi'_n(0, \omega) \mid \not\to 0$ for all $\omega \in D$, which, D being of a positive measure, leads to a contradiction. □

3. Minimum points in uniformly convex Banach space. Consider a uniformly convex Banach space $(X, ||\cdot||)$, where X is a subspace of $L_0([a, b])$, the space of measurable functions, and suppose that X contains the constants, which we shall denote by $f_t \equiv t$.

The following theorem is true.

Theorem 3.1. *In the setting introduced, suppose that the norm is of class* $\mathcal{C}^1$ *and* X^1, X^2, $\dots$ *are* X*-valued independent random variables such that* $||X^n(\omega)|| \leq M = M(\omega)$, $n = 1, 2, \dots$ *Let* $S_n(\omega) = f_{S_n(\omega)}$ *denote the minimum point of the strictly convex function*

$$\phi_n \colon \mathbb{R} \ni t \mapsto \phi_n(t, \omega) := ||f_t - X^n(\omega)|| \in \mathbb{R}$$

and suppose that $\phi'_n(0) \xrightarrow{1} 0$*.*

If for all $t \in \mathbb{R}$ $\liminf_{n \to +\infty} ||f_t - X^n(\omega)|| > 0$ *for almost all* ω*, then also* $S_n(\omega) \xrightarrow{1} 0$*.*

Proof. By the uniform convexity of X, for all $x_1, x_2 \in X$ with $||x_1|| = ||x_2|| = 1$, if $||x_1 - x_2|| \geq \varepsilon$, then

$$\left|\left|\frac{x_1 + x_2}{2}\right|\right| \leq (1 - \delta(\varepsilon)).$$

For all $x, y \in X$ such that $||x|| = ||y|| = d > 0$ put $x_1 = \frac{x}{||x||}$ i $x_2 = \frac{y}{||y||}$. Then we obtain the following characterization of the uniform convexity

$$||x - y|| \geq \varepsilon \Longrightarrow \left|\left|\frac{x + y}{2}\right|\right| \leq \left(1 - \delta\left(\frac{\varepsilon}{d}\right)\right)\left(\frac{||x|| + ||y||}{2}\right),$$

which follows from straightforward computation.

Now take $x, y \in X$ and suppose that $||y|| \geq ||x|| > 0$. There exists $\alpha \in (0, 1]$ such that $||x|| = ||\alpha y||$. By the uniform convexity condition we get

$$||x + \alpha y|| \leq \left(1 - \delta\left(\frac{||x - \alpha y||}{||x||}\right)\right)(||x|| + \alpha||y||).$$

Consequently,

$$\begin{aligned}
||x + y|| &\leq ||x + \alpha y|| + (1 - \alpha)||y|| \\
&\leq \left(1 - \delta\left(\frac{||x - \alpha y||}{||x||}\right)\right)(||x|| + \alpha||y||) + (1 - \alpha)||y|| \\
&= \left(1 - \delta\left(\frac{||x - \alpha y||}{||x||}\right)\right)||x|| + \left(1 - \alpha\delta\left(\frac{||x - \alpha y||}{||x||}\right)\right)||y|| \\
&\leq \left(1 - \alpha\delta\left(\frac{||x - \alpha y||}{||x||}\right)\right)(||x|| + ||y||).
\end{aligned}$$

Now take $x_n = S_n(\omega) - X^n(\omega)$ and $y_n = -X^n(\omega)$. Since $S_n(\omega)$ is the element of best approximation among constants for $X^n(\omega)$, there holds $||x_n|| \leq ||y_n||$ and we can find $\alpha_n \in (0, 1]$ such that the norms of x_n and $\alpha_n y_n$ are equal. By the above inequality we obtain

$$\begin{aligned}
&||S_n(\omega) - 2X^n(\omega)|| \\
&\leq \left(1 - \alpha_n\delta\left(\frac{||S_n(\omega) - (1 - \alpha_n)X^n(\omega)||}{||S_n(\omega) - X^n(\omega)||}\right)\right)(||S_n(\omega) - X^n(\omega)|| + ||X^n(\omega)||).
\end{aligned}$$

Since $2S_n(\omega)$ is the element of best approximation among constants for $2X^n(\omega)$, there also is

$$2||S_n(\omega) - X^n(\omega)|| \leq ||S_n(\omega) - 2X^n(\omega)||.$$

On the other hand $||S_n(\omega)-X^n(\omega)||+||-X^n(\omega)|| \geq ||S_n(\omega)-2X^n(\omega)||$. Thus we finally obtain

$$\begin{aligned} &||X^n(\omega)|| - ||S_n(\omega) - X^n(\omega)|| \\ &\qquad \geq \alpha_n \delta\Big(\frac{||S_n(\omega) - (1-\alpha_n)X^n(\omega)||}{||S_n(\omega) - X^n(\omega)||}\Big)||S_n(\omega) - 2X^n(\omega)||. \end{aligned}$$

Observe that on the left-hand side we have in fact $\phi_n(0,\omega) - \phi_n(S_n(\omega),\omega)$, which is obviously not greater than $|\phi'_n(0,\omega)||S_n(\omega)|$ (by the strict convexity of $\phi_n(\cdot,\omega)$). Thus

$$|\phi'_n(0,\omega)||S_n(\omega)| \geq \alpha_n \delta\Big(\frac{||S_n(\omega) - (1-\alpha_n)X^n(\omega)||}{||S_n(\omega) - X^n(\omega)||}\Big)2||S_n(\omega) - X^n(\omega)||.$$

Set $D_S := \{\omega \in \Omega;\ S_n(\omega) \not\to 0\}$, $D_\phi := \{\omega \in \Omega;\ \phi'_n(0,\omega) \not\to 0\}$ and $D := \{\omega \in \Omega;\ \exists n_k \to +\infty\colon ||S_{n_k}(\omega) - X^{n_k}(\omega)|| \to 0\}$. By the assumptions, there is $P(D) = P(D_\phi) = 0$. Thus $P(D_S) = P(D_S \setminus (D_\phi \cup D))$.

Suppose that $P(D_S) > 0$ and take $\omega \in D_S \setminus (D_\phi \cup D)$. Then there exists $d = d(\omega) > 0$ and a sequence $n_k \to +\infty$ such that $|S_{n_k}(\omega)| \geq d$. By the choice of ω, there is a constant $c = c(\omega) > 0$ such that $||S_{n_k}(\omega) - X^{n_k}(\omega)|| \geq c$.

Since $||X^n(\omega)|| > 0$, then $\alpha_n = \frac{||S_n(\omega)-X^n(\omega)||}{||X^n(\omega)||}$, and thus $\alpha_{n_k} \geq \frac{c}{M(\omega)} > 0$.

Observe also that

$$\begin{aligned} ||S_n(\omega)|| &\leq ||S_n(\omega) - X^n(\omega)|| + ||X^n(\omega)|| \\ &= ||S_n(\omega) - X^n(\omega)|| + \frac{||S_n(\omega) - X^n(\omega)||}{\alpha_n} \\ &\leq \frac{2||S_n(\omega) - X^n(\omega)||}{\alpha_n}, \end{aligned}$$

whence finally

$$||S_n(\omega) - X^n(\omega)|| \geq \frac{\alpha_n}{2}||S_n(\omega)|| = \frac{\alpha_n}{2}||f_1|| \cdot |S_n(\omega)|.$$

Besides, $\delta(\varepsilon)$ is decreasing with ε; thus, since $||S_n(\omega)-X^n(\omega)|| \leq ||X^n(\omega)|| \leq M(\omega)$, then

$$\delta\Big(\frac{||S_n(\omega) - (1-\alpha_n)X^n(\omega)||}{||S_n(\omega) - X^n(\omega)||}\Big) \geq \delta\Big(\frac{||S_n(\omega) - (1-\alpha_n)X^n(\omega)||}{M(\omega)}\Big).$$

Therefore, we obtain

$$|\phi'_{n_k}(0,\omega)||S_{n_k}(\omega)| \geq \delta\Big(\frac{||S_{n_k}(\omega) - (1-\alpha_{n_k})X^{n_k}(\omega)||}{M(\omega)}\Big)\frac{c^2}{M(\omega)^2}|S_{n_k}(\omega)|.$$

Since $\alpha_{n_k} \in (0,1]$, we may assume (possibly extracting a subsequence) that $\alpha_{n_k} \to \alpha$. Obviously, $\alpha \in [0,1]$, but we already know that $\alpha \neq 0$.

If there were $\alpha = 1$, then for any sufficiently large k we would obtain

$$\begin{aligned}||S_{n_k}(\omega) - (1-\alpha_{n_k})X^{n_k}(\omega)|| &\geq \frac{||S_{n_k}(\omega)||}{2} = \frac{||f_1||}{2}|S_{n_k}(\omega)| \\ &\geq \frac{||f_1||}{2}d =: N(\omega).\end{aligned}$$

On the other hand, if $\alpha < 1$, then in view of the fact that

$$||S_n(\omega)|| \leq ||S_n(\omega) - X^n(\omega)|| + ||X^n(\omega)|| \leq 2M(\omega),$$

we may assume (possibly extracting subsequences from $S_{n_k}(\omega)$ and from $X^{n_k}(\omega)$) that $S_{n_k}(\omega)$ and $X^{n_k}(\omega)$ are convergent. Let $S(\omega)$, $X(\omega)$ denote the corresponding limits.

Now if there were (for these subsequences) $||S_{n_k}(\omega)-(1-\alpha_{n_k})X^{n_k}(\omega)|| \to 0$, then

$$\begin{aligned}||S(\omega) - (1-\alpha)X^{n_k}(\omega)|| \leq &||S(\omega) - S_{n_k}(\omega)|| + ||S_{n_k}(\omega) - (1-\alpha_{n_k})X^{n_k}(\omega)|| \\ &+ ||(1-\alpha_{n_k})X^{n_k}(\omega) - (1-\alpha)X^{n_k}(\omega)||\end{aligned}$$

would lead to

$$\left|\left|\frac{S(\omega)}{(1-\alpha)} - X^{n_k}(\omega)\right|\right| \to 0,$$

which contradicts our assumptions, since $\frac{S(\omega)}{1-\alpha}$ is a constant.

Thus there exists $N(\omega) > 0$ such that

$$||S_{n_k}(\omega) - (1-\alpha_{n_k})X^{n_k}(\omega)|| \geq N(\omega).$$

This finally yields

$$|\phi'_{n_k}(0,\omega)||S_{n_k}(\omega)| \geq \delta\left(\frac{N(\omega)}{M(\omega)}\right)\frac{c^2}{M(\omega)^2}|S_{n_k}(\omega)|,$$

whence

$$|\phi'_{n_k}(0,\omega)| \geq \delta\left(\frac{N(\omega)}{M(\omega)}\right)\frac{c^2}{M(\omega)^2} > 0.$$

But that implies $|\phi'_{n_k}(0,\omega)| \not\to 0$ for $\omega \in D_S \setminus (D_\phi \cup D)$ and since this set is of a positive measure we get a contradiction. $\square$

4. Strong law of large numbers in Orlicz spaces. In this section, we will consider independent random variables $X_n(\omega)$ with values in an Orlicz space defined as follows:

Consider the measure space $([0,1],\mu)$ with some Borel finite measure μ. If Φ is an Orlicz function, then by L_Φ we will denote the space of all μ-measurable functions $f:[0,1]\to\mathbb{R}$ such that $\lim_{\lambda\to 0^+}\rho_\Phi(\lambda f)=0$, where ρ_Φ is the modular defined by

$$\rho_\Phi(f) := \int_0^1 \Phi(|f(\tau)|)d\mu(\tau).$$

This obviously means that for some $\lambda>0$ there is $\int_0^1 \Phi(|\lambda f(\tau)|)d\mu(\tau) < +\infty$. Under some assumptions on Φ this will be satisfied for all $\lambda>0$. All the previously introduced definitions have their respectives analogues in this case too. The space L_Φ with the norm $||\cdot||_\Phi$ defined as earlier is a Banach space. Moreover, since Φ is convex and tends to infinity when $t\to+\infty$, $L_\Phi\subset L_1$. Actually, if $\Phi(t)=t^p$ for some $p>1$, then $L_\Phi=L_p$.

We now consider a function

$$\varphi_n(t,\omega)=\frac{1}{n}\sum_{k=1}^{n}\rho_\Phi(|f_t - X_k(\omega)|)\ for\ t\in\mathbb{R},\omega\in\Omega,$$

where $f_t\equiv t$ is a constant function (constants obviously belong to L_Φ). As earlier, by $S_n(\omega)\in L_\Phi$ we denote the point at which the convex function $\varphi_n(\cdot,\omega)$ attains its minimum. Note that $S_n(\omega)$ is a constant function.

From now on, assume that Φ is of class $\mathcal{C}^1$. If we calculate the derivative of $t\mapsto\varphi_n(t,\omega)$ (with ω fixed), we obviously get

$$\begin{aligned}\frac{d}{dt}\varphi_n(t,\omega) &= \lim_{s\to t}\frac{\varphi_n(s,\omega)-\varphi_n(t,\omega)}{s-t}\\ &= \int_0^1 \frac{1}{n}\sum_{k=1}^{n}\Phi'(|t-X_k(\omega)(\tau)|)\mathrm{sgn}(t-X_k(\omega)(\tau))d\mu(\tau).\end{aligned}$$

To be able to apply usual versions of the strong law of large numbers, we have to recall the following definitions (cf. [**6**]):

Definition 4.1. Let $\{X_n\}$ be a sequence of X-valued random variables satisfying $EX_n=0$. $\{X_n\}$ is said to satisfy the strong law of large numbers if

$$\frac{1}{n}\sum_{j=1}^{n}X_j \xrightarrow{1} 0.$$

To be able to show that $\varphi_n'(0)\xrightarrow{1}0$ (which is crucial for our purposes), it would be enough to know that the variables $X_k'(\omega):=\Phi'(|X_k(\omega)|)\mathrm{sgn}(X_k(\omega))$ satisfy the strong law of large numbers. The first problem we meet is, however, the fact that $X_k'(\omega)$ have their values in a function space. Then the problem is: under what kind of assumptions X_k' will satisfy the strong law of large numbers.

Observe that for L_p spaces, $X_k'(\omega)\in L_q$ for $q=\frac{p}{p-1}$ (cf. [**3**]). We will need more preparatory work.

Recall that a Rademacher sequence is a sequence $\{r_i\}$ of independent random variables taking the values 1 and -1 with probability $1/2$. We take r_i on $[0, 1]$.

DEFINITION 4.2. A Banach space X is said to be of (Rademacher) type p for some $1 \le p < +\infty$ if there is a constant $C > 0$ such that for all finite sequences $\{x_i\}_1^n \subset X$,

$$(\star) \qquad E\Big|\Big|\sum_{i=1}^{n} r_i x_i\Big|\Big|^p \le C \sum_{i=1}^{n} ||x_i||^p.$$

It is clear (by the triangle inequality) that every Banach space is of type 1. Besides, Khintchine's classic inequalities imply that the definition makes sense only for $p \le 2$.

Equivalently, X is of type p with $1 \le p \le 2$ if for any sequence $\{x_i\}$ in X such that $\{||x_i||\} \in l_p$ the series $\sum_{i=1}^{\infty} r_i x_i$ converges a.e. (i.e. with probability 1) on $[0, 1]$. The equivalence follows from a closed graph argument (namely the inclusion of $l_p(X) \subset C(X) := \{\{x_i\};\ \sum_i r_i x_i$ *converges in probability*$\}$ has a closed graph whenever X is of type p).

Moreover, by the Kahane inequality, condition $(\star)$ is equivalent to

$$E\Big|\Big|\sum_{i=1}^{n} r_i x_i\Big|\Big| \le \text{const.}\Big(\sum_{i=1}^{n} ||x_i||^p\Big)^{\frac{1}{p}}.$$

Let us finally note that a Banach space of type p is also of type $p' \le p$ and that every Hilbert space is of type 2.

Then the following is true.

THEOREM 4.3. ([**6**] *Thm 2.1*) *Let* $1 \le p \le 2$, *then the strong law of large numbers holds for all sequences* $\{X_n\}$ *of independent* X*-valued Radon variables satisfying*

$$EX_n = 0 \text{ and } \sum_{n=1}^{+\infty} \frac{E||X_n||^p}{n^p} < \infty$$

iff the space X *is of type* p.

Recall that a Radon variable ξ is by definition a variable regular with respect to compact sets, i.e., for each $\varepsilon > 0$ there is a compact K such that $P(\xi \in K) \ge 1-\varepsilon$. Equivalently, one may assume that the space X is separable.

Before we turn to proving the main theorem of this section we give an example to illustrate it.

EXAMPLE 4.4. Consider the following Orlicz function: $\Phi_q(t) = e^{t^q} - 1$, for $q \ge 1$. For $f \in L_{\Phi_q}$ it is clear that $\rho_{\Phi_q}(\lambda f) < \infty$ iff $\int_0^1 e^{|f(t)|^q}\, d\mu(t) < \infty$. Since $|f|^q \le e^{|f|^q}$, then $f \in L_q$.

We compute $\Phi_q'(|f|)\mathrm{sgn}(f) = q|f|^{q-1}e^{|f|^q}\mathrm{sgn}(f)$ and so it is clear that it suffices to consider $|f|^{q-1}e^{|f|^q}$ only. This obviously belongs to L_p for $p = \frac{q}{q-1}$. Thus, $\Psi(t) := t^p$ is the Orlicz function sought for, if $q > 1$.

THEOREM 4.5. *If $X_1(\omega), X_2(\omega), \ldots$ are independent random variables with values in L_Φ such that $\rho_\Phi(X_k(\omega)) \le M = M(\omega)$ for all k and if, moreover,*

(i) $E(\Phi_k'(|X_k(\omega)(\tau)|)\mathrm{sgn}(X_k(\omega)(\tau))) = 0$,

(ii) *there is a Banach function space X of type p such that*

$$\Phi'(X_k(\omega))\mathrm{sgn}(X_k(\omega)) \in X$$

and

$$\sum_{k=1}^{\infty} \frac{E||\Phi'(|X_k(\omega)|)||_X}{k^p} < +\infty,$$

then $\rho_\Phi(S_n(\omega)) \xrightarrow{1} 0$, where $\varphi_n(t,\omega)$ is the function introduced at the beginning of this section and $S_n(\omega)$ is its minimum point.

PROOF. Suppose that for some ω there is a sequence $n_k \to +\infty$ such that $S_{n_k}(\omega) \not\to 0$. Then there exists $d > 0$ such that $|S_{n_k}(\omega)| \ge d$ for all k. Fix such $\alpha > 0$ that $\alpha d \ge 1$. Then by Lemma 2.1

$$\Phi\left(|S_{n_k}(\omega)|\alpha\frac{1}{\alpha}\right) \ge \alpha|S_{n_k}(\omega)|\Phi\left(\frac{1}{\alpha}\right)$$

whence

$$\begin{aligned}\rho_\Phi(|S_{n_k}(\omega)|) &\ge \Phi\left(\frac{1}{\alpha}\right)\alpha\int_0^1 |S_{n_k}(\omega)|d\mu(\tau)\\ &= \Phi\left(\frac{1}{\alpha}\right)\alpha|S_{n_k}(\omega)|\mu([0,1])\\ &\ge \Phi\left(\frac{1}{\alpha}\right)\alpha d\mu([0,1]) > 0,\end{aligned}$$

since S_{n_k} is a constant function on [0,1].

Thus $S_n(\omega) \not\to 0$ implies $\rho_\Phi(|S_n(\omega)|) \not\to 0$. On the other hand, if $S_n(\omega) \to 0$, then $\Phi(|S_n(\omega)|) \to 0$ (by continuity), and so

$$\int_0^1 \Phi(|S_n(\omega)|)d\mu(\tau) \to 0,$$

that is $\rho_\Phi(|S_n(\omega)|) \to 0$. In other words $S_n(\omega) \xrightarrow{1} 0$ iff $\rho_\Phi(|S_n(\omega)|) \xrightarrow{1} 0$. Hence we only need to show that $S_n(\omega) \xrightarrow{1} 0$.

Applying a standard argument, in view of the convexity of φ_n and computations similar to those of Lemma 2.2, we obtain

$$\frac{2\rho_\Phi(|\frac{S_n(\omega)}{2}|)\delta_M(|S_n(\omega)|)}{|S_n(\omega)|} \le |\varphi_n'(0,\omega)|.$$

So if $\omega \in \Omega$ is such that $S_n(\omega) \not\to 0$, there is a subsequence $n_k \to +\infty$ and a constant $d = d(\omega) > 0$ such that $|S_{n_k}(\omega)| \ge d$.

Now

$$\begin{aligned}\frac{2\rho_\Phi\left(\frac{|S_{n_k}(\omega)|}{2}\right)\delta_M(|S_{n_k}(\omega)|)}{|S_{n_k}(\omega)|} &\ge 2\frac{\delta_M(d)}{|S_{n_k}(\omega)|}\int_0^1 \Phi\left(\frac{|S_{n_k}(\omega)|}{2}\right) d\mu \\ &\ge 2\frac{\delta_M(d)}{|S_{n_k}(\omega)|}\Phi\left(\frac{|S_{n_k}(\omega)|}{2}\right)\mu([0,1]),\end{aligned}$$

since the integrand is a constant.

Choose now $\alpha > 0$ such that $\frac{\alpha d}{2} \ge 1$. Then

$$\Phi\left(\frac{|S_{n_k}(\omega)|}{2}\right) = \Phi\left(\frac{1}{\alpha}\alpha\frac{|S_{n_k}(\omega)|}{2}\right) \ge \alpha\frac{|S_{n_k}(\omega)|}{2}\Phi\left(\frac{1}{\alpha}\right)$$

by Lemma 2.1.

Finally we obtain

$$|\varphi_{n_k}'(0,\omega)| \ge 2\frac{\delta_M(d)}{|S_{n_k}(\omega)|}\alpha\frac{|S_{n_k}(\omega)|}{2}\Phi\left(\frac{1}{\alpha}\right)\mu([0,1]) = \delta_M(d)\alpha\Phi\left(\frac{1}{\alpha}\right)\mu([0,1]) > 0,$$

whence $\varphi_n'(0,\omega) \not\to 0$. But as $\varphi_n'(0,\omega) \xrightarrow{1} 0$ (cf. our assumptions), we get $S_n(\omega) \xrightarrow{1} 0$. □

To get a full counterpart of Proposition 2 from [**3**], we now consider the convex function

$$\psi_n(f,\omega) := \frac{1}{n}\sum_{k=1}^{n} \rho_\Phi(f - X_k(\omega)), \quad \textit{for } f \in L_\Phi, \omega \in \Omega.$$

It is (when we fix ω) the distance (computed in the sum of modulars) from the point $(X_1(\omega), \ldots, X_n(\omega)) \in (L_\Phi)^n$ to the diagonal. Thus there exists the unique $S_n(\omega) \in L_\Phi$ realizing the minimum of $\psi_n(\cdot,\omega)$ (in this case it may be non-constant).

It is clear that if by $\widetilde{S}_n(\omega)(\tau)$ we denote the minimum point of the strictly convex function

$$\widetilde{\psi}_n(\tau,\omega) \colon \mathbb{R} \ni t \mapsto \frac{1}{n}\sum_{k=1}^{n} \Phi(|t - X_k(\omega)(\tau)|) \in \mathbb{R},$$

then $\widetilde{S}_n(\omega) = S_n(\omega)$ μ-almost everywhere on $[0,1]$.

Therefore, if we apply the previous ideas to these functions, we easily get an extension of the last theorem:

THEOREM 4.6. *Under the assumptions of the previous theorem, for* $S_n(\omega) \in L_\Phi$ *being the minimum point of* ψ_n*, there is* $\rho_\Phi(S_n(\omega)) \xrightarrow{1} 0$.

Acknowledgements. I am grateful to prof. G. Lewicki for suggesting the problem and deeply indebted to him for his constant help and most valuable advice during the preparation of this paper.

References

1. Beauzamy B., *Points minimaux dans les espaces de Banach*, Comptes Rendus Acad. Sci. Paris, t. **280**, Série A (1975), 717–720.
2. Beauzamy B., Maurey B., *Points minimaux et ensembles optimaux dans les espaces de Banach*, J. Funct. Anal., **24** (1977), 107–139.
3. Beauzamy B., Guerre S., *Une loi forte des grands nombres pour des variables aléatoires non intégrables*, Comptes Rendus Acad. Sci. Paris, t. **286**, Série A (1978), 67–69.
4. Bian S., Hudzik H., Wang T., *Smooth, very smooth and strongly smooth points in Musielak–Orlicz spaces*, Bull. Austral. Math. Soc., **63** (2001), 441–457.
5. Cui Y., Hudzik H., Nowak M., Płuciennik R., *Some geometric properties in Orlicz sequence spaces equipped with the Orlicz norm*, Convex Anal., **6.1** (1999), 91–113.
6. Hoffman-Jørgensen J., Pisier G., *The law of large numbers and the central limit theorem in Banach spaces*, Ann. Probab., Vol. **4**, No. **4** (1976), 587–599.
7. Hudzik H., Kurc W., *Monotonicity properties of Musielak–Orlicz spaces and dominated best approximation in Banach lattices*, J. Approx. Theory, **95** (1998), 353–368.
8. Hudzik H., Ye Y., *Support functionals and smoothness in Musielak–Orlicz sequence spaces equipped with the Luxemburg norm*, Comment. Math. Univ. Carolin., **31.4** (1990), 661–684.
9. Kamińska A., Lewicki G., *Contractive and optimal sets in modular spaces*, Math. Nachr., **268** (2004), 74–95.
10. Katirtzoglou E., *Type and cotype of Musielak–Orlicz sequence spaces*, J. Math. Anal. Appl., **226** (1998), 431–455.
11. Ledoux M., Talagrand M., *Probability in Banach Spaces*, Ergebnisse der Math. u. ihrer Grenzgebiete, 3. Folge, Band 23, Springer-Verlag, 1991.

Received December 19, 2004

Jagiellonian University
Institute of Mathematics
ul. Reymonta 4
30-059 Kraków
Poland
e-mail: denkowsk@im.uj.edu.pl