WEAK NORM INEQUALITIES FOR MARTINGALES AND GEOMETRY OF BANACH SPACES

ADAM OSĘKOWSKI

Abstract. Let f, g be two Hilbert-space-valued martingales such that g is differentially subordinate to f. The paper contains the proof of the estimate

$$||g||_{p,\infty} \leq \frac{2p(p+1)}{p-1}||f||_{p,\infty}, \qquad 1 < p < \infty.$$

The constant is shown to be of optimal order for $p \to \infty$ and for $p \to 1$. Related results for transforms of UMD-valued martingales are also established.

1. Introduction

Martingale theory is a powerful tool in the study of the geometry of Banach spaces: see the survey [4] of Burkholder for the wide overview of the subject. The objective of this paper is to establish some novel martingale inequalities and to explore their connections with the structure of the Banach space in which the martingales take values. In particular, this will yield some new and interesting characterizations of UMD spaces.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, filtered by $(\mathcal{F}_n)_{n\geq 0}$, a non-decreasing sequence of sub-σ-algebras of $\mathcal{F}$. Let $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ be two adapted martingales with values in a given separable Banach space $(\mathbb{B}, |\cdot|)$. We define $df = (df_n)_{n\geq 0}$ and $dg = (dg_n)_{n\geq 0}$, the difference sequences of f and g, by $df_0 = f_0$, $df_n = f_n - f_{n-1}$, $n = 1, 2, \ldots$, and similarly for dg. We will use the notation $f_n^* = \sup_{0\leq k\leq n} |f_k|$, $n = 0, 1, 2, \ldots$, and for $1 \leq p < \infty$, we will write

$$||f||_p = \sup_{n\geq 0} ||f_n||_p \quad \text{and} \quad ||f||_{p,\infty} = \sup_{n\geq 0} ||f_n||_{p,\infty} = \sup_{n\geq 0}\sup_{\lambda>0} \left[\lambda^p \mathbb{P}(|f_n| \geq \lambda)\right]^{1/p}$$

for the strong and the weak p-th norm of f. Following Burkholder [2], we say that g is *differentially subordinate* to f, if for any $n \geq 0$ we have

$$|dg_n| \leq |df_n|$$

with probability 1. For example, this holds if g is a transform of f by a predictable sequence $v = (v_n)_{n\geq 0}$ taking values in $[-1, 1]$ (that is, we have $dg_n = v_n df_n$ for all n and by predictability of v we mean that each term v_n is measurable with respect to $\mathcal{F}_{(n-1)\vee 0}$). If $\mathbb{B}$ is a Hilbert space, then the differential subordination implies many interesting inequalities between f and g, which can be applied in many areas of mathematics. In addition, there is a beautiful method, due to Burkholder, which allows to determine optimal constants in such estimates. The method rests on the existence of a certain special function, having appropriate convexity-type

2010 *Mathematics Subject Classification.* Primary: 60G42. Secondary: 46C15.

Key words and phrases. martingale, martingale transform, differential subordination, UMD space.

Partially supported by MNiSW Grant N N201 397437.

properties; for the detailed description and more on the subject, see the survey [3]. We will only mention here two types of inequalities, which are closely related to the results obtained in this paper. A celebrated L^p-inequality of Burkholder [2] states that if $\mathbb{B}$ is a Hilbert space and g is differentially subordinate to f, then for any $1 < p < \infty$ we have the sharp bound

$$||g||_p \leq \max\left\{p-1, (p-1)^{-1}\right\}||f||_p. \tag{1.1}$$

For $p = 1$ the above inequality does not hold with any finite constant, but we have the weak-type bound $||g||_{1,\infty} \leq 2||f||_1$, and the constant 2 is optimal: see [2]. In fact, Burkholder proved the sharp weak-type estimate for a wider range of parameters: if $1 \leq p \leq 2$, then for f, g as above,

$$||g||_{p,\infty} \leq \left(\frac{2}{\Gamma(p+1)}\right)^{1/p} ||f||_p. \tag{1.2}$$

What are the best constants for $p > 2$? The answer is due to Suh [10]:

$$||g||_{p,\infty} \leq \left(p^{p-1}/2\right)^{1/p} ||f||_p. \tag{1.3}$$

We continue this line of research and study the following novel estimates between the weak p-th norms of f and g. Until the end of this section, the filtration and the probability space are assumed to vary.

Theorem 1.1. *Suppose that $\mathbb{B}$ is a Hilbert space and f, g are $\mathbb{B}$-valued martingales such that g is differentially subordinate to f. Then for any $1 < p < \infty$ we have*

$$||g||_{p,\infty} \leq \frac{2p(p+1)}{p-1}||f||_{p,\infty}. \tag{1.4}$$

The constant is of optimal order $O(p)$ as $p \to \infty$ and $O((p-1)^{-1})$ as $p \to 1$, even in the special case when $\mathbb{B} = \mathbb{R}$ and g is a transform of f by a deterministic sequence with values in $[-1, 1]$.

Unfortunately, we have not managed to determine the sharp version of the estimate above. This is due to the fact that Burkholder's method, which is so efficient in the proofs of (1.1), (1.2) and (1.3), does not seem to be applicable here (nevertheless, we use it to obtain some intermediate estimate; see Section 2 below).

One may wonder whether the above result can be carried over to a wider class of Banach spaces. Of course, if $\mathbb{B}$ is isomorphic to a Hilbert space, then the inequality (1.4) still holds true, possibly with a different constant. We will show that this implication can be reversed.

Theorem 1.2. *Suppose that a Banach space $\mathbb{B}$ has the following property: there are $p \in (1, \infty)$ and $C < \infty$ satisfying*

$$||g||_{p,\infty} \leq C||f||_{p,\infty} \tag{1.5}$$

for all $\mathbb{B}$-valued martingales f, g such that g is differentially subordinate to f. Then $\mathbb{B}$ is isomorphic to a Hilbert space.

However, if we restrict ourselves to the class of martingale transforms, we obtain a larger class of Banach spaces which are well-behaved with respect to (1.4). Recall that $\mathbb{B}$ is a UMD space if there is a constant $K = K(\mathbb{B})$ with the following property: if f is a $\mathbb{B}$-valued martingale and g is its transform by a real predictable sequence bounded in absolute value by 1, then

$$\lambda\mathbb{P}(g_n^* \geq \lambda) \leq K||f_n||_1$$

for every integer n and all $\lambda > 0$. For alternative definitions of UMD spaces and their geometrical characterizations, see e.g. [1] and [4].

Theorem 1.3. *The following conditions are equivalent.*
(i) $\mathbb{B}$ *is UMD.*
(ii) There is a finite constant $\kappa(\mathbb{B})$ *depending only on* $\mathbb{B}$ *such that*

$$||g||_{p,\infty} \leq \frac{\kappa(\mathbb{B})p^2}{p-1}||f||_{p,\infty}, \qquad 1 < p < \infty, \tag{1.6}$$

whenever f *is a* $\mathcal{B}$*-valued martingale and* g *is the transform of* f *by a predictable sequence with values in* $[-1, 1]$.
(iii) There are $p \in (1,\infty)$ *and* $\kappa < \infty$ *such that*

$$||g||_{p,\infty} \leq \kappa||f||_{p,\infty} \tag{1.7}$$

whenever f *is a* $\mathcal{B}$*-valued martingale and* g *is the transform of* f *by a deterministic sequence with values in* $\{-1, 1\}$.

A few words about the organization of the paper. Theorem 1.1 is established in the next section, while the remaining results are proved in Section 3.

2. Proof of Theorem 1.1

Let $r \geq 2$ be a fixed number. We start with defining several special functions. First, let $H_r : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be given by

$$H_r(x, y) = (|y| - (r-1)|x|)(|x| + |y|)^{r-1}.$$

This is the famous function of Burkholder [3], who used it to establish the moment estimate (1.1). Next, let V_r, $U_r : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be given by $V_r(x, y) = 1 - r|x|$ and

$$U_r(x,y) = \begin{cases} \min\left\{H_r(x,y), V_r(x,y)\right\} & \text{if } |x| + |y| \leq 1, \\ V_r(x,y) & \text{if } |x| + |y| > 1. \end{cases}$$

Finally, introduce $F_r : [0,\infty) \to \mathbb{R}$ by

$$F_r(t) = \begin{cases} (r-1)^{r-1}t^r & \text{if } t \leq (r-1)^{-1}, \\ rt - 1 & \text{if } t > (r-1)^{-1}. \end{cases}$$

Lemma 2.1. *If* $|x| + |y| \leq 1$*, then* $U_r(x, y) \geq -F_r(|x|)$.

Proof. We may assume that $\mathbb{B} = \mathbb{R}$. It is easy to check that for a fixed x, the function $s \mapsto V(x, s)$ is decreasing on $[0,(r-2)|x|]$ and increasing on $[(r-2)|x|, \infty)$. This immediately gives the desired bound for $|x| \geq (r-1)^{-1}$, since then we have $(r-2)|x| \geq 1-|x|$ and hence $H_r(x, y) \geq H_r(x, 1-|x|) = V_r(x, y) = -F_r(|x|)$. So, suppose that $|x| < (r-1)^{-1}$: then $H_r(x, y) \geq H_r(x,(r-2)|x|) = -F_r(|x|)$ and $V_r(x, y) \geq -F_r(|x|)$; the latter estimate is equivalent to

$$(1-r)(1-|x|^{-1}) \geq 1 - (|x|^{-1})^{1-r},$$

which follows from the mean-value property for the function $t \mapsto t^{1-r}$, $t \geq 1$. □

Lemma 2.2. *Suppose that* $r = p + 1$*. Then*

$$\int_0^\infty F_r\left(\min\{t^{-1/p}, 1\}\right) dt \leq \frac{p+1}{p-1}p^p.$$

Proof. This is straightforward: the integral on the left equals

$$\begin{aligned}
&\int_0^1 F_r(1)\mathrm{d}t + \int_1^\infty F_r(t^{-1/p})\mathrm{d}t \\
&= F_r(1) + \int_1^{(r-1)^p} \big(rt^{-1/p} - 1\big)\mathrm{d}t + \int_{(r-1)^p}^\infty (r-1)^{r-1}t^{-r/p}\mathrm{d}t \\
&= r - 1 + \frac{rp}{p-1}\big((r-1)^{p-1} - 1\big) + 1 - (r-1)^p + \frac{p}{r-p}(r-1)^{p-1}.
\end{aligned}$$

Since $r = p+1$, the last two summands cancel out. It suffices to use the trivial bound $r \leq rp/(p-1)$ to get the claim. □

Now we will need auxiliary functions ϕ_r, $\psi_r : \mathcal{H}\times\mathcal{H} \to \mathcal{H}$, defined by

$$\phi_r(x,y) = \begin{cases} \frac{\partial U_r}{\partial x}(x,y) & \text{if the partial derivative exists,} \\ -rx' & \text{otherwise,} \end{cases}$$

and

$$\psi_r(x,y) = \begin{cases} \frac{\partial U_r}{\partial y}(x,y) & \text{if the partial derivative exists,} \\ 0 & \text{otherwise.} \end{cases}$$

Here we have used the notation $x' = x/|x|$ for $x \neq 0$, and $x' = 0$ for $x = 0$.

Lemma 2.3. *Suppose that x, y, h, $k \in \mathcal{H}$ satisfy $|h| \geq |k|$. Then*

$$U_r(x+h, y+k) \leq U_r(x,y) + \langle \phi_r(x,y), h\rangle + \langle \psi_r(x,y), k\rangle. \tag{2.1}$$

Proof. There is a well-known procedure of proving such estimates. Consider a function $G = G_{x,y,h,k} : \mathbb{R} \to \mathbb{R}$, given by $G(t) = U_r(x+th, y+tk)$; we will show that this function is concave. Observe that if we replace the function U_r with V_r or H_r, then the corresponding functions G^V and G^H do have the concavity property. Indeed, G^V is concave since V_r is concave on $\mathcal{H}\times\mathcal{H}$, while the concavity of G^H is due to Burkholder (see page 17 of [3]). Note that $T = \{t : |x+th| + |y+tk| \leq 1\}$ is a closed bounded interval, $G^V = G^H$ at the boundary of T and

$$G = \begin{cases} \min\{G^V, G^H\} & \text{on } T, \\ G^V & \text{outside } T. \end{cases}$$

Hence G is concave, and this in turn yields (2.1). Indeed: if G is differentiable at 0, we have $\langle \phi_r(x,y), h\rangle + \langle \psi_r(x,y), k\rangle = G'(0)$; if $G'(0)$ does not exist, then $x = 0$, $y = 0$ or $G(0) = G^H(0) = G^V(0)$, and in all the cases one easily checks that $\langle \phi_r(x,y), h\rangle + \langle \psi_r(x,y), k\rangle \in [G'_+(0), G'_-(0)]$. □

Lemma 2.4. *If g is differentially subordinate to f, then for any integer n we have*

$$\mathbb{E}U_r(f_n, g_n) \leq 0.$$

Proof. The previous lemma implies that the sequence $(U_r(f_n,g_n))_{n\geq 0}$ forms a supermartingale with respect to $(\mathcal{F}_n)_{n\geq 0}$. Indeed, for any $n \geq 0$ we have

$$\begin{aligned}
U_r(f_{n+1}, g_{n+1}) &= U_r(f_n + df_{n+1}, g_n + dg_{n+1}) \\
&\leq U_r(f_n, g_n) + \big\langle \phi_r(f_n,g_n), df_{n+1}\big\rangle + \big\langle \psi_r(f_n,g_n), dg_{n+1}\big\rangle.
\end{aligned}$$

Both sides are integrable, for there is an absolute constant κ such that $|U_r(x,y)| \leq \kappa(1+|x|)$ and $|\phi_r(x,y)| + |\psi(x,y)| \leq \kappa$ for all $(x,y) \in \mathcal{H}\times\mathcal{H}$. Taking expectation with respect to $\mathcal{F}_n$ yields the supermartingale property. In consequence,

$\mathbb{E}U_r(f_n, g_n) \leq \mathbb{E}U_r(f_0, g_0) \leq 0$, where the latter estimate holds even pointwise (apply the previous lemma to $x = y = 0$ and $h = f_0$, $k = g_0$). □

We are ready to establish the assertion of Theorem 1.1.

Proof of (1.4). By homogeneity, we may and do assume that $||f||_{p,\infty} \leq 1$. For any $\lambda > 0$, the martingale g/λ is differentially subordinate to f/λ. The bound $\mathbb{E}U_r(f_n/\lambda, g_n/\lambda) \leq 0$, guaranteed by Lemma 2.4, is equivalent to

$$\mathbb{P}(|f_n| + |g_n| \geq \lambda) \leq \mathbb{E}\left(\frac{r|f_n|}{\lambda}\right) 1_{\{|f_n|+|g_n|\geq\lambda\}} - \mathbb{E}U_r\left(\frac{f_n}{\lambda}, \frac{g_n}{\lambda}\right) 1_{\{|f_n|+|g_n|<\lambda\}}.$$

By Lemma 2.1, the right-hand side does not exceed

$$\begin{aligned}
\mathbb{E}\left(\frac{r|f_n|}{\lambda}\right)&1_{\{|f_n|+|g_n|\geq\lambda\}} + \mathbb{E}F_r\left(\frac{|f_n|}{\lambda}\right) 1_{\{|f_n|+|g_n|<\lambda\}}\\
&= \mathbb{E}\left[\frac{r|f_n|}{\lambda} - F_r\left(\frac{|f_n|}{\lambda}\right)\right] 1_{\{|f_n|+|g_n|\geq\lambda\}} + \mathbb{E}F_r\left(\frac{|f_n|}{\lambda}\right).
\end{aligned}$$

The inequality of Hardy and Littlewood [6] implies that for any nonnegative random variable X and any $A \in \mathcal{F}$ we have $\mathbb{E}X1_A \leq \int_0^{\mathbb{P}(A)} X^{\#}(t)\mathrm{d}t$, where

$$X^{\#}(t) = \inf\{s > 0 : \mathbb{P}(X > s) \leq t\}, \qquad t \in [0, 1],$$

is the nonincreasing rearrangement of X. Moreover, $G_r(t) := rt - F_r(t)$ is easily checked to be nondecreasing. Combining this with the bound above, we get that

$$\begin{aligned}
\mathbb{P}(|f_n| + |g_n| \geq \lambda) &\leq \int_0^{\mathbb{P}(|f_n|+|g_n|\geq\lambda)} G_r(|f_n|/\lambda)^{\#}(t)\mathrm{d}t + \mathbb{E}F_r\left(\frac{|f_n|}{\lambda}\right)\\
&= \int_0^{\mathbb{P}(|f_n|+|g_n|\geq\lambda)} G_r(|f_n|^{\#}(t)/\lambda)\mathrm{d}t + \int_0^1 F_r\left(\frac{|f_n|^{\#}(t)}{\lambda}\right)\mathrm{d}t\\
&= \frac{r}{\lambda}\int_0^{\mathbb{P}(|f_n|+|g_n|\geq\lambda)} |f_n|^{\#}(t)\mathrm{d}t + \int_{\mathbb{P}(|f_n|+|g_n|\geq\lambda)}^1 F_r\left(\frac{|f_n|^{\#}(t)}{\lambda}\right)\mathrm{d}t\\
&\leq \frac{r}{\lambda}\int_0^{\mathbb{P}(|f_n|+|g_n|\geq\lambda)} |f_n|^{\#}(t)\mathrm{d}t + \int_{\mathbb{P}(|f_n|\geq\lambda)}^1 F_r\left(\frac{|f_n|^{\#}(t)}{\lambda}\right)\mathrm{d}t.
\end{aligned}$$

Since $||f_n||_{p,\infty} \leq 1$, we have $|f_n|^{\#}(t) \leq t^{-1/p}$ for all $t \in (0, 1]$. Furthermore, by the definition of the nonincreasing rearrangement, we see that $|f_n|^{\#}(t) \leq \min\{t^{-1/p}, \lambda\}$ for $t \in [\mathbb{P}(|f_n| \geq \lambda), 1]$. These two observations yield

$$\mathbb{P}(|f_n| + |g_n| \geq \lambda) \leq \frac{r}{\lambda}\int_0^{\mathbb{P}(|f_n|+|g_n|\geq\lambda)} t^{-1/p}\mathrm{d}t + \int_0^1 F_r\left(\frac{\min\{t^{-1/p}, \lambda\}}{\lambda}\right)\mathrm{d}t.$$

Denote $c(\lambda) = \left(\lambda^p\mathbb{P}(|f_n| + |g_n| \geq \lambda)\right)^{1/p}$ and assume that $r = p + 1$. Multiply both sides of the estimate above by λ^p and derive the first integral on the right to obtain

$$\begin{aligned}
c(\lambda)^p &\leq \frac{rp}{p-1}c(\lambda)^{p-1} + \lambda^p\int_0^1 F_r\left(\frac{\min\{t^{-1/p}, \lambda\}}{\lambda}\right)\mathrm{d}t\\
&\leq \frac{p(p+1)}{p-1}c(\lambda)^{p-1} + \int_0^\infty F_r\left(\min\{t^{-1/p}, 1\}\right)\mathrm{d}t\\
&\leq \frac{p(p+1)}{p-1}c(\lambda)^{p-1} + \frac{p+1}{p-1}p^p,
\end{aligned}$$

where in the last line we have used Lemma 2.2. This implies $c(\lambda)\leq 2p(p+1)/(p-1)$, since the function $c\mapsto c^p-\frac{p(p+1)}{p-1}c^{p-1}$ is increasing on $[p+1,\infty)$ and its value at $c=2p(p+1)/(p-1)\in(p+1,\infty)$ equals

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

In other words, we have proved that

$$||g_n||_{p,\infty}\leq\big|\big||f_n|+|g_n|\big|\big|_{p,\infty}\leq\frac{2p(p+1)}{p-1},$$

and taking supremum over n yields the claim. □

On the order of the constant. The order $O(p)$ is optimal for $p\to\infty$ because of the bound $||f||_{p,\infty}\leq||f||_p$ and the sharpness of (1.3). The case $p\to1$ is more complicated, and we have decided to split the reasoning into two parts. Assume that the underlying probability space is the interval $[0,1)$ equipped with its Borel subsets and Lebesgue's measure. Let $1<p<2$ and $\varepsilon>0$ be fixed numbers.

Step 1. A martingale pair. Pick $q\in(p,2)$ and $\delta>0$. It is easy to verify that

$$(1+2\delta)^{p-1}(1+(2-q)\delta)<1+q\delta \tag{2.2}$$

provided δ is sufficiently small. Consider a Markov martingale (F,G) with the distribution determined by the conditions

(i) $(F_0,G_0)\equiv(0,0)$ and $(F_1,G_1)\in\{(-1,0),(1,0)\}$,
(ii) any state of the form $(x,0)$, $x\neq0$, leads to $(x(1+\delta),-\delta x)$ or to the point $((q-1)x/q,x/q)$,
(iii) any state of the form $(x(1+\delta),-\delta x)$, $x\neq0$, leads to $(x(1+2\delta),0)$ or to $((q-1)(x+2\delta)/q,(x+2\delta)/q)$,
(iv) the states not mentioned in (ii) and (iii) are absorbing.

It is easy to see that G is a transform of F by a deterministic sequence of numbers in $\{-1,0,1\}$. Now let N be an odd number, which will be specified in a moment. Directly from (i)–(iv), we have either $|F_N|=(q-1)|G_N|\in\big[1-q^{-1},(1+2\delta)^{(N-1)/2}\big)$ or $(F_N,G_N)=((1+2\delta)^{(N-1)/2},0)$. In addition,

$$\mathbb{P}\big((F_N,G_N)=\big((1+2\delta)^{(N-1)/2},0\big)\big)=\left(\frac{1+(2-q)\delta}{(1+2\delta)(1+q\delta)}\right)^{(N-1)/2}. \tag{2.3}$$

Thus $\lambda^p\mathbb{P}(|F_N|\geq\lambda)=0$ for $\lambda>(1+2\delta)^{(N-1)/2}$, while for $0<\lambda\leq(1+2\delta)^{(N-1)/2}$,

$$\begin{aligned}\lambda^p\,\mathbb{P}(|F_N|\geq\lambda)&\leq\lambda^p\,\mathbb{P}((q-1)|G_N|\geq\lambda)+\lambda^p\left(\frac{1+(2-q)\delta}{(1+2\delta)(1+q\delta)}\right)^{(N-1)/2}\\&\leq(q-1)^p||G_N||_{p,\infty}^p+\left(\frac{(1+2\delta)^{p-1}(1+(2-q)\delta)}{1+q\delta}\right)^{(N-1)/2}.\end{aligned}$$

By (2.2), if N is sufficiently large, then the last term is smaller than ε and hence

$$||F_N||_{p,\infty}^p\leq(q-1)^p||G_N||_{p,\infty}^p+\varepsilon. \tag{2.4}$$

Finally, note that $\mathbb{P}(G_N=0\text{ or }|G_N|\geq q^{-1})=1$, so by (2.3),

$$q^{-p}\mathbb{P}(|G_N|\geq q^{-1})=q^{-p}\left[1-\left(\frac{1+(2-q)\delta}{(1+2\delta)(1+q\delta)}\right)^{(N-1)/2}\right]>q^{-p}(1-\varepsilon),$$

which gives the bound

$$||G_N||_{p,\infty} \geq (1-\varepsilon)^{1/p}/q. \tag{2.5}$$

Step 2. A "portioning" argument. The idea is to split the probability space into a few parts and to use an appropriate time-shifted copy of (F, G) on each part. To be more precise, let K be a fixed integer. Define a pair (f, g) by $(f_0, g_0) \equiv (0, 0)$ and, for $\ell = 0, 1, \ldots, K-1$, $m = 1, 2, \ldots, N$ and $\omega \in [0, 1)$,

$$(f_{\ell N+m}, g_{\ell N+m})(\omega) = \begin{cases} \big(F_N(\{K\omega\}), G_N(\{K\omega\})\big) & \text{if } K\omega < \ell, \\ \big(F_m(\{K\omega\}), G_m(\{K\omega\})\big) & \text{if } K\omega \in [\ell, \ell+1), \\ (0,0) & \text{if } K\omega \geq \ell+1, \end{cases}$$

where $\{x\} = x - [x]$ denotes the fractional part of a number x. Then f, g are martingales with respect to the filtration generated by the pair (f, g), and g is a transform of f by a deterministic sequence with values in $\{-1, 0, 1\}$. Since $|f_{\ell N+m}(\omega)| \leq |F_N(\{K\omega\})|$ outside $[\ell/K, (\ell+1)/K)$, the inequality (2.4) implies

$$\begin{aligned} \lambda^p\, \mathbb{P}(|f_{\ell N+m}| \geq \lambda) &\leq \frac{K-1}{K}\, \lambda^p\, \mathbb{P}(|F_N| \geq \lambda) + \frac{1}{K}\, \lambda^p\, \mathbb{P}(|F_m| \geq \lambda) \\ &\leq (q-1)^p ||G_N||_{p,\infty}^p + \varepsilon + \frac{1}{K} \sup_{0 \leq n \leq N} ||F_n||_{p,\infty}^p. \end{aligned}$$

Now, if K is sufficiently large, then $\sup_{0\leq n\leq N} ||F_n||_{p,\infty}^p/K \leq \varepsilon ||G_N||_{p,\infty}^p$; furthermore, we have $\varepsilon \leq \varepsilon q^p ||G_N||_{q,\infty}^p/(1-\varepsilon)$ by virtue of (2.5). Since $||G_N||_{p,\infty} = ||g_{KN}||_{p,\infty} \leq ||g||_{p,\infty}$, we have proved that

$$||f_{\ell N+m}||_{p,\infty}^p \leq \left[(q-1)^p + \frac{\varepsilon q^p}{1-\varepsilon} + \varepsilon\right] ||g||_{p,\infty}^p.$$

But ε, $q \in (p, 2)$, ℓ and m were arbitrary. Thus, for $1 < p < 2$ the best constant in (1.4) is at least $(p-1)^{-1}$ and its order $O((p-1)^{-1})$ is optimal as $p \to 1$. $\square$

3. Weak-type inequalities and the geometry of $\mathbb{B}$

3.1. **Proof of Theorem 1.2.** We split the reasoning into three parts.

Step 1. A good-λ inequality. We use the extrapolation method of Burkholder and Gundy [5]. Recall that a $\mathbb{B}$-valued martingale is called *dyadic* if $df_0 \equiv b_0$ for some $b_0 \in \mathbb{B}$ and for any $n \geq 1$ and any nonempty set of the form $\{df_1 = b_1, df_2 = b_2, \ldots, df_{n-1} = b_{n-1}\}$, the restriction of df_n to this set either vanishes identically or has its values in $\{-b_n, b_n\}$ for some $b_n \in \mathbb{B} \setminus \{0\}$. Let $\delta > 0$, $\beta > 2\delta + 1$ be numbers to be specified later, and put $\alpha = (2\delta C/(\beta - 2\delta - 1))^p$. We will show that if f, g are $\mathbb{B}$-valued dyadic martingales and g is differentially subordinate to f, then

$$\mathbb{P}(g_n^* > \beta\lambda,\ f_n^* \leq \delta\lambda) \leq \alpha \mathbb{P}(g_n^* > \lambda) \tag{3.1}$$

for any integer n and any $\lambda > 0$. To do this, define

$$\begin{aligned} \mu &= \inf\{k : |g_k| > \lambda\}, \\ \nu &= \inf\{k : |g_k| > \beta\lambda\}, \\ \sigma &= \inf\{k : |f_k| > \delta\lambda \text{ or } |df_{k+1}| > 2\delta\lambda\}. \end{aligned}$$

Then μ, ν and σ are adapted stopping times: this is obvious for the first two variables, and to deal with the third one, observe that $|df_{k+1}|$ is $\mathcal{F}_k$-measurable since f is dyadic. Introduce the sequence $w = \big(1_{\{\mu<k\leq\nu\wedge\sigma\}}\big)_{k=0}^{\infty}$. Clearly, it is

predictable, so $(w_k df_k)_{k=0}^\infty$ and $(w_k dg_k)_{k=0}^\infty$ are martingale difference sequences. Denoting the corresponding martingales by F and G, we see that G is differentially subordinate to F and $|dG_k| \leq |dF_k| \leq 2\delta\lambda$ almost surely for all k, so

$$\begin{aligned}\mathbb{P}(g_n^* > \beta\lambda,\, f_n^* \leq \delta\lambda) &\leq \mathbb{P}(\mu \leq \nu \leq n,\, \sigma \geq n).\\ &\leq \mathbb{P}(|G_n| > \beta\lambda - 2\delta\lambda - \lambda)\\ &\leq C^p \frac{||F||_{p,\infty}^p}{(\beta\lambda - 2\delta\lambda - \lambda)^p},\end{aligned} \tag{3.2}$$

where in the latter passage we have exploited (1.5). It is not difficult to see that for any $0 \leq m \leq n$, $|F_m| \leq 2\delta\lambda 1_{\{\mu<n\}} \leq 2\delta\lambda 1_{\{g_n^*>\lambda\}}$, which gives $||F||_{p,\infty}^p \leq (2\delta\lambda)^p \mathbb{P}(g_n^* > \lambda)$. Plugging this into (3.2) yields (3.1).

Step 2. An L^2-estimate. The good-λ inequality (3.1) implies that

$$\mathbb{P}(g_n^* > \beta\lambda) \leq \alpha \mathbb{P}(g_n^* > \lambda) + \mathbb{P}(f_n^* > \delta\lambda),$$

and thus, by the standard integration argument,

$$\beta^{-2}||g_n^*||_2^2 \leq \alpha ||g_n^*||_2^2 + \delta^{-2}||f_n^*||_2^2,$$

or $(1-\alpha\beta^2)||g_n^*||_2^2 \leq \beta^2\delta^{-2}||f_n^*||_2^2$. If δ is suitably small, then $1-\alpha\beta^2$ is positive and we obtain, by Doob's maximal inequality,

$$||g_n||_2^2 \leq ||g_n^*||_2^2 \leq \frac{\beta^2\delta^{-2}}{1-\alpha\beta^2}||f_n^*||_2^2 \leq \frac{4\beta^2\delta^{-2}}{1-\alpha\beta^2}||f_n||_2^2. \tag{3.3}$$

Step 3. Kwapień's characterization theorem. Let $(h_k)_{k\geq 0}$ be the Haar system on $[0,1)$, $(a_k)_{k\geq 0}$ be a sequence in $\mathbb{B}$, and pick $b \in \mathbb{B}$ with $|b| = 1$. Consider the martingale difference sequences $(a_k h_k)_{k\geq 0}$ and $(b|a_k|h_k)_{k\geq 0}$ (on the probability space $([0,1), \mathcal{B}([0,1)), |\cdot|)$, equipped with the filtration generated by $(h_k)_{k\geq 0}$). Then the associated martingales f, g are dyadic and differentially subordinate to each other. Hence, by (3.3), for any integer n we have

$$\left|\left|\sum_{k=0}^n a_k h_k\right|\right|_2 \approx \left|\left|\sum_{k=0}^n b|a_k|h_k\right|\right|_2 = \left|\left|\sum_{k=0}^n |a_k|h_k\right|\right|_2 = \left(\sum_{k=0}^n |a_k|^2 ||h_k||_2^2\right)^{1/2}.$$

Now, let $(r_k)_{k\geq 0}$ be the sequence of Rademacher functions on $[0,1)$ and let $(c_k)_{k\geq 0}$ be a sequence in $\mathbb{B}$. We have $r_k = h_{2^k} + h_{2^k+1} + \ldots + h_{2^{k+1}-1}$, so putting $a_0 = 0$, $a_{2^k+\ell} = c_k$ for all k and $0 \leq \ell < 2^k$, and using the above bound with $n = 2^m$, yields

$$\left|\left|\sum_{k=0}^m c_k r_k\right|\right|_2 \approx \left(\sum_{k=0}^m |c_k|^2\right)^{1/2}.$$

Thus, by the result of Kwapień [7], $\mathbb{B}$ is isomorphic to a Hilbert space.

3.2. **Proof of Theorem 1.3.** (i)⇒(ii) Again, we will use the extrapolation method. Let $K = K(\mathbb{B})$ be the constant coming from the definition of UMD. We have $K \geq K(\mathbb{R}) = 2$ (see [2]). Let

$$\delta = \frac{p^p}{K(4(p+1)^{p+1} + p^{p+1})}, \quad \alpha = \frac{4\delta K}{\beta - 2\delta - 1} \quad \text{and} \quad \beta = 1 + \frac{1}{p}.$$

Arguing as above, we see that if f is a $\mathcal{B}$-valued dyadic martingale and g is its transform by a predictable sequence with values in $[-1,1]$, then for any λ and n,

$$\mathbb{P}(g_n^* > \beta\lambda, f_n^* \leq \delta\lambda) \leq \alpha \mathbb{P}(g_n^* > \lambda).$$

In consequence,

$$\lambda^p\,\mathbb{P}(g_n^*\geq\beta\lambda)\leq\alpha\lambda^p\,\mathbb{P}(g_n^*>\lambda)+\lambda^p\,\mathbb{P}(f_n^*>\delta\lambda)\leq\alpha||g_n^*||_{p,\infty}^p+\delta^{-p}||f_n^*||_{p,\infty}^p.$$

Taking supremum over λ gives $\beta^{-p}||g_n^*||_{p,\infty}^p\leq\alpha||g_n^*||_{p,\infty}^p+\delta^{-p}||f_n^*||_{p,\infty}^p$, or

$$||g_n^*||_{p,\infty}\leq\frac{\beta}{\delta}(1-\alpha\beta^p)^{1/p}||f_n^*||_{p,\infty}\leq 68Kp||f_n^*||_{p,\infty}.$$

Here we have used the estimates

$$\frac{\beta}{\delta}=K(p+1)\left[4\left(\frac{p+1}{p}\right)^{p+1}+1\right]\leq 34Kp\quad\text{and}\quad(1-\alpha\beta^p)^{-1/p}\leq(p+1)^{1/p}\leq 2.$$

It remains to use the trivial bound $||g_n||_{p,\infty}\leq||g_n^*||_{p,\infty}$ and the Doob-type estimate $||f_n^*||_{p,\infty}\leq\frac{p}{p-1}||f_n||_{p,\infty}$ (see Nikolidakis [9]), to get (1.6) for dyadic martingales. To pass to general martingales, repeat word-by-word the argument of Maurey [8].

(ii)⇒(iii) This is trivial.

(iii)⇒(i) Let f be a bounded nonzero martingale and g is its transform by a deterministic sequence with values in $\{-1,1\}$. Let $\lambda=2\beta/||f||_\infty$ and introduce the stopping time $\tau=\inf\{n:|g_n|\geq\lambda\}$. The stopped martingale $g^\tau=(g_{\tau\wedge n})_{n\geq 0}$ is a transform of $f^\tau=(f_{\tau\wedge n})_{n\geq 0}$ by the same sequence as previously and $||f^\tau||_{p,\infty}\leq||f^\tau||_\infty\leq||f||_\infty$. Therefore, by (1.7),

$$\mathbb{P}(g^*\geq 2\lambda)\leq\lim_{n\to\infty}\mathbb{P}(|g_n^\tau|\geq\lambda)\leq\frac{\beta}{\lambda}||f^\tau||_{p,\infty}\leq\frac{1}{2}.$$

Thus, we have proved that the inequality $||f||_\infty<\infty$ implies $\mathbb{P}(g^*<\infty)>0$. By Theorem 3.2 of Burkholder [1], $\mathbb{B}$ is a UMD space.

References

[1] Burkholder, D. L. (1981). Martingale transforms and the geometry of Banach spaces, Probability in Banach Spaces, III (Medford, 1980), Lecture Notes in Math. 860, Springer, Berlin, 35–50.

[2] Burkholder, D. L. (1984). Boundary value problems and sharp inequalities for martinale transforms. Ann. Probab. 12, 647–702.

[3] Burkholder, D. L. (1991). Extrapolations in martingale theory and its applications. École d'Été de Probabilités de Saint-Flour XIX - 1989, Lecture Notes in Math. 1464, Springer, Berlin, 1–66.

[4] Burkholder, D. L. (2001). Martingales and singular integrals in Banach spaces. Handbook of the Geometry of Banach Spaces, Vol. 1, pp. 233-269.

[5] Burkholder, D. L. and Gundy, R. F. (1970). Extrapolation and interpolation of quasi-linear operators on martingales. Acta Math. 124, 249–304.

[6] Hardy, G. H., Littlewood, J. E. and Pólya, G. (1952). Inequalities. 2nd ed., Cambridge University Press, Cambridge, UK, 1952.

[7] Kwapień, S. (1972). Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. Studia Math. 44, 583–595.

[8] Maurey, B. (1975). Système de Haar. Séminaire Maurey-Schwartz, 1974–1975, École Polytechnique, Paris.

[9] Nikolidakis, E. N. (2010). Extremal problems related to maximal dyadic-like operators. J. Math. Anal. Appl. 369, 377-385.

[10] Suh, Y. (2005). A sharp weak type (p,p) inequality $(p>2)$ for martingale transforms and other subordinate martingales. Trans. Amer. Math. Soc. 357, 1545–1564.

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

E-mail address: ados@mimuw.edu.pl