# On the Convergence in Mean of Martingale Difference Sequences 

by Paul Abraham, John Alexopoulos and S. J. Dilworth

September 1, 1998


#### Abstract

In [6] Freniche proved that any weakly null martingale difference sequence in $L_{1}[0,1]$ has arithmetic means that converge in norm to 0 . We show any weakly null martingale difference sequence in an Orlicz space whose N -function belongs to $\nabla_{3}$ has arithmetic means that converge in norm to 0 . Then based on a theorem in Stout [13][Theorem 3.3 .9 (i) and (iii)], we give necessary and sufficient conditions for a bounded martingale difference sequence in an Orlicz space whose N -function belongs to a large class of $\Delta_{2}$ functions to have means that converge to 0 a.s. Finally, we conclude with some expository comments including an easy proof of Komlos' theorem [9] for $L_{p}[0,1], 1<p<\infty$.


## 1 Introduction

Recall that an N -function $\Phi$ is in $\Delta_{3}$ if there is a constant $c>0$ so that $\Phi(c x) \geq x \Phi(x)$ for all large values of $x$. Using the notation in [14] we say that an N -function is in $\nabla_{3}$ if its complementary N -function is in $\Delta_{3}$. In [6] Freniche proved that any weakly null martingale difference sequence (mds) in $L_{1}[0,1]$ has arithmetic means that converge to 0 in norm. In Section 2, we first establish a similar result for the class of Orlicz spaces whose generating N -functions belong to $\nabla_{3}$. In particular, we show in theorem 2.1 if $F$ is an $N$-function in $\nabla_{3}$ then any weakly null mds in $L_{F}$ has arithmetic means that converge to 0 in norm. $\nabla_{3}$ is a rather large class of N -functions. It includes, for example, all N -functions $F$ whose principal parts are defined by $F(x)=x(\log x)^{p}$ for $p \geq 1$. It is worth noting that the union of all Orlicz Spaces whose generating N -functions belong to $\nabla_{3}$ is $L_{1}[0,1][10$, pp. 60-62].

Next recall that an N -function $F$ belongs to $\Delta_{2}$ if and only if there is a constant $K>2$ so that $F(2 x) \leq K F(x)$ for sufficiently large values of $x$. The other new result in Section 2, theorem 2.3, gives several necessary and sufficient conditions for all bounded martingale difference sequences in an Orlicz space with N -function F in $\Delta_{2}$ that satisfies $\frac{F(x)}{x^{2}}$ non-increasing on $[0, \infty)$ to have means that converge to 0 a.s. Theorem 2.3 is based largely on a result stated in Stout [13, Theorem 3.3.9 (i) and (iii)]. However, the result in Stout's book is more than required. For the sake of completeness, clarity and ease of proof, in lemma 2.2 we state and prove what we need to justify theorem 2.3 .

At this stage we note that if $F$ is in $\nabla_{3}$ then $F$ is in $\Delta_{2}[10$, p. 44, Theorem 6.5]. Furthermore if $G$ denotes the complement of $F$ then $F$ is equivalent to an $N$-function $Q$ whose principal part is given by $Q(x)=x G^{-1}(x)\left[10, \mathrm{p} .37\right.$, Theorem 6.2]. In light of this fact there is no loss in assuming that if $F$ is in $\nabla_{3}$ then $F(x)=x G^{-1}(x)$ for some $G$ in $\Delta_{3}$ and all sufficiently large values of $x$. Furthermore for all such $x$ we have:

$$
\frac{d}{d x}\left(\frac{F(x)}{x^{2}}\right)=\frac{d}{d x}\left(\frac{G^{-1}(x)}{x}\right)=\frac{\frac{x}{G^{\prime}\left(G^{-1}(x)\right)}-G^{-1}(x)}{x^{2}}=\frac{x-G^{-1}(x) G^{\prime}\left(G^{-1}(x)\right)}{x^{2} G^{\prime}\left(G^{-1}(x)\right)}
$$

By letting $x=G(u)$ we get:

$$
\frac{d}{d x}\left(\frac{F(x)}{x^{2}}\right)=\frac{G(u)-u G^{\prime}(u)}{G^{2}(u) G^{\prime}(u)}<\frac{G(u)-\int_{0}^{u} G^{\prime}(t) d t}{G^{2}(u) G^{\prime}(u)}=\frac{G(u)-G(u)}{G^{2}(u) G^{\prime}(u)}=0
$$

Thus $\frac{F(x)}{x^{2}}$ is non-increasing for all sufficiently large values of $x$. But for any $F$ in $\nabla_{3}$ and any $p>1$, we have that $F(x) \leq|x|^{p}$ for large values of $x\left[10\right.$, p. 38]. Hence $\frac{F(x)}{x^{2}} \searrow 0$. It follows then that if $F$ is in $\nabla_{3}$ then $F$
is in $\Delta_{2}$ and $\frac{F(x)}{x^{2}} \searrow 0$.
In section 3, which is expository, theorem 3.1 gives an alternate proof of the following special case of theorem 2.3: any bounded $m d s$ in $L_{p}[0,1], 1<p<\infty$, must have arithmetic means that converge to 0 a.s. Note that theorem 3.1 is a known consequence of the important fact that any mds $\left(d_{n}\right)$ in $L_{p}[0,1], 1<p<2$, has upper $p$-estimates; that is, there exists a constant $C_{p}>0$ depending only on $p$ such that for all $n$ and all choices of scalars $c_{1}, \ldots, c_{n},\left\|\sum_{k=1}^{n} c_{k} d_{k}\right\|_{p}^{p} \leq C_{p}\left(\sum_{k=1}^{n}\left\|c_{k} d_{k}\right\|_{p}^{p}\right)$. The fact that any mds has upper $p$-estimates for $1<p<2$ follows from Burkholder's inequality [3, theorem 9] $\left\|\sum_{k=1}^{n} c_{k} d_{k}\right\|_{p} \leq C_{p}\left\|\left(\sum_{k=1}^{n}\left|c_{k} d_{k}\right|^{2}\right)^{\frac{1}{2}}\right\|_{p}$ and the triangle inequality $\int\left(\sum_{k=1}^{n}\left|c_{k} d_{k}\right|^{2}\right)^{\frac{p}{2}} \leq \sum_{k=1}^{n} \int\left|c_{k} d_{k}\right|^{p}$ for $\frac{p}{2}<1$. Our proof of theorem 3.1 is based on the work of Banach and Saks [2] and Freniche [6] and shows in an elementary fashion that any mds in $L_{p}[0,1], 1<p<2$, has upper $p$-estimates. It is perhaps interesting to note that versions of theorem 3.1 for sequences of independent random variables date back to Kolmogoroff [8] in the case $p=2$, Marcinkiewicz and Zygmund [12] for $p>1$ and Chung [4] for more general function spaces.

Finally, Section 3 concludes with corollary 3.2 which shows in an elementary fashion that Komlos' theorem [9] holds for bounded sequences in $L_{p}[0,1], 1<p<\infty$. In its full force, Komlos' theorem promises for any bounded sequence $\left(f_{n}\right)$ in $L_{1}[0,1]$, there exists an integrable $f$ and a subsequence $\left(g_{n}\right)$ of $\left(f_{n}\right)$ each subsequence of which has arithmetic means that converge a.s. to $f$. An impressive but surprisingly easy-to-prove result of Gaposhkin [7] will be recalled in Section 3 to help establish corollary 3.2.

We conclude the introduction with some background material required for theorem 2.1. A subset $\mathcal{K}$ of $L^{1}[0,1]$ is called uniformly integrable if: given $\varepsilon>0$ there is a $\delta>0$ so that $\sup \left\{\int_{E}|f| d \lambda: f \in \mathcal{K}\right\}<\varepsilon$ whenever $\lambda(E)<\delta$. Alternatively $\mathcal{K}$ is bounded and uniformly integrable if and only if given $\varepsilon>0$ there is an $N>0$ so that $\sup \left\{\int_{[|f|>N]}|f| d \lambda: f \in \mathcal{K}\right\}<\varepsilon$. A concept similar to uniform integrability is that of equi-absolute continuity of norms. We say that a collection $\mathcal{K} \subset L_{F}$ has equi-absolutely continuous norms if:

$$
\forall \varepsilon>0 \exists \delta>0 \text { so that } \sup \left\{\left\|\chi_{E} f\right\|_{F}: f \in \mathcal{K}\right\}<\varepsilon \text { whenever } \lambda(E)<\delta
$$

Alternatively $\mathcal{K}$ is bounded and has equi-absolutely continuous norms if: given $\varepsilon>0$ there is an $N>0$ so that $\sup \left\{\left\|f \cdot \chi_{[|f|>N]}\right\|_{F}: f \in \mathcal{K}\right\}<\varepsilon$. The next result resembles the theorem of Dunford and Pettis. For its proof the reader should consult [1, Corollary 2.9].

Theorem 1.1 Let $F \in \Delta_{2}$ and suppose that its complement $G$ satisfies:

$$
\lim _{t \rightarrow \infty} \frac{G(c t)}{G(t)}=\infty \text { for some } c>0
$$

Then a bounded set $\mathcal{K} \subset L_{F}$ is relatively weakly compact if and only if $\mathcal{K}$ has equi-absolutely continuous norms.

Note that if $F$ is in $\nabla_{3}$ then $F$ satisfies the hypothesis of Theorem 1.1.

## 2 Mean Convergence Results for MDSs in Two Large Classes of Orlicz Spaces

Theorem 2.1 If $F$ is an $N$-function in $\nabla_{3}$ then any weakly null mds in $L_{F}$ has arithmetic means that converge to 0 in norm.

Proof: The argument is essentially that of Freniche in [6].
Let $\left(d_{n}\right)$ be a weakly null mds in $L_{F}$. Let $\varepsilon>0$. For any $M>0$ let $e_{n}=d_{n} \cdot \chi_{\left[\left|d_{n}\right| \leq M\right]}, \tilde{e}_{n}=d_{n} \cdot \chi_{\left[\left|d_{n}\right|>M\right]}$ and for $n \geq 2$ let $f_{n}=\mathbb{E}\left(\mathrm{e}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)$. Notice that $\left(e_{n}-f_{n}\right)$ is a mds, uniformly bounded by $2 M$ and hence in $L_{2}[0,1]$. The conditioning now ensures that the sequence $\left(e_{n}-f_{n}\right)$ is orthogonal. Notice that:

$$
0=\mathbb{E}\left(\mathrm{d}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)=\mathbb{E}\left(\mathrm{e}_{\mathrm{n}}+\tilde{\mathrm{e}}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)=\mathbb{E}\left(\mathrm{e}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)+\mathbb{E}\left(\tilde{\mathrm{e}}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)
$$

and so:

$$
f_{n}=\mathbb{E}\left(\mathrm{e}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)=-\mathbb{E}\left(\tilde{\mathrm{e}}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right)
$$

Now for any $\sigma\left(d_{1}, \ldots, d_{n-1}\right)$ measurable $g$ in $L_{G}$, with $\|g\|_{G} \leq 1$ we have:

$$
\begin{aligned}
\int_{0}^{1} g f_{n} d \lambda & =-\int_{0}^{1} g \mathbb{E}\left(\tilde{\mathrm{e}}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right) \mathrm{d} \lambda \\
& =-\int_{0}^{1} \mathbb{E}\left(\tilde{\mathrm{~g}}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right) \mathrm{d} \lambda \\
& =-\int_{0}^{1} g \tilde{e}_{n} d \lambda \\
& \leq\|g\|_{G} \cdot\left\|\tilde{e}_{n}\right\|_{F} \\
& =\|g\|_{G} \cdot\left\|d_{n} \cdot \chi_{\left[\left|d_{n}\right|>M\right]}\right\|_{F}
\end{aligned}
$$

Hence,

$$
\left\|f_{n}\right\|_{F} \leq\left\|d_{n} \cdot \chi_{\left[\left|d_{n}\right|>M\right]}\right\|_{F} .
$$

Now use theorem 1.1 to choose $M>0$ big enough to ensure that $\sup _{n}\left\|d_{n} \cdot \chi_{\left.\|\left|d_{n}\right|>M\right]}\right\|_{F}<\varepsilon$. Since $\frac{F(x)}{x^{2}} \searrow 0$, the map $L_{2} \hookrightarrow L_{F}$ is continuous and thus there is a constant $K>0$ so that $\|f\|_{F} \leq K\|f\|_{2}$ for all $f$ in $L_{2}$. For any positive integer $N$ we then have:

$$
\begin{aligned}
\left\|\sum_{n=1}^{N} d_{n}\right\|_{F} & \leq\left\|\sum_{n=1}^{N}\left(d_{n}-e_{n}\right)\right\|_{F}+\left\|\sum_{n=1}^{N}\left(e_{n}-f_{n}\right)\right\|_{F}+\left\|\sum_{n=1}^{N} f_{n}\right\|_{F} \\
& \leq \sum_{n=1}^{N}\left\|d_{n} \cdot \chi_{\left\lfloor\left|d_{n}\right|>M\right]}\right\|_{F}+K\left\|\sum_{n=1}^{N}\left(e_{n}-f_{n}\right)\right\|_{2}+\sum_{n=1}^{N}\left\|d_{n} \cdot \chi_{\left[\left|d_{n}\right|>M\right]}\right\|_{F} \\
& \leq N \varepsilon+2 M K \sqrt{N}+N \varepsilon
\end{aligned}
$$

and so,

$$
\frac{1}{N}\left\|\sum_{n=1}^{N} d_{n}\right\|_{F} \leq 2 \varepsilon+\frac{2 M K}{\sqrt{N}} .
$$

Therefore,

$$
\lim _{N} \frac{1}{N}\left\|\sum_{n=1}^{N} d_{n}\right\|_{F}=0
$$

The following lemma is a special case of a result stated in Stout [13, Theorem 3.3 .9 (i) and (iii)] suitable for our needs. The proof is included for the sake of completeness.

Lemma 2.2 Suppose $\phi$ is a positive even function with $\frac{\phi(x)}{x}$ non-decreasing and $\frac{\phi(x)}{x^{2}}$ non-increasing on $[0, \infty)$ and $\left(a_{n}\right)$ is a sequence of positive numbers with $a_{n} \nearrow \infty$.
(a) If $\left(d_{n}\right)$ is an mds on $[0,1]$ such that $\sum_{n=1}^{\infty} \frac{\mathbb{E}\left(\phi\left(\mathrm{d}_{\mathrm{n}}\right)\right)}{\phi\left(a_{n}\right)}<\infty$ then $\frac{1}{a_{n}} \sum_{i=1}^{n} d_{i}$ converges to 0 a.s.
(b) On the other hand, given any sequence of positive numbers $\left(m_{i}\right)$ such that $\sum_{n=1}^{\infty} \frac{m_{n}}{\phi\left(a_{n}\right)}=\infty$ there exist independent symmetric random variables $\left(X_{n}\right)$ on $[0,1]$ such that $\mathbb{E}\left(\phi\left(\mathrm{X}_{\mathrm{n}}\right)\right)=\mathrm{m}_{\mathrm{n}}$ for all $n$ and $\frac{1}{a_{n}} \sum_{i=1}^{n} X_{i}$ diverges a.s.

Proof: For part (a) we adapt Chung's argument (see [5, Theorem 5.4.1] or [4]) for a sequence of independent random variables to hold for an mds. Suppose $\phi$ and $\left(a_{n}\right)$ are as stated and $\left(d_{n}\right)$ is an mds on $[0,1]$ such that $\Sigma_{n=1}^{\infty} \frac{\mathbb{E}\left(\phi\left(\mathrm{d}_{\mathrm{n}}\right)\right)}{\phi\left(a_{n}\right)}<\infty$. For all $n$, let $F_{n}$ denote the distribution function of $d_{n}$ and $Y_{n}=d_{n} \chi_{\left[\left|d_{n}\right| \leq a_{n}\right]}$. Arguing exactly as Chung,

$$
\sum_{n=1}^{\infty} \sigma^{2}\left(\frac{Y_{n}}{a_{n}}\right) \leq \sum_{n=1}^{\infty} \mathbb{E}\left(\frac{\mathrm{Y}_{\mathrm{n}}^{2}}{\mathrm{a}_{\mathrm{n}}^{2}}\right) \leq \sum_{\mathrm{n}=1}^{\infty} \int_{\left[|\mathrm{x}| \leq \mathrm{a}_{\mathrm{n}}\right]} \frac{\phi(\mathrm{x})}{\phi\left(\mathrm{a}_{\mathrm{n}}\right)} \mathrm{dF}_{\mathrm{n}}(\mathrm{x}) \leq \sum_{\mathrm{n}=1}^{\infty} \frac{\mathbb{E}\left(\phi\left(\mathrm{d}_{\mathrm{n}}\right)\right)}{\phi\left(\mathrm{a}_{\mathrm{n}}\right)}<\infty
$$

noting the second inequality holds since $\frac{\phi(x)}{x^{2}}$ non-increasing implies $\frac{x^{2}}{a_{n}^{2}} \leq \frac{\phi(x)}{\phi\left(a_{n}\right)}$ for $|x| \leq a_{n}$. Thus, by the Kolmogoroff Maximal Inequality for martingale difference sequences (see [11, Theorem Ia p. 236 and pp. 386-387]) it follows that $\sum_{n=2}^{\infty} \frac{\left.Y_{n}-\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)\right)}{a_{n}}$ converges a.s.
Next observe that for all $n \geq 2, \mathbb{E}\left(\mathrm{Y}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)=\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}}-\mathrm{d}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)$ a.s. since $\mathbb{E}\left(\mathrm{d}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)=$ 0 a.s. Therefore, $\left\|\frac{\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)}{a_{n}}\right\|_{1}=\left\|\frac{\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}}-\mathrm{d}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)}{a_{n}}\right\|_{1} \leq\left\|\frac{Y_{n}-d_{n}}{a_{n}}\right\|_{1}$. Now,

$$
\left\|\frac{Y_{n}-d_{n}}{a_{n}}\right\|_{1}=\int_{\left[\left|d_{n}\right|>a_{n}\right]} \frac{\left|d_{n}\right|}{a_{n}} d \lambda=\int_{\left[|x|>a_{n}\right]} \frac{|x|}{a_{n}} d F_{n}(x) \leq \int_{\left[|x|>a_{n}\right]} \frac{\phi(x)}{\phi\left(a_{n}\right)} d F_{n}(x) \leq \frac{\mathbb{E}\left(\phi\left(\mathrm{d}_{\mathrm{n}}\right)\right)}{\phi\left(a_{n}\right)},
$$

noting the first inequality holds since $\frac{\phi(x)}{x}$ non-decreasing implies $\frac{|x|}{a_{n}} \leq \frac{\phi(x)}{\phi\left(a_{n}\right)}$ for $|x|>a_{n}$. Hence, $\sum_{n=2}^{\infty}\left\|\frac{\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)}{a_{n}}\right\|_{1}<\infty$ forcing $\sum_{n=2}^{\infty} \frac{\mathbb{E}\left(\mathrm{Y}_{\mathrm{n}} \mid \mathrm{Y}_{1}, \ldots, \mathrm{Y}_{\mathrm{n}-1}\right)}{a_{n}}$ to be convergent a.s. Therefore, $\sum_{n=1}^{\infty} \frac{Y_{n}}{a_{n}}$ converges a.s.

Following Chung's argument again, we get:

$$
\sum_{n=1}^{\infty} \lambda\left[d_{n} \neq Y_{n}\right]=\sum_{n=1}^{\infty} \int_{\left[|x|>a_{n}\right]} d F_{n}(x) \leq \sum_{n=1}^{\infty} \int_{\left[|x|>a_{n}\right]} \frac{\phi(x)}{\phi\left(a_{n}\right)} d F_{n}(x) \leq \sum_{n=1}^{\infty} \frac{\mathbb{E}\left(\phi\left(\mathrm{d}_{n}\right)\right.}{\phi\left(a_{n}\right)}<\infty
$$

where the first inequality holds since $\frac{\phi(x)}{x}$ non-decreasing forces $\phi$ to be non-decreasing. Therefore, the Borel-Cantelli lemma implies that $\sum_{n=1}^{\infty} \frac{d_{n}}{a_{n}}$ converges a.s. Finally, Kronecker's Lemma gives $\frac{1}{a_{n}} \sum_{i=1}^{n} d_{i}$ converges to 0 a.s.

The proof of (b) was given by Chung in [4].

Theorem 2.3 Let $F$ be an $N$-function satisfying the $\Delta_{2}$-condition such that $\frac{F(x)}{x^{2}}$ is non-increasing on $[0, \infty)$ and $\left(a_{n}\right)$ is a sequence of positive numbers with $a_{n} \nearrow \infty$. The following are equivalent:
(a) $\frac{1}{a_{n}} \sum_{k=1}^{n} d_{k} \rightarrow 0$ a.s. for every $m d s$ bounded in the $L_{F}$ norm.
(b) $\frac{1}{a_{n}} \sum_{k=1}^{n} d_{k} \rightarrow 0$ a.s. for every weakly null mds in $L_{F}$.
(c) $\frac{1}{a_{n}} \sum_{k=1}^{n} X_{k} \rightarrow 0$ a.s. for every norm null sequence of mean zero independent symmetric random variables in $L_{F}$.
(d) $\sum_{n=1}^{\infty} \frac{1}{F\left(a_{n}\right)}<\infty$.

Proof: Clearly $(a) \Rightarrow(b) \Rightarrow(c)$ regardless of the $N$-function $F$. Suppose $F$ is $\Delta_{2}$ and satisfies $\frac{F(x)}{x^{2}}$ is non-increasing on $[0, \infty)$. Note that since $F$ is an N-function, by definition $\frac{F(x)}{x} \nearrow \infty$.

Now, we prove $(c) \Rightarrow(d)$ by contraposition. Suppose $\sum_{n=1}^{\infty} \frac{1}{F\left(a_{n}\right)}=\infty$. Then there is a sequence $\left(b_{n}\right)$ of positive numbers such that $b_{n} \rightarrow 0$ and $\sum_{n=1}^{\infty} \frac{b_{n}}{F\left(a_{n}\right)}=\infty$. Then by Lemma 2.2(b), there exist independent
symmetric mean 0 random variables $\left(X_{i}\right)$ such that $\mathbb{E}\left(\mathrm{F}\left(\mathrm{X}_{\mathrm{i}}\right)\right)=\mathrm{b}_{\mathrm{i}}$ for all $i$ and $\frac{1}{a_{n}} \sum_{i=1}^{n} X_{i}$ diverges a.s. Since $b_{n} \rightarrow 0$ we have that $\int_{0}^{1} F\left(X_{n}\right) d \lambda \rightarrow 0$ and as $F$ is in $\Delta_{2},\left\|X_{n}\right\|_{F} \rightarrow 0$. Hence $(c) \Rightarrow(d)$.

Finally, we proceed to establish $(d) \Rightarrow(a)$. Let $\left(d_{n}\right)$ be a mds bounded in the $L_{F}$ norm and suppose that $\sum_{n=1}^{\infty} \frac{1}{F\left(a_{n}\right)}<\infty$. Since $\left(\left\|d_{n}\right\|_{F}\right)_{n}$ is bounded and $F$ is in $\Delta_{2}$ we have that $\left(\int_{0}^{1} F\left(d_{n}\right) d \lambda\right)_{n}$ is bounded as well. Hence $\sum_{i=2}^{\infty} \int_{0}^{1} \frac{F\left(d_{i}\right)}{F\left(a_{i}\right)} d \lambda<\infty$. So by Lemma $2.2(\mathrm{a}), \frac{1}{a_{n}} \sum_{k=1}^{n} d_{k} \rightarrow 0$ a.s. Hence $(d) \Rightarrow(a)$.

Note that if $F$ is in $\nabla_{3}$ then $F$ satisfies the hypothesis of Theorem 2.3. The following Corollary is then immediate:

Corollary 2.4 If $F$ is an $N$-function in $\nabla_{3}$ then any weakly null mds in $L_{F}$ has arithmetic means that converge to 0 in norm. Furthermore, the arithmetic means of a weakly null mds converge to 0 a.s. if and only if $\sum_{n=1}^{\infty} \frac{1}{F(n)}<\infty$.

## 3 Some Remarks on Convergence in Arithmetic Means of Bounded

 Sequences in $L_{p}[0,1], 1<p<\infty$As noted in the introduction, the fact that any bounded martingale difference sequence in $L_{p}[0,1], 1<p<\infty$, has arithmetic means that converge almost surely to 0 follows from the fact that martingale difference sequences in $L_{p}[0,1], 1<p<2$, possess upper $p$-estimates. From weak compactness and Egorov's theorem, it then readily follows that any bounded martingale difference sequence in $L_{p}[0,1], 1<p<\infty$, has arithmetic means that converge in norm to 0 . However based on the work of Banach and Saks [2], Freniche [6] was able to give a surprisingly simple proof of the latter result. With a little more work, in theorem 3.1 we extend Freniche's method of proof to derive the existence of upper $p$-estimates for martingale difference sequences in $L_{p}[0,1], 1<p<2$, from which the stronger conclusion of almost sure convergence of the arithmetic means of a bounded mds to 0 easily follows.

Theorem 3.1 If $\left(d_{k}\right)$ is a bounded $m d s$ in $L_{p}[0,1], 1<p<\infty$, then $\frac{1}{n} \sum_{i=1}^{n} d_{i}$ converges almost surely to 0 .

Proof: Note that by monotonicity of the $L_{p}$ norms, it suffices to prove the theorem for $1<p<2$. Let $1<p<2$ be given. Let $q$ be such that $\frac{1}{p}+\frac{1}{q}=1$. We begin by establishing the existence of the well-known
upper $p$-estimates for ( $d_{n}$ ) for all $n$ and all scalars $c_{1}, \ldots, c_{n}$ :

$$
\left\|\sum_{k=1}^{n} c_{k} d_{k}\right\|_{p}^{p} \leq C_{p}^{p} \sum_{k=1}^{n}\left|c_{k}\right|^{p}\left\|d_{k}\right\|_{p}^{p}
$$

for some constant $C_{p}$ depending only on $p$. Let $s_{n}=\sum_{k=1}^{n} c_{k} d_{k}$. Thanks to the work of Banach-Saks [2, (1), p. 52], we know there exists a constant $A \geq 1$ such that:

$$
|a+b|^{p} \leq|a|^{p}+p|a|^{p-1} \operatorname{sgn}(a) b+A|b|^{p}
$$

holds for all real numbers $a$ and $b$. So, it follows that a.s.:

$$
\begin{aligned}
\left|s_{n}\right|^{p} & =\left|s_{n-1}+c_{n} d_{n}\right|^{p} \\
& \leq\left|s_{n-1}\right|^{p}+p\left|s_{n-1}\right|^{p-1} \operatorname{sgn}\left(s_{n-1}\right) c_{n} d_{n}+A\left|c_{n} d_{n}\right|^{p}
\end{aligned}
$$

Note that $d_{n}$ is in $L_{p}$ and $\left|s_{n-1}\right|^{p-1}$ is in $L_{q}$. Hence, $c_{n}\left|s_{n-1}\right|^{p-1} \operatorname{sgn}\left(s_{n-1}\right) d_{n}$ is in $L_{1}$. Furthermore, observe that $c_{n}\left|s_{n-1}\right|^{p-1} \operatorname{sgn}\left(s_{n-1}\right)$ is $\sigma\left(d_{1}, \ldots, d_{n-1}\right)$ measurable. Hence,

$$
\begin{aligned}
\int_{0}^{1} c_{n}\left|s_{n-1}\right|^{p-1} \operatorname{sgn}\left(s_{n-1}\right) d_{n} d \lambda & =\int_{0}^{1} \mathbb{E}\left(\mathrm{c}_{\mathrm{n}}\left|\mathrm{~s}_{\mathrm{n}-1}\right|^{\mathrm{p}-1} \operatorname{sgn}\left(\mathrm{~s}_{\mathrm{n}-1}\right) \mathrm{d}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right) \mathrm{d} \lambda \\
& =\int_{0}^{1} c_{n}\left|s_{n-1}\right|^{p-1} \operatorname{sgn}\left(s_{n-1}\right) \mathbb{E}\left(\mathrm{d}_{\mathrm{n}} \mid \mathrm{d}_{1}, \ldots, \mathrm{~d}_{\mathrm{n}-1}\right) \mathrm{d} \lambda \\
& =0
\end{aligned}
$$

Therefore, we have:

$$
\left\|s_{n}\right\|_{p}^{p} \leq\left\|s_{n-1}\right\|_{p}^{p}+A\left\|c_{n} d_{n}\right\|_{p}^{p}
$$

Reapplying the previous inequality to $s_{n-1}$ gives $\left\|s_{n-1}\right\|_{p}^{p} \leq\left\|s_{n-2}\right\|_{p}^{p}+A\left\|c_{n-1} d_{n-1}\right\|_{p}^{p}$, and so:

$$
\left\|s_{n}\right\|_{p}^{p} \leq\left\|s_{n-2}\right\|_{p}^{p}+A\left(\left\|c_{n-1} d_{n-1}\right\|_{p}^{p}+\left\|c_{n} d_{n}\right\|_{p}^{p}\right)
$$

Continuing in the same fashion the upper p-estimates are established:

$$
\left\|s_{n}\right\|_{p}^{p} \leq A\left(\sum_{k=1}^{n}\left|c_{k}\right|^{p}\left\|d_{k}\right\|_{p}^{p}\right)
$$

for all $n$. With $M=\sup _{n}\left\|d_{n}\right\|_{p}$ and $c_{k}=\frac{1}{k}$, it follows that $\sup _{n}\left\|s_{n}\right\|_{p}^{p} \leq A M^{p}\left(\sum_{k=1}^{\infty} \frac{1}{k^{p}}\right)<\infty$. Hence, $\left(s_{n}\right)$ is an $L_{p}$-bounded martingale and so $\left(s_{n}\right)$ converges a.s. Since $\frac{1}{c_{k}}=k \nearrow \infty$, Kronecker's lemma gives $\sum_{k=1}^{n} \frac{d_{k}}{c_{n}}=\sum_{k=1}^{n} \frac{d_{k}}{n}$ converges to 0 a.s.

The ultimate theorem in almost sure convergence in arithmetic means is that of Komlos in [9]. Theorem 3.1 can be used to obtain this result in the special case where $\left(f_{n}\right)$ is bounded in $L_{p}[0,1]$ for some $1<p<\infty$. A theorem of Gaposhkin in [7] is needed:

Theorem (Gaposhkin): If $\left(f_{n}\right)$ is a weakly null sequence in $L_{p}[0,1], 1 \leq p<\infty$ then there is a subsequence $\left(f_{n_{k}}\right)$ of $\left(f_{n}\right)$ and a mds $\left(d_{k}\right)$ in $L_{p}[0,1]$ so that

$$
\sum_{k=1}^{\infty}\left\|f_{n_{k}}-d_{k}\right\|_{p}<\infty
$$

With this theorem in hand, we proceed into establishing the following:

Corollary 3.2 Let $\left(f_{n}\right)$ be a bounded sequence in $L_{p}[0,1], 1<p<\infty$. Then there is a function $f$ in $L_{p}[0,1]$ and a subsequence $\left(h_{n}\right)$ of $\left(f_{n}\right)$, each subsequence of which converges in arithmetic mean to $f$ a.s.

Proof: Let $\left(f_{n}\right)$ be a bounded sequence in $L_{p}[0,1]$. Since $L_{p}[0,1]$ is reflexive, there is a function $f$ in $L_{p}[0,1]$ and a subsequence $\left(g_{n}\right)$ of $\left(f_{n}\right)$, so that $g_{n} \rightarrow f$ weakly. The sequence $\left(g_{n}-f\right)$ is weakly null and so by Gaposhkin's theorem, there is a subsequence $\left(h_{n}\right)$ of $\left(g_{n}\right)$ and a mds $\left(d_{n}\right)$ in $L_{p}[0,1]$ so that

$$
\sum_{n=1}^{\infty}\left\|\left(h_{n}-f\right)-d_{n}\right\|_{p}<\infty
$$

Let $\left(h_{n_{k}}\right)$ be any subsequence of $\left(h_{n}\right)$ and let $c_{k}=h_{n_{k}}-f$. Then, $\sum_{k=1}^{\infty}\left\|c_{k}-d_{n_{k}}\right\|_{p}<\infty$, and so the series $\sum_{k=1}^{\infty}\left(c_{k}-d_{n_{k}}\right)$ converges to some function in $L_{p}[0,1]$ a.s. Thus, $\left(c_{k}-d_{n_{k}}\right) \rightarrow 0$ a.s. Now,

$$
\frac{1}{N} \sum_{k=1}^{N} c_{k}=\frac{1}{N} \sum_{k=1}^{N}\left(c_{k}-d_{n_{k}}\right)+\frac{1}{N} \sum_{k=1}^{N} d_{n_{k}}
$$

Notice that $\left(d_{n_{k}}\right)$ is still a mds in $L_{p}[0,1]$ and so $\frac{1}{N} \sum_{k=1}^{N} d_{n_{k}} \rightarrow 0$ almost surely, thanks to theorem 3.1. Hence $\frac{1}{N} \sum_{k=1}^{N} c_{n_{k}} \rightarrow 0$ almost surely. So,

$$
\frac{1}{N} \sum_{k=1}^{N} h_{n_{k}} \rightarrow f \text { almost surely. }
$$

## References

[1] J. Alexopoulos, De La Vallée Poussin's Theorem and Weakly Compact Sets in Orlicz Spaces, QM 17 (1994), pp. 231-248.
[2] S. Banach and S. Saks, Sur la convergence forte dans les champs L ${ }^{p}$, Studia Math, 2 (1930), pp. 51-57.
[3] D. L. Burkholder, Martingale Transforms, Annals of Mathematical Statistics, 37 (1966), pp. 1494-1504.
[4] K. L. Chung, Note on Some Strong Laws of Large Numbers, American Journal of Mathematics, 69 (1947), pp.189-192.
[5] K. L. Chung, A Course in Probability Theory, 2nd edition, Academic Press, 1974.
[6] F. J. Freniche, Cesaro convergence of martingale difference sequences and the Banach-Saks and Szlenk Theorems, Proceedings of the American Math Society, 103 (1988), pp. 234-236.
[7] V. F. Gaposhkin, Convergence and Limit Theorems for Sequences of Random Variables, Theory of Probability and its Applications, 17 (1972), pp. 379-400.
[8] A. Kolmogoroff, Uber die Summen durch den Zufall bestimmten unabhangiger Grossen, Mathematische Annalen, 99 (1928), pp. 309-319.
[9] J. Komlos, A Generalization of a Problem of Steinhaus, Acta Mathematica Academiae Scientiarum Hungaricae, 18 (1967), pp. 217-229.
[10] M. A. Krasnoselskii and Ya. B. Rutickii, Convex functions and Orlicz Spaces, Noorhoff Ltd., Groningen, 1961.
[11] M. Loeve, Probability Theory, 3rd edition, Van Nostrand, 1963.
[12] J. Marcinkiewicz and A. Zygmund, Sur les fonctiones independantes, Fundamenta Mathematicae, 29 (1937), pp. 60-90.
[13] W. F. Stout, Almost Sure Convergence, Academic Press, 1974.
[14] M. M. Rao and Z. D. Ren, Theory of Orlicz Spaces, Marcel Dekker, 1991.

Paul Abraham, pabraham@stark.kent.edu
Department of Mathematics \& Computer Science
Kent State University, Stark Campus
6000 Frank Avenue N.W.

John Alexopoulos, jalexopoulos@stark.kent.edu Department of Mathematics \& Computer Science Kent State University, Stark Campus 6000 Frank Avenue N.W.
S. J. Dilworth, dilworth@math.sc.edu

Department of Mathematics
University of South Carolina
Columbia, SC 29208

