PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 18, Fasc. 1 (1998), pp. 149–171

OPERATORS ON MARTINGALES, Φ-SUMMING OPERATORS, AND THE CONTRACTION PRINCIPLE

BY

STEFAN GEISS (JENA)

Abstract. For the absolutely Φ-summing operators $T: X \to Y$ between Banach spaces X and Y we consider martingale inequalities of the type

$$\Big\| \sup_{1 \leqslant k \leqslant N} \Big\| \sum_{l=1}^{k} T d_l \Big\|_Y \Big\|_{L_2} \leqslant c \Big\| \sup_{i=1,2,\ldots} \Big(\sum_{k=1}^{N} |\langle d_k, a_i \rangle|^2 \Big)^{1/2} \Big\|_{L_2},$$

where $(d_k)_{k=0}^{N} \subset L_1^X(\Omega, \mathscr{F}, P)$ is a martingale difference sequence and $(a_i)_{i=1}^{\infty}$ is a sequence of normalized functionals on X, and we show that these inequalities are useful in different directions. For example, for a Banach space X, $x_1, \ldots, x_n \in X$, independent standard Gaussian variables $g_1, \ldots, g_n$, and $1 \leqslant r < \infty$ we deduce that

$$\Big\| \sum_{i=1}^{n} \Big[\sum_{k=\tau_{i-1}+1}^{\tau_i} d_k \Big] x_i \Big\|_{L_r^X} \leqslant c \sqrt{r} \Big\| \sup_{1 \leqslant i \leqslant n} S_2({}^{\tau_{i-1}}f^{\tau_i}) \Big\|_{L_r} \Big\| \sum_{i=1}^{n} g_i x_i \Big\|_{L_1^X},$$

where $f = (d_k)_{k=0}^{N}$ is a scalar-valued martingale difference sequence such that $(|d_k|)_{k=1}^{N}$ is predictable, $0 = \tau_0 \leqslant \tau_1 \leqslant \ldots \leqslant \tau_n = N$ is a sequence of stopping times, and

$$S_2({}^{\tau_{i-1}}f^{\tau_i}) := \Big(\sum_{k=\tau_{i-1}+1}^{\tau_i} |d_k|^2 \Big)^{1/2}.$$

Introduction. There are several reasons to extend inequalities involving operators defined on martingales from the scalar-valued setting to the Banach space valued setting. For example, one possible variant of the Burkholder–Davis–Gundy inequality in the vector-valued setting is

$$(1) \qquad \Big\| \sup_{1 \leqslant k \leqslant N} \Big\| \sum_{l=1}^{k} d_l \Big\|_X \Big\|_{L_2} \leqslant c \Big\| \Big(\sum_{k=1}^{N} \|d_k\|_X^2 \Big)^{1/2} \Big\|_{L_2},$$

where X is a Banach space and $(d_k)_{k=0}^{N} \subset L_1^X(\Omega, \mathscr{F}, P)$ is a martingale difference sequence. This inequality can be used to characterize and to handle those Banach spaces X which admit renorming with the modulus of smoothness of power type 2 (see [29]). There is also another way to consider a vector-valued

Burkholder–Davis-Gundy inequality. Instead of (1) we take a bounded and linear operator $T: X \to Y$ between Banach spaces X and Y and regard

$$(2) \qquad \Big\| \sup_{1 \leqslant k \leqslant N} \Big\| \sum_{l=1}^{k} T d_l \Big\|_Y \Big\|_{L_2} \leqslant c \Big\| \sup_{i=1,2,\ldots} \Big(\sum_{k=1}^{N} |\langle d_k, a_i \rangle|^2 \Big)^{1/2} \Big\|_{L_2},$$

where $(a_i)_{i=1}^{\infty}$ is some normalized sequence of linear functionals. First of all, the consideration of inequality (2) requires the usage of operators T since the validity of (2) for all $N = 1, 2, \ldots$ for an identity $T = I_X$ of a Banach space X implies $\dim(X) < \infty$ in general.

The subject of the paper is to show that inequalities of type (2) are useful in different situations and to develop a general approach for such inequalities.

The paper is organized as follows. In Section 1 we recall some facts about the absolutely Φ-summing operators. These operators are used to state in Theorem 3.2 the basic result of the paper, which is an abstract version of (2). Since the BMO_ψ-L_∞ estimates, the starting point of Theorem 3.2, are based on Lorentz norms, whereas the notion of absolutely Φ-summing operators is based on Orlicz norms, we show in Section 2 that the BMO_ψ-spaces have a representation by Orlicz norms. Besides the applications of Theorem 3.2 given in Section 3 we derive in Section 4 contraction principles for vector-valued Gaussian random variables. A corresponding contraction principle for Rademacher variables is proved in Section 5 by using a different technique.

Throughout this paper $[\Omega, \mathscr{F}, P]$ stands for a probability space, and $(\mathscr{F}_k)_{k=0}^N$ for a filtration with $\mathscr{F}_k \subseteq \mathscr{F}$ and $\mathscr{F}_0 = \{\emptyset, \Omega\}$. All random variables and Banach spaces are assumed to be real. By *standard Gaussian random variables* we mean symmetric random variables distributed like $\mathcal{N}(0, 1)$. A random variable $\varepsilon \in L_2(M, \mu)$ is called a *Rademacher variable* if $\mu(\varepsilon = 1) = \mu(\varepsilon = -1) = 1/2$. The *Haar functions* $(h_k)_{k=0}^{\infty} \subset L_1[0, 1)$ are given by

$$h_0 = 1, h_1 = \chi_{[0,1/2)} - \chi_{[1/2,1)}, h_2 = \chi_{[0,1/4)} - \chi_{[1/4,1/2)}, h_3 = \chi_{[1/2,3/4)} - \chi_{[3/4,1)}, \ldots,$$

where $\mathscr{F}_k^h := \sigma(h_0, \ldots, h_k)$. Given a Banach space X its dual is denoted by X', and its closed unit ball by B_X. Moreover, $L_0^X(\Omega, \mathscr{F}, P)$ is the space of all Borel measurable $h: \Omega \to X$ such that there is a separable and closed subspace $X_0 \subseteq X$ with $P(h \in X_0) = 1$, where $L_0(\Omega, \mathscr{F}, P) = L_0^{\mathbb{R}}(\Omega, \mathscr{F}, P)$. The symbol $\mathscr{L}(X, Y)$ stands for the linear and continuous operators $T: X \to Y$ between the Banach spaces X and Y equipped with the operator norm $\|T\| := \sup\{\|Tx\|_Y : x \in B_X\}$. Given quantities $\|\cdot\|$ and $|||\cdot|||$ we use

$$\|\cdot\| \sim_c |||\cdot||| \qquad \text{for } c^{-1}\|\cdot\| \leqslant |||\cdot||| \leqslant c\|\cdot\|.$$

1. Absolutely Φ-summing operators. The introduction of the absolutely Φ-summing operators, where Φ is an exponential Young function, was motivated by the consideration of majorizing measures for Gaussian processes (cf. Corollary 3.11). The results of this section are folklore.

DEFINITION 1.1. (1) A *Young function* Φ: $[0, \infty) \to [0, \infty)$, that means an increasing and convex bijection, is said to be *sup-multiplicative* if there is some $c > 0$ such that $\Phi(\lambda)\Phi(\mu) \leqslant \Phi(c\lambda\mu)$ for all $\lambda, \mu \geqslant 1$. We write $\Phi \in \mathscr{Y}_{\sup}$ and let $\Delta_{\sup}(\Phi) := \inf c$.

(2) Given a Young function Φ, the space $L_\Phi^X(\Omega, \mathscr{F}, P)$ consists of all $h \in L_0^X(\Omega, \mathscr{F}, P)$ with

$$\|h\|_{L_\Phi^X} := \inf\left\{c > 0 \,\middle|\, E\Phi\left(\frac{\|h\|_X}{c}\right) \leqslant 1\right\} < \infty,$$

where $L_\Phi(\Omega, \mathscr{F}, P) := L_\Phi^{\mathbf{R}}(\Omega, \mathscr{F}, P)$.

(3) For $\Phi \in \mathscr{Y}_{\sup}$ an operator $T \in \mathscr{L}(X, Y)$ is *absolutely Φ-summing* if there is a constant $c > 0$ such that for all probability spaces $[\Omega, \mathscr{F}, P]$ and all $h \in L_0^X(\Omega, \mathscr{F}, P)$

$$\|Th\|_{L_\Phi^Y} \leqslant c \sup_{a \in B_{X'}} \|\langle h, a\rangle\|_{L_\Phi^Y}. \tag{3}$$

We write $T \in \Pi_\Phi(X, Y)$ and let $\pi_\Phi(T) := \inf c$.

In particular, we use $\Phi_q(\lambda) := \exp\{\lambda^q\} - 1 \in \mathscr{Y}_{\sup}$ for $1 \leqslant q < \infty$. The absolutely Φ-summing operators form a Banach operator ideal in the sense of [27]. In the case $L_\Phi = L_p$ we obtain the absolutely p-summing operators $\Pi_p(X, Y)$. We restrict ourselves to the sup-multiplicative Young functions for two reasons. First, according to (4) and Lemma 2.2 this case is of only interest in our situation. Secondly, this condition on Φ ensures that the typical absolutely Φ-summing operators are the embeddings $C(K) \to L_\Phi(K, \mu)$, where K is a compact Hausdorff space and μ a normalized Borel measure (see Theorem 1.2 and Remark 1.5 (1)). From this latter fact one can deduce $\Pi_\Phi(X, Y) \subseteq \Pi_\Psi(X, Y)$ if and only if $L_\Psi[0, 1] \subseteq L_\Phi[0, 1]$. Let us start with the basic example of an absolutely Φ-summing operator.

THEOREM 1.2. *For $\Phi \in \mathscr{Y}_{\sup}$, a compact Hausdorff space K, and a normalized Borel measure μ on K, we have for the embedding I: $C(K) \to L_\Phi(K, \mu)$*

$$\pi_\Phi(I) \leqslant \big(1 + \Phi(1)\big)^2 \Delta_{\sup}(\Phi).$$

Proof. We use standard arguments from the theory of the Orlicz spaces which can be exploited to prove Fubini type theorems. The only point is that we do not assume the sup-multiplicativity of Φ for all $\lambda, \mu \geqslant 0$.

(1) For $g \in L_\Phi(K, \mu)$ with $\|g\|_{L_\Phi(K,\mu)} > c_0 := \big(1 + \Phi(1)\big)\Delta_{\sup}(\Phi)$ we show

$$\Phi\left(\frac{\|g\|_{L_\Phi(K,\mu)}}{c_0}\right) \leqslant \int_K \Phi(|g|)\, d\mu.$$

Indeed, by convexity,

$$\Phi\left(\frac{\nu}{\lambda}\right) \leqslant \frac{\Phi(\nu)}{\lambda} \qquad (\nu \geqslant 0,\ \lambda \geqslant 1)$$

so that for $1 < b < \|g\|_{L_\Phi(K,\mu)}/c_0$ we get

$$1 < \int_K \Phi\left(\frac{|g|}{bc_0}\right) d\mu \leqslant \frac{1}{1+\Phi(1)} \int_K \Phi\left(\frac{|g|}{b\Delta_{\sup}(\Phi)}\right) d\mu$$
$$\leqslant \frac{1}{1+\Phi(1)} \left[\int_{|g| \geqslant b\Delta_{\sup}(\Phi)} \Phi\left(\frac{|g|}{b\Delta_{\sup}(\Phi)}\right) d\mu + \Phi(1)\right]$$
$$\leqslant \frac{1}{1+\Phi(1)} \left[\frac{1}{\Phi(b)} \int_K \Phi(|g|)\, d\mu + \Phi(1)\right].$$

(2) Now let $h \in L_0^{C(K)}(\Omega, \mathscr{F}, P)$ be a step function taking a finite number of values (see Remark 1.5 (2) below). For

$$\Omega' := \{\|h\|_{L_\Phi(K,\mu)} > c_0\} \subseteq \Omega \quad \text{and} \quad c_0(1+\Phi(1)) < \big\|\|h\|_{L_\Phi(K,\mu)}\big\|_{L_\Phi(\Omega,P)}$$

we deduce with the help of the first step

$$1 < \int_\Omega \Phi\left(\frac{\|h(\omega)\|_{L_\Phi(K,\mu)}}{c_0(1+\Phi(1))}\right) dP(\omega) \leqslant \frac{1}{1+\Phi(1)} \left[\int_{\Omega'} \Phi\left(\frac{\|h(\omega)\|_{L_\Phi(K,\mu)}}{c_0}\right) dP(\omega) + \Phi(1)\right]$$
$$\leqslant \frac{1}{1+\Phi(1)} \left[\int_{\Omega' \times K} \Phi(|\langle h(\omega), \delta_a\rangle|)\, d(\mu \times P)(a, \omega) + \Phi(1)\right]$$
$$\leqslant \frac{1}{1+\Phi(1)} \left[\sup_{a\in K} \int_\Omega \Phi(|\langle h, \delta_a\rangle|)\, dP + \Phi(1)\right]$$

and $1 < \sup_{a\in K} \int_\Omega \Phi(|\langle h, \delta_a\rangle|)\, dP$. ∎

To obtain a special case of Theorem 1.2 we need

LEMMA 1.3. *Let* $1 \leqslant q < \infty$, $K := \{1, 2, \ldots\}$, *and*

$$\mu := \sum_{i=1}^{\infty} \frac{1}{i(i+1)} \delta_{\{i\}}.$$

Then

$$\sup_{i=1,2,\ldots} \frac{|\alpha_i|}{\sqrt[q]{\log(i+1)}} \sim_c \|(\alpha_i)_{i=1}^{\infty}\|_{L_{\Phi_q}(K,\mu)},$$

where $c > 0$ *depends on* q *only.*

Proof. If we have

$$\sup_{i=1,2,\ldots} \frac{|\alpha_i|}{\sqrt[q]{\log(i+1)}} > 1,$$

then we get some i_0 with $|\alpha_{i_0}| > \sqrt[q]{\log(i_0+1)}$ and

$$\|(\alpha_i)_{i=1}^{\infty}\|_{L_{\Phi_q}(K,\mu)} > \left\|\chi_{\left[0, \frac{1}{i_0(i_0+1)}\right]} \sqrt[q]{\log(i_0+1)}\right\|_{L_{\Phi_q}[0,1]}$$
$$= \frac{\sqrt[q]{\log(i_0+1)}}{\sqrt[q]{\log(i_0(i_0+1)+1)}} \geqslant \frac{1}{\sqrt[q]{2}},$$

so that

$$\sup_{i=1,2,\dots} \frac{|\alpha_i|}{\sqrt[q]{1+\log i}} \leqslant \sqrt[q]{2}\,\|(\alpha_i)_{i=1}^{\infty}\|_{L_{\Phi_q}(K,\mu)}.$$

The remaining inequality is left to the reader. ∎

COROLLARY 1.4. *For* $1 \leqslant q < \infty$ *we have*

$$D_q \in \Pi_{\Phi_q}(l_\infty, l_\infty) \quad \textit{if } D_q\big((\xi_i)_{i=1}^{\infty}\big) := \left(\frac{\xi_i}{\sqrt[q]{\log(i+1)}}\right)_{i=1}^{\infty}.$$

Proof. Either we go a direct way or we use Theorem 1.2. For the latter we observe that it is sufficient to show $D_q \in \Pi_{\Phi_q}(\mathscr{C}, \mathscr{C})$, where $\mathscr{C}$ is the space of convergent sequences. Since $\mathscr{C} = C(K)$ in a canonical way, where $K = \{1, 2, \dots, \infty\}$ is equipped with the metric $d(k, l) = |1/k - 1/l|$, we apply Theorem 1.2 and Lemma 1.3. ∎

Remark 1.5. (1) Theorem 1.2 is a part of the basic characterization of the absolutely Φ-summing operators: For $\Phi \in \mathscr{Y}_{\sup}$ an operator $T \in \mathscr{L}(X, Y)$ is absolutely Φ-summing if and only if T can be factorized through a restriction of an embedding $C(K) \to L_\Phi(K, \mu)$ like in Theorem 1.2. We have seen the "if" part; the "only if" part follows from the corresponding result of Assouad [1] about the Φ-0-summing operators. In particular, it turns out that the Φ-0-summing and the absolutely Φ-summing operators coincide whenever $\Phi \in \mathscr{Y}_{\sup}$.

(2) There are some straightforward reductions in Definition 1.1 (3). First we have for a norming sequence $(a_i)_{i=1}^{\infty} \subset B_{X'}$, which means $\|x\|_X = \sup_{i=1,2,\dots} |\langle x, a_i\rangle|$ for all $x \in X$,

$$\sup_{a \in B_{X'}} \|\langle h, a\rangle\|_{L_\Phi} = \sup_{i=1,2,\dots} \|\langle h, a_i\rangle\|_{L_\Phi}.$$

Secondly, it is enough to consider in inequality (3) step functions h taking a finite number of values.

2. BMO_ψ-spaces

DEFINITION 2.1. (1) Let $\mathscr{D}$ be the set of all increasing bijections

$$\psi\colon [1, \infty) \to [1, \infty)$$

and let $\overline{\mathscr{D}}$ be the subset of those $\psi \in \mathscr{D}$ for which

$$\psi(\lambda+\mu)+1 \geqslant \psi(\lambda)+\psi(\mu) \quad \text{for } \lambda, \mu \geqslant 1.$$

(2) For $\psi \in \mathscr{D}$ the *Lorentz space* $M_\psi(\Omega, \mathscr{F}, P)$ consists of all $h \in L_0(\Omega, \mathscr{F}, P)$ with

$$|h|_{M_\psi} := \inf\{c > 0 \mid P(|h| > \lambda) \leqslant \exp\{1 - \psi(\lambda/c)\} \text{ for } \lambda \geqslant c\} < \infty.$$

(3) Let $\psi \in \overline{\mathscr{D}}$ and $(f_k)_{k=0}^{N} \subset L_0(\Omega, \mathscr{F}, P)$ be adapted to $(\mathscr{F}_k)_{k=0}^{N}$. Then

$$\|(f_k)_{k=0}^{N}\|_{BMO_\psi} := \sup_{0 \leqslant k \leqslant l \leqslant N} \sup_{\substack{C \in \mathscr{F}_k \\ P(C) > 0}} |f_l - f_{k-1}|_{M_\psi(C, P_C)},$$

where $f_{-1} := 0$ and $P_C := P/P(C)$ is the normalized restriction of P to C.

In view of [13] (Theorem 4.6 and Lemma 4.4 (1)) the restriction to the subset $\overline{\mathscr{D}}$ of $\mathscr{D}$ in the definition of the above BMO-spaces is of no loss of generality. The typical examples for elements of $\overline{\mathscr{D}}$ are given by $\psi_q(\lambda) := \lambda^q$ for $1 \leqslant q < \infty$. Lemma 4.4 (2) of [13] implies

$$\text{(4)} \qquad \sup_{a>1} \inf_{\lambda \geqslant 1} \frac{\psi(a\lambda)}{\psi(\lambda)} > 1 \quad \text{whenever } \psi \in \overline{\mathscr{D}},$$

so that the next lemma shows that the BMO_ψ-spaces have a representation with the help of Orlicz norms. This gives the link between the BMO_ψ-spaces and the absolutely Φ-summing operators, which is behind Theorem 3.2. This also complements [13] (Remark 4.14) where some relations between the BMO-definition in [2], which uses Orlicz norms, and our BMO-definition are outlined.

LEMMA 2.2. *For* $\psi \in \mathscr{D}$ *with* $\sup_{a>1} \inf_{\lambda \geqslant 1} [\psi(a\lambda)/\psi(\lambda)] > 1$ *there are* $\Phi \in \mathscr{Y}_{\sup}$ *and* $c \geqslant 1$ *such that*

$$|\cdot|_{M_\psi} \sim_c \|\cdot\|_{L_\Phi}.$$

Proof. We extend ψ to $\psi: [0, \infty) \to [0, \infty)$ by $\psi(\lambda) = \lambda$ for $0 \leqslant \lambda \leqslant 1$ and find $a > 1$ and $\varepsilon > 0$ such that $\psi(\lambda)(1+\varepsilon) \leqslant \psi(a\lambda)$ for $\lambda \geqslant 0$. Choosing $0 < p < \infty$ such that $a^{1/p} = 1+\varepsilon$ we get, for $\mu \geqslant 1$ with $a^n \leqslant \mu \leqslant a^{n+1}$, $n \in \{0, 1, 2, \ldots\}$, and $\lambda \geqslant 0$,

$$\psi(\mu\lambda) \geqslant \psi(a^n\lambda) \geqslant (1+\varepsilon)^n \psi(\lambda) = \frac{1}{1+\varepsilon}(a^{n+1})^{1/p}\psi(\lambda) \geqslant \frac{1}{1+\varepsilon}\mu^{1/p}\psi(\lambda).$$

For $s \geqslant 0$ and $t \geqslant 1/\sqrt[p]{a}$ this gives

$$\text{(5)} \qquad \psi^{-1}(ts) \leqslant at^p \psi^{-1}(s).$$

Setting $\Phi_0(\lambda) := e^{\psi(\lambda)} - 1$ for $\lambda \geqslant 0$ and observing that $\Phi_0^{-1}(t) = \psi^{-1}(\log(t+1))$ we see that inequality (5) implies

$$\sup_{s_0 \leqslant s \leqslant t} \frac{s}{t} \frac{\Phi_0^{-1}(t)}{\Phi_0^{-1}(s)} \leqslant \sup_{0 < s \leqslant t} \left(\frac{\log(s+1)}{\log(t+1)}\right)^p \frac{\Phi_0^{-1}(t)}{\Phi_0^{-1}(s)} \leqslant a,$$

where $s_0 > 0$ depends on p. On the other hand, assuming that $s_1 := e-1 < s_0$ we have

$$\sup_{0 < s \leqslant t \leqslant s_1} \frac{s}{t} \frac{\Phi_0^{-1}(t)}{\Phi_0^{-1}(s)} \leqslant 1 \quad \text{and} \quad \sup_{s_1 \leqslant s \leqslant t \leqslant s_0} \frac{s}{t} \frac{\Phi_0^{-1}(t)}{\Phi_0^{-1}(s)} = b < \infty,$$

so that for $c = ab$

$$\frac{\Phi_0^{-1}(t)}{t} \leqslant c\, \frac{\Phi_0^{-1}(s)}{s} \quad \text{for } 0 < s < t < \infty.$$

Putting $h(t) := \inf_{s>0}(1+cts^{-1})\Phi_0^{-1}(s)$ we obtain a concave $h: [0, \infty) \to [0, \infty)$ satisfying $h(0) = 0$ and

$$\frac{1}{c+1} h(t) \leqslant \Phi_0^{-1}(t) \leqslant h(t) \quad \text{for all } 0 \leqslant t < \infty;$$

cf. [3] (Proposition 2.5.10). h is continuous at the origin. Moreover, since h is increasing, concave, and satisfies $\lim_{t\to\infty} h(t) = \infty$, it must be continuous on $(0, \infty)$ and strictly increasing on $[0, \infty)$. Setting $\Phi(\lambda) := h^{-1}(\lambda)$ we get a convex bijection $\Phi\colon [0, \infty) \to [0, \infty)$ satisfying

$$\Phi(\lambda) \leqslant \Phi_0(\lambda) \leqslant \Phi\big((c+1)\lambda\big).$$

To show that $\Phi \in \mathscr{Y}_{\sup}$ we choose $\Delta \geqslant 1$ such that $\inf_{\lambda\geqslant 1} \psi(\Delta\lambda)/\psi(\lambda) \geqslant 2$. This implies for $\lambda, \mu \geqslant 1$

$$\psi(\lambda)+\psi(\mu) \leqslant 2\psi(\lambda\mu) \leqslant \psi(\Delta\lambda\mu) \quad \text{and} \quad e^{\psi(\lambda)} e^{\psi(\mu)} \leqslant e^{\psi(\Delta\lambda\mu)} + [e^{\psi(\lambda)} + e^{\psi(\mu)} - 2],$$

which means that $\Phi_0(\lambda)\Phi_0(\mu) \leqslant \Phi_0(\Delta\lambda\mu)$. Consequently, we can deduce that for $\lambda, \mu \geqslant 1$

$$\Phi(\lambda)\Phi(\mu) \leqslant \Phi_0(\lambda)\Phi_0(\mu) \leqslant \Phi_0(\Delta\lambda\mu) \leqslant \Phi\big((c+1)\Delta\lambda\mu\big).$$

Moreover, assuming that $\|f\|_{L_\Phi} \leqslant 1/(c+1)$, we get for $\lambda > 0$

$$\lambda P(e^{\psi(|f|)} > \lambda) \leqslant E e^{\psi(|f|)} \leqslant E\Phi\big((c+1)|f|\big)+1 \leqslant 2,$$

so that $P(|f| > \lambda) \leqslant e^{1-\psi(\lambda)}$ and $|f|_{M_\psi} \leqslant 1$. Now let $|f|_{M_\psi} \leqslant 1$ so that we have $P(|f| > \lambda) \leqslant e^{1-\psi(\lambda)}$ for $\lambda \geqslant 0$. Choosing some $d > 1$ with $\psi(d\lambda) \geqslant (1+e)\psi(\lambda)$ for $\lambda \geqslant 0$, we get

$$E\Phi_0\left(\frac{|f|}{d}\right) = \int_0^\infty P\left(\exp\left\{\psi\left(\frac{|f|}{d}\right)\right\} > \lambda\right) d\lambda - 1 \leqslant \int_1^\infty P\left(\exp\left\{\psi\left(\frac{|f|}{d}\right)\right\} > \lambda\right) d\lambda$$
$$\leqslant \int_1^\infty e\left(\frac{1}{\lambda}\right)^{1+e} d\lambda = 1$$

and $\|\cdot\|_{L_\Phi} \leqslant d|\cdot|_{M_\psi}$. ∎

3. A martingale inequality. Assume a subset E of sequences $f = (d_k)_{k=0}^N \subset L_0^X(\Omega, \mathscr{F}, P)$ adapted to $(\mathscr{F}_k)_{k=0}^N$ and

$$S\colon E \to L_0^+(\Omega, \mathscr{F}, P) := \{f \in L_0(\Omega, \mathscr{F}, P) \mid f \geqslant 0 \text{ a.s.}\}.$$

If for all $f = (d_k)_{k=0}^N \in E$, $g = (e_k)_{k=0}^N \in E$, and all stopping times σ, τ we have

(A1) $-f = (-d_k)_{k=0}^N \in E$, $d_0 = 0$, and ${}^\sigma f^\tau := (d_k \chi_{\{\sigma < k \leqslant \tau\}})_{k=0}^N \in E$;

(A2) $S(f+g) \leqslant \gamma_S [Sf + Sg]$ a.s. for some $\gamma_S \geqslant 1$ if $f+g := (d_k+e_k)_{k=0}^N \in E$;

(A3) $Sf = 0$ a.s. on $\{0 = E(\|d_k\| \mid \mathscr{F}_{k-1}),\ k = 1, \ldots, n\}$ and $Sf = S(-f)$ a.s.;

(A4) $Sf^k \leqslant Sf$ a.s. for $k = 0, \ldots, N$, where $f^k := (d_l \chi_{\{l \leqslant k\}})_{l=0}^N$;

(A5) Sf^k is $\mathscr{F}_{k-1}$-measurable for $k = 1, \ldots, N$;

then we say that (E, S) *satisfies* (A). Moreover, we use

$$f_k := \sum_{l=0}^{k} d_l, \quad f^* = \sup_{0 \leqslant k \leqslant N} \|f_k\|_X,$$

$$S^*f := \sup_{0 \leqslant k \leqslant N} Sf^k, \quad \text{and} \quad T^*f := \sup_{0 \leqslant k \leqslant N} \|Tf_k\|_Y, \text{ where } T \in \mathscr{L}(X, Y).$$

DEFINITION 3.1. Let E be a set of sequences $f = (d_k)_{k=0}^N \subset L_0(\Omega, \mathscr{F}, P)$ with $d_0 = 0$ adapted to the filtration $(\mathscr{F}_k)_{k=0}^N$ and let X be a Banach space. A sequence $F = (D_k)_{k=0}^N \subset L_0^X(\Omega, \mathscr{F}, P)$ adapted to $(\mathscr{F}_k)_{k=0}^N$ *belongs to* E^X if $D_0 = 0$ and if there is a sequence $(a_i)_{i=1}^\infty \subset B_{X'}$ and a closed subspace $X_0 \subseteq X$ such that

$$P(D_l \in X_0) = 1 \quad \text{and} \quad \langle F, a_i \rangle := (\langle D_k, a_i \rangle)_{k=0}^N \in E$$

for $l = 1, \ldots, N$, $i = 1, 2, \ldots$, and $\|x\|_X = \sup_{i=1,2,\ldots} |\langle x, a_i \rangle|$ for $x \in X_0$. We say that $(a_i)_{i=1}^\infty$ is *norming* for F.

THEOREM 3.2. *Assume that* (E, S) *satisfies* (A) *and let* $\psi \in \overline{\mathscr{D}}$ *and* $\Phi \in \mathscr{Y}_{\sup}$ *with* $|\cdot|_{M_\psi} \sim_c \|\cdot\|_{L_\Phi}$ *for some* $c > 0$. *If*

$$\|(f_k)_{k=0}^N\|_{BMO_\psi} \leqslant \|Sf\|_{L_\infty} \quad \text{for } f \in E,$$

then for $T \in \Pi_\Phi(X, Y)$, $f \in E^X$ *with a norming sequence* $(a_i)_{i=1}^\infty \subset B_{X'}$, *and* $1 \leqslant r < \infty$ *we have*

$$\|T^*f\|_{L_r} \leqslant c\psi^{-1}(r)\, \pi_\Phi(T) \Big\| \sup_{i=1,2,\ldots} S(\langle f, a_i \rangle) \Big\|_{L_r},$$

where $c > 0$ *depends on* γ_S, ψ, *and* c *only.*

Proof. Fix $(a_i)_{i=1}^\infty \subset B_{X'}$ and a closed subspace $X_0 \subseteq X$ such that for all $x \in X_0$ we have $\|x\| = \sup_{i=1,2,\ldots} |\langle x, a_i \rangle|$. Let $\mathscr{E}$ be the set of $f = (d_k)_{k=0}^N \subset L_0^X(\Omega, \mathscr{F}, P)$ adapted to $(\mathscr{F}_k)_{k=0}^N$ with $d_0 = 0$,

$$P(d_k \in X_0) = 1, \quad \langle f, a_i \rangle \in E, \quad \text{and} \quad \sup_{j=1,2,\ldots} S(\langle f, a_j \rangle) < \infty \text{ a.s.}$$

for $k = 0, \ldots, N$ and $i = 1, 2, \ldots$, and let $A, B\colon \mathscr{E} \to L_0^+(\Omega, \mathscr{F}, P)$ be given by

$$Af := \|Tf_N\|_Y \quad \text{and} \quad Bf := \sup_{i=1,2,\ldots} S(\langle f, a_i \rangle),$$

where $Bf := 0$ on $\{\sup_{i=1,2,\ldots} S(\langle f, a_i \rangle) = \infty\}$. The triple $(\mathscr{E}, A, B)$ satisfies the assumptions of [13] (Proposition 7.3, $C = 0$). For example, ${}^\sigma g^\tau \in \mathscr{E}$ if $g \in \mathscr{E}$ since $S({}^\sigma g^\tau) \leqslant a\, Sg$ a.s. for some $a > 0$ depending on γ_S only [9] (Lemma 2.1) (cf. [13], Lemma 7.1). Now, from the definition of $\pi_\Phi(T)$ and Remark 1.5 (2) with $f_{-1} := 0$ and $Af^{-1} := 0$ we get

$$c^{-1} \|(Af^k)_{k=0}^N\|_{BMO_\psi} \leqslant \sup_{\substack{0 \leqslant k \leqslant l \leqslant N \\ C \in \mathscr{F}_k, P(C) > 0}} \|T(f_l - f_{k-1})\|_{L_\Phi^Y(C, P_C)}$$

$$\leqslant \pi_\Phi(T|X_0 \to Y) \sup_{\substack{0\leqslant k\leqslant l\leqslant N \\ C\in\mathscr{F}_k, P(C)>0}} \sup_{i=1,2,\ldots} \|\langle(f_l - f_{k-1}), a_i\rangle\|_{L_\Phi(C,P_C)}$$

$$\leqslant \pi_\Phi(T) \sup_{i=1,2,\ldots} c\,\|(\langle f_k, a_i\rangle)_{k=0}^N\|_{BMO_\psi}$$

$$\leqslant c\pi_\Phi(T) \sup_{i=1,2,\ldots} \|S(\langle f, a_i\rangle)\|_{L_\infty} \leqslant c\pi_\Phi(T)\|Bf\|_{L_\infty}.$$

Hence we can apply [13] (Theorem 1.7) and are done. ∎

Combining Theorem 3.2 with [13] (Theorem 4.6 (23)) we obtain

COROLLARY 3.3. *Assume that* (E, S) *satisfies* (A) *and let* $0 < s < \frac{1}{2}$ *be such that*

$$\sup_{0\leqslant k\leqslant l\leqslant N} \sup_{\substack{C\in\mathscr{F}_k \\ P(C)>0}} P_C(|f_l - f_{k-1}| > \|Sf\|_{L_\infty}) \leqslant s \quad \textit{for } f\in E.$$

Then for $T\in\Pi_\Phi(X, Y)$ *with* $\Phi(\lambda) = e^\lambda - 1$, $f\in E^X$ *with a norming sequence* $(a_i)_{i=1}^\infty \subset B_{X'}$, *and* $1 \leqslant r < \infty$ *we have*

$$\|T^* f\|_{L_r} \leqslant cr\pi_\Phi(T)\Big\| \sup_{i=1,2,\ldots} S(\langle f, a_i\rangle)\Big\|_{L_r},$$

where $c > 0$ *depends on* γ_S *and* s *only.*

For some further applications we need

DEFINITION 3.4. (1) For martingale difference sequences

$$f = (d_k)_{k=0}^N \subset L_1(\Omega, \mathscr{F}, P) \quad \text{and} \quad F = (D_k)_{k=0}^N \subset L_1^X(\Omega, \mathscr{F}, P)$$

we let

$$S_2 f := \Big(\sum_{k=0}^N |d_k|^2\Big)^{1/2} \quad \text{and} \quad S_2^w F := \sup_{a\in B_{X'}} S_2(\langle F, a\rangle),$$

$$S_{p,\infty} f := \sup_{1\leqslant k\leqslant N} \sqrt[p]{k}\, d_k^*, \quad \text{and} \quad S_{p,\infty}^w F := \sup_{a\in B_{X'}} S_{p,\infty}(\langle F, a\rangle),$$

where $1 < p < 2$ and $\big(d_k^*(\omega)\big)_{k=1}^N$ is a non-increasing rearrangement of $\big(|d_k(\omega)|\big)_{k=1}^N$.

(2) The set of all martingale difference sequences $f = (d_k)_{k=0}^N \subset L_1(\Omega, \mathscr{F}, P)$ with respect to $(\mathscr{F}_k)_{k=0}^N$ such that $d_0 = 0$ and $|d_k|$ is $\mathscr{F}_{k-1}$-measurable for $k = 1, \ldots, N$ is denoted by $\mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$(¹).

Note that for example $(h_k x_k)_{k=0}^N \in \mathscr{P}^X\big((\mathscr{F}_k^h)_{k=0}^N\big)$, where $(x_k)_{k=0}^N \subset X$ with $x_0 = 0$. For $S\in\{S_2, S_{p,\infty}\}$ the function $S^w F$ is measurable as a composition of $\Omega \to l_\infty^N(X)$ with $\omega \to \big(D_1(\omega), \ldots, D_N(\omega)\big)$ and a continuous map from $l_\infty^N(X)$

(¹) We will write $\mathscr{P}^X\big((\mathscr{F}_k)_{k=0}^N\big)$ instead of $\big(\mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)\big)^X$.

into $\boldsymbol{R}$. Moreover, given $(a_i)_{i=1}^{\infty} \subset B_{X'}$, norming for X, by duality we get

$$\sup_{i=1,2,\ldots} S(\langle F, a_i\rangle)(\omega) \leqslant S^w F(\omega) \leqslant c \sup_{i=1,2,\ldots} S(\langle F, a_i\rangle)(\omega)$$

with $c = 1$ if $S = S_2$ and $c = c_p^2$ if $S = S_{p,\infty}$, where $c_p > 1$ is a constant such that

$$\sup_{k=1,2,\ldots} \sqrt[p]{k}\xi_k^* \sim_{c_p} \|(\xi_k)_{k=1}^{\infty}\|$$

for an equivalent norm $\|\cdot\|$ on $l_{p,\infty}$. In order to describe tail estimates for $S_{p,\infty}$ we use the notion of the K-functional, which is defined for a compatible couple of Banach spaces (X_0, X_1) and $x \in X_0 + X_1$ as

$$K(x; t; X_0, X_1)$$
$$:= \inf\{\|x_0\|_{X_0} + t\|x_1\|_{X_1} \mid x = x_0 + x_1,\ x_0 \in X_0,\ x_1 \in X_1\} \quad (t \geqslant 0).$$

LEMMA 3.5. *Let* $1 < p \leqslant 2 \leqslant q < \infty$ *with* $1 = 1/p + 1/q$, $S = S_2$ *if* $p = 2$, *and* $S = S_{p,\infty}$ *if* $1 < p < 2$. *Then there is a constant* $c > 0$, *depending on p only, such that*

$$\|(f_k)_{k=0}^N\|_{BMO_{\psi_q}} \leqslant c\|Sf\|_{L_\infty} \quad \textit{for } f \in \mathscr{P}((\mathscr{F}_k)_{k=0}^N).$$

Proof. According to a result of Hitczenko [16] (Theorem 4.1), for $\lambda > 0$ and $f = (d_k)_{k=0}^N \in \mathscr{P}((\mathscr{F}_k)_{k=0}^N)$ we have

$$\boldsymbol{P}\Big(\Big|\sum_{k=1}^N d_k\Big| > c\,\big\|K\big((d_k)_{k=1}^N, \lambda; l_1^N, l_2^N\big)\big\|_{L_\infty}\Big) \leqslant 2\exp\left\{-\frac{\lambda^2}{c}\right\}, \tag{6}$$

where $c > 0$ is an absolute constant. Hitczenko proved this inequality for a transform $(v_k \varepsilon_k)_{k=1}^N$ of a Rademacher sequence $(\varepsilon_k)_{k=1}^N$ by some predictable sequence $(v_k)_{k=1}^N$. If we consider $(d_k)_{k=0}^N \in \mathscr{P}((\mathscr{F}_k)_{k=0}^N)$, then $d_k = |d_k| \operatorname{sgn} d_k$, where $\operatorname{sgn} d_k(\omega) := d_k(\omega)/|d_k(\omega)|$ if $d_k(\omega) \neq 0$ and $\operatorname{sgn} d_k(\omega) := 0$ if $d_k(\omega) = 0$. Since $(\operatorname{sgn} d_k)_{k=0}^N \in \mathscr{P}((\mathscr{F}_k)_{k=0}^N)$ and $(|d_k|)_{k=1}^N$ is predictable, we replace ε_k by $\operatorname{sgn} d_k$, and v_k by $|d_k|$. Now looking at Hitczenko's proof we realize that this proof works as well without any changes. In particular, we can also use for a predictable sequence $(w_k)_{k=1}^N$ the inequality

$$\boldsymbol{P}\Big(\Big|\sum_{k=1}^N w_k \operatorname{sgn} d_k\Big| > \lambda\|S_2 g\|_{L_\infty}\Big) \leqslant 2\exp\left\{-\frac{\lambda^2}{2}\right\},$$

which follows from [10] or [15] (Lemma 4.3) and an approximation argument with respect to the w_k. It is known that there is an absolute constant $c_q > 0$, depending on q only, such that

$$K(x, \lambda^{q/2}; l_1^N, l_2^N) \leqslant \lambda c_q \|x\|_{\mathscr{E}_p^N} \quad \text{for } \lambda \geqslant 0, \tag{7}$$

where $\mathscr{E}_p^N := l_2^N$ if $p = 2$ and $\mathscr{E}_p^N := l_{p,\infty}^N$ if $1 < p < 2$. Inequalities (6) and (7) imply, for $\lambda \geqslant 0$ and $f = (d_k)_{k=0}^N \in \mathscr{P}((\mathscr{F}_k)_{k=0}^N)$,

$$\boldsymbol{P}\Big(\Big|\sum_{k=1}^N d_k\Big| > \lambda c_q c\|Sf\|_{L_\infty}\Big) \leqslant 2\exp\left\{-\frac{\lambda^q}{c}\right\}.$$

Considering $0 \leqslant k \leqslant l \leqslant N$ and $C \in \mathscr{F}_k$ with $P(C) > 0$ we obtain

$$|\sum_{i=k}^{l} d_i|_{M_{\psi_q}(C,P_C)} \leqslant c_q' [|\sum_{i=k+1}^{l} d_i|_{M_{\psi_q}(C,P_C)} + \|Sf\|_{L_\infty}] \leqslant c_q'' \|Sf\|_{L_\infty}. \blacksquare$$

COROLLARY 3.6. *Let* $1 < p \leqslant 2 \leqslant q < \infty$ *with* $1 = 1/p + 1/q$, $S = S_2$ *if* $p = 2$, *and* $S = S_{p,\infty}$ *if* $1 < p < 2$. *Then there is a constant* $c > 0$, *depending on* p *only, such that for* $T \in \Pi_{\Phi_q}(X, Y)$, $f \in \mathscr{P}^X((\mathscr{F}_k)_{k=0}^N)$, *and* $1 \leqslant r < \infty$

$$\|T^* f\|_{L_r} \leqslant c \sqrt[q]{r}\, \pi_{\Phi_q}(T) \|S^w f\|_{L_r}.$$

Proof. We take $E = \mathscr{P}((\mathscr{F}_k)_{k=0}^N)$ and use Theorem 3.2 and Lemma 3.5. ■

COROLLARY 3.7. *Let* $1 < p \leqslant 2 \leqslant q < \infty$ *be such that* $1 = 1/p + 1/q$, $T \in \Pi_{\Phi_q}(X, Y)$, *and* $f = (d_k)_{k=0}^N \in \mathscr{P}^X((\mathscr{F}_k)_{k=0}^N)$. *Then, for some* $c > 0$, *depending on* p *only,*

$$\|T^* f\|_{L_p} \leqslant c\, \pi_{\Phi_q}(T) \Big(\int_\Omega \sum_{k=1}^N \|d_k\|_X^p \, dP\Big)^{1/p}. \tag{8}$$

Proof. Use Corollary 3.6 and $S^w f \leqslant \big(\sum_{k=1}^N \|d_k\|_X^p\big)^{1/p}$. ■

Remark 3.8. (1) Pisier has shown in [29] that for $T = I_X$ inequality (8) is equivalent to a renorming of the Banach space X such that the modulus of smoothness is of power type p. The same arguments apply in the operator case (see e.g. the forthcoming book [28]), so that Corollary 3.7 implies smoothness properties of the absolutely Φ_q-summing operators.

(2) Inequality (8) fails to be true for the absolutely Φ_r-summing operators whenever $2 \leqslant q < r < \infty$. In fact, for the embedding

$$I_r \colon C[0, 1) \to L_{\Phi_r}[0, 1) \in \Pi_{\Phi_r}$$

inequality (8) would imply type p, which means

$$\int_M \Big\| \sum_{k=1}^N \varepsilon_k x_k \Big\|_{L_{\Phi_r}[0,1)}^p d\mu \leqslant c^p \sum_{k=1}^N \|x_k\|_{C[0,1)}^p$$

for all $x_1, \ldots, x_N \in C[0, 1)$ and independent Rademacher variables $\varepsilon_1, \ldots, \varepsilon_N$. Approximating $r_k(t) = \sum_{l=2^{k-1}}^{2^k-1} h_l(t) \in L_1[0, 1)$ by $x_k \in C[0, 1)$ in an appropriate way we obtain a contradiction to the type p property of I_r.

COROLLARY 3.9. *For* $T \in \Pi_{\Phi_2}(X, Y)$ *and* $f = (d_k)_{k=0}^N \in \mathscr{P}^X((\mathscr{F}_k)_{k=0}^N)$ *we have*

$$\|T^* f\|_{L_2} \leqslant c\, \pi_{\Phi_2}(T) \Big\| \sum_{k=1}^N \varepsilon_k d_k \Big\|_{L_2^X(M \times \Omega)},$$

where $\varepsilon_1, \ldots, \varepsilon_N$ *are independent Rademacher variables and* $c > 0$ *is an absolute constant.*

Proof. Use Corollary 3.6 and $S_2^w f \leqslant \big\|\sum_{k=1}^N \varepsilon_k d_k\big\|_{L_2^X(M,\mu)}$. ■

Note that we have only used that the Rademacher variables form an orthonormal system. To explain another application let us consider for $t \geqslant 1$ and $2 \leqslant q < \infty$ the weight $w_t^q\colon [0, 1] \to [0, 1]$,

$$w_t^q(s) := \begin{cases} \dfrac{1}{\sqrt[q]{1+\log(st)}} & \text{for } 1/t \leqslant s \leqslant 1, \\ 1 & \text{for } 0 < s < 1/t, \end{cases}$$

so that

$$\frac{1}{\sqrt[q]{1+\log t}} \leqslant w_t^q \leqslant 1,$$

and for $1 \leqslant r < \infty$ and $h \in L_r(\Omega, \mathscr{F}, P)$ the weighted K-functional

$$K^{w_t^q}(h, t; L_\infty, L_r) := K\big(w_t^q(s)\, h(\omega), t; L_\infty(\Omega'), L_r(\Omega')\big) \quad \text{with} \quad \Omega' = [0, 1] \times \Omega.$$

The next corollary is contained and motivated for $p = 2$ in [13].

COROLLARY 3.10. *Let* $1 < p < 2 < q < \infty$ *with* $1 = 1/p + 1/q$ *and* $f \in \mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$. *Then*

$$K^{w_t^q}(f^*, t^{1/r}; L_\infty, L_r) \leqslant c \sqrt[q]{r}\, K(S_{p,\infty} f, t^{1/r}; L_\infty, L_r)$$

for $t \geqslant 1$ *and* $1 \leqslant r < \infty$, *where* $c > 0$ *depends on* p *only.*

Proof. We can easily see that it is enough to prove the statement for $t \in \{1, 2, \ldots\}$. Consider

$$[\Omega^t, \mathscr{F}^t, P^t] := \times_1^t [\Omega, \mathscr{F}, P]$$

and the product filtration

$$(\mathscr{F}_k^t)_{k=0}^N := (\times_1^t \mathscr{F}_k)_{k=0}^N.$$

Fix $f = (d_k)_{k=0}^N \in \mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$ and let $f^j := (d_k^j)_{k=0}^N \in \mathscr{P}\big((\mathscr{F}_k^t)_{k=0}^N\big)$ be given by

$$d_k^j(\omega_1, \ldots, \omega_t) := d_k(\omega_j).$$

Then [13] (Theorem 1.8 and the proof of Theorem 1.7) gives

$$K^{w_t^q}(f^*, t^{1/r}; L_\infty, L_r) \sim_c \left\| \sup_{1 \leqslant j \leqslant t} \frac{f^*(\omega_j)}{\sqrt[q]{1+\log j}} \right\|_{L_r(\Omega^t)} \sim_c \left\| \sup_{1 \leqslant j \leqslant t} \frac{f^*(\omega_j)}{\sqrt[q]{\log(j+1)}} \right\|_{L_r(\Omega^t)},$$

where $c > 0$ depends on q only. Now Theorem 3.2 ($X = l_\infty^t$, $(a_i)_{i=1}^t$ is the unit vector basis of l_1^t), Lemma 3.5, Corollary 1.4, and once more [13] (Theorem 1.8) yield

$$\left\| \sup_{1 \leqslant j \leqslant t} \frac{f^*(\omega_j)}{\sqrt[q]{\log(j+1)}} \right\|_{L_r(\Omega^t)} \leqslant c_{(3.2)} \sqrt[q]{r}\, \pi_{\Phi_q}(D_q)\, c_{(3.5)}^{(q)} \Big\| \sup_{1 \leqslant j \leqslant t} S_{p,\infty} f^j \Big\|_{L_r(\Omega^t)},$$

$$\leqslant c_{(3.2)} \sqrt[q]{r}\, \pi_{\Phi_q}(D_q)\, c_{(3.5)}^{(q)} K(S_{p,\infty} f, t^{1/r}; L_\infty, L_r),$$

where we have used the notation of Corollary 1.4. ∎

For $T \in \mathscr{L}(l_2^n, Y)$ we set $l(T) := \|\sum_{i=1}^n g_i Tv_i\|_{L_1^Y}$, where $(v_i)_{i=1}^n$ is the unit vector basis of l_2^n. From Talagrand's majorizing measure theorem it should be folklore that $\pi_{\Phi_2}(\cdot) \sim l(\cdot)$. Now we easily extend this equivalence to

COROLLARY 3.11. *For some absolute $c > 0$ we have, for all $T \in \mathscr{L}(l_2^n, Y)$ $(n = 1, 2, \ldots)$,*

$$\pi_{\Phi_2}(T) \sim_c l(T) \sim_c \sup\{\|T^* f\|_{L_2} \mid \|S_2^w f\|_{L_2} = 1,\ f \in \mathscr{P}^{l_2^n}((\mathscr{F}_k)_{k=0}^N)\}.$$

Proof. Let us denote the last item in the assertion by $\sigma(T)$. The estimate $\sigma(T) \leqslant c\pi_{\Phi_2}(T)$ follows from Corollary 3.6. To get $l(T) \leqslant \sigma(T)$ take $x_1, \ldots, x_N \in l_2^n$ and martingale differences $d_k := \varepsilon_k x_k$, where ε_k are independent Rademacher variables. We obtain

$$\|\sum_{k=1}^N \varepsilon_k Tx_k\|_{L_2^Y} \leqslant \sigma(T) \sup_{a \in B_{l_2^n}} \Big(\sum_{k=1}^N |\langle x_k, a\rangle|^2\Big)^{1/2}. \tag{9}$$

By the consideration of blocks $s^{-1/2}(\varepsilon_{(k-1)s+1} + \ldots + \varepsilon_{ks})x_k$ in the above inequality and by letting $s \to \infty$ the central limit theorem (cf. [31], p. 90) implies that we can replace in (9) the Rademacher variables by independent standard Gaussian variables so that $l(T) \leqslant \sigma(T)$. To deduce $\pi_{\Phi_2}(T) \leqslant cl(T)$ we can assume that $Y = l_\infty$. It is known that the majorizing measure theorem for Gaussian variables [30] ([23], Theorem 12.10) implies the existence of $\|u_t\|_{l_2^n} \leqslant 1$ $(t = 1, 2, \ldots)$ such that

$$\|Ta\| \leqslant cl(T) \sup_{t=1,2,\ldots} \frac{|\langle a, u_t\rangle|}{\sqrt{\log(t+1)}} \quad \text{for } a \in l_2^n,$$

where $c > 0$ is an absolute constant (cf. the arguments of the proof of Lemma 3.3 in [14]). Hence we can conclude with Corollary 1.4 in the case $p = 2$. ∎

Finally, for UMD-transforms (2), from Corollary 3.3 we get

COROLLARY 3.12. *For $T \in \Pi_\Phi(X, Y)$ with $\Phi(\lambda) = e^\lambda - 1$, $(x_k)_{k=1}^N \subset X$, and $\theta_k = \pm 1$ we have*

$$\|\sum_{k=1}^N h_k Tx_k\|_{L_2^Y[0,1)} \leqslant c\,\pi_\Phi(T) \|\sum_{k=1}^N \theta_k h_k x_k\|_{L_2^X[0,1)},$$

where $c > 0$ is an absolute constant.

Proof. Consider $E := \mathscr{P}((\mathscr{F}_k^h)_{k=0}^N)$ and the operator $S: E \to L_0^+[0, 1)$ given by

$$S((d_k)_{k=0}^N) := \sup_{0 \leqslant k < N} \Big[\Big|\sum_{l=0}^k \theta_l d_l\Big| + |d_{k+1}|\Big].$$

(2) UMD stands for 'unconditional martingale differences.'

The pair (E, S) satisfies condition (A). For $f = (d_k)_{k=0}^N \in E$, $0 \leqslant k \leqslant l \leqslant N$, $C \in \mathscr{F}_k^h$ of positive measure, and the Lebesgue measure λ we get (we can assume that $\|Sf\|_\infty > 0$)

$$\lambda_C\Big(\Big|\sum_{i=k}^{l} d_i\Big| > 8\,\|Sf\|_\infty\Big) \leqslant \frac{1}{8\,\|Sf\|_\infty}\Big\|\sum_{i=k}^{l} d_i\Big\|_{L_2(C,\lambda_C)}$$
$$= \frac{1}{8\,\|Sf\|_\infty}\Big\|\sum_{i=k}^{l} \theta_i d_i\Big\|_{L_2(C,\lambda_C)} \leqslant \frac{1}{4}.$$

Hence Corollary 3.3 applies for $s = 1/4$ so that for $F = (D_k)_{k=0}^N \in E^X$ with a norming sequence $(a_i)_{i=1}^\infty$ we obtain

$$\|T^* F\|_{L_2} \leqslant 8\, c_{(3.3)}\, \pi_\Phi(T)\Big\| \sup_{i=1,2,\dots} S(\langle F, a_i\rangle)\Big\|_{L_2}$$
$$\leqslant 24\, c_{(3.3)}\, \pi_\Phi(T)\Big\| \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} \theta_l D_l\Big\|_X\Big\|_{L_2}$$
$$\leqslant 48\, c_{(3.3)}\, \pi_\Phi(T)\Big\|\sum_{k=1}^{N} \theta_k D_k\Big\|_{L_2^X},$$

where we have used Doob's maximal inequality. ∎

4. The contraction principle and Gaussian variables. For a symmetric random vector $(d_1, \dots, d_n)$, where $d_1, \dots, d_n \in L_2(\Omega, \mathscr{F}, P)$, a Banach space X, and $x_1, \dots, x_n \in X$, the contraction principle states that

$$\Big\|\sum_{i=1}^{n} d_i x_i\Big\|_{L_2^X} \leqslant \Big\|\sup_{1\leqslant i\leqslant n} |d_i|\Big\|_{L_2}\Big\|\sum_{i=1}^{n} \varepsilon_i x_i\Big\|_{L_2^X}. \tag{10}$$

For basic information the reader is referred to [21], [17] and [18]. As we will see in Theorem 5.1 inequality (10) remains true with some additional multiplicative constant if $(d_i)_{i=1}^n$ is a martingale difference sequence. Now we ask for a similar inequality for the Gaussian variables instead of the Rademacher variables. Since

$$\Big\|\sum_i \varepsilon_i x_i\Big\|_{L_2^X} \leqslant \sqrt{\pi/2}\,\Big\|\sum_i g_i x_i\Big\|_{L_2^X}$$

for independent standard Gaussian variables $g_1, \dots, g_n$, from Theorem 5.1 for a martingale difference sequence $(d_i)_{i=1}^n$ we also get

$$\Big\|\sum_{i=1}^{n} d_i x_i\Big\|_{L_2^X} \leqslant c\Big\|\sup_{1\leqslant i\leqslant n} |d_i|\Big\|_{L_2}\Big\|\sum_{i=1}^{n} g_i x_i\Big\|_{L_2^X}. \tag{11}$$

Analyzing (11) we observe that $\|\sup_{1\leqslant i\leqslant n} |d_i|\|_{L_2}$ is far from being an optimal factor since for $d_i = g_i$

$$\Big\|\sum_{i=1}^{n} g_i x_i\Big\|_{L_2^X} \leqslant c\Big\|\sup_{1\leqslant i\leqslant n} |g_i|\Big\|_{L_2}\Big\|\sum_{i=1}^{n} g_i x_i\Big\|_{L_2^X} \sim \sqrt{\log(n+1)}\,\Big\|\sum_{i=1}^{n} g_i x_i\Big\|_{L_2^X}. \tag{12}$$

In Corollary 4.2 we remove this defect in (12). The corollary will follow from

THEOREM 4.1. *Let $g_1, \ldots, g_n$ be independent standard Gaussian random variables, X be a Banach space, and $x_1, \ldots, x_n \in X$. If $(v_i)_{i=1}^n$ is the unit vector basis of l_2^n and if we have $(d_k^i)_{k=0}^N \subset L_1(\Omega, \mathscr{F}, P)$ such that*

$$f = \Big(\sum_{i=1}^n d_k^i v_i\Big)_{k=0}^N \in \mathscr{P}^{l_2^n}\big((\mathscr{F}_k)_{k=0}^N\big),$$

then for $1 \leqslant r < \infty$

$$\Big(E \sup_{1 \leqslant k \leqslant N} \Big\| \sum_{i=1}^n \Big(\sum_{l=1}^k d_l^i\Big) x_i \Big\|_X^r\Big)^{1/r} \leqslant c\sqrt{r}\, \big\| \|A\|_{\mathscr{L}(l_2^n, l_2^N)} \big\|_{L_r} \Big\| \sum_{i=1}^n g_i x_i \Big\|_{L_1^X},$$

where $A(\omega) := \big(d_k^i(\omega)\big)_{i=1,k=1}^{n,N}$ and $c > 0$ is an absolute constant.

Proof. We have to combine Corollary 3.11 for the operator $T \in \mathscr{L}(l_2^n, X)$ defined by $Tv_i := x_i$ with $S_2^w f = \|A\|_{\mathscr{L}(l_2^n, l_2^N)}$. ■

To discuss some special cases we use the following. If $(V_i)_{i=1}^L \subset l_2^n$ and $(y_i)_{i=1}^L \subset l_2^N$ are vectors having pairwise disjoint supports, respectively, and if $T = \sum_{i=1}^L V_i \otimes y_i \in \mathscr{L}(l_2^n, l_2^N)$ is given by $Tx := \sum_{i=1}^L \langle x, V_i \rangle y_i$, then

$$\Big\| \sum_{i=1}^L V_i \otimes y_i \Big\|_{\mathscr{L}(l_2^n, l_2^N)} = \sup_{1 \leqslant i \leqslant L} \|V_i\|_{l_2^n} \|y_i\|_{l_2^N}. \tag{13}$$

Moreover, for an adapted sequence $f = (d_k)_{k=0}^N \subset L_0(\Omega, \mathscr{F}, P)$ and stopping times σ, τ we write

$${}^{\sigma}\!\Delta^{\tau} f := \sum_{\sigma < k \leqslant \tau} d_k.$$

COROLLARY 4.2. *Let $g_1, \ldots, g_n$ be independent standard Gaussian random variables. Then for all Banach spaces X, $x_1, \ldots, x_n \in X$, $f = (d_k)_{k=0}^N \in \mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$, all sequences of stopping times $0 = \tau_0 \leqslant \tau_1 \leqslant \ldots \leqslant \tau_n = N$, and $1 \leqslant r < \infty$, we have*

$$\Big(E \sup_{1 \leqslant k \leqslant N} \Big\| \sum_{i=1}^n \big[{}^{\tau_{i-1}}\!\Delta^{\tau_i \wedge k} f\big] x_i \Big\|_X^r\Big)^{1/r} \leqslant c\sqrt{r} \Big\| \sup_{1 \leqslant i \leqslant n} S_2\big({}^{\tau_{i-1}}\!f^{\tau_i}\big)\Big\|_{L_r} \Big\| \sum_{i=1}^n g_i x_i \Big\|_{L_1^X}, \tag{14}$$

where $c > 0$ is an absolute constant.

Proof. The matrix $A(\omega)$ of Theorem 4.1 can be written as

$$A(\omega) = \sum_{i=1}^n v_i(\omega) \otimes y_i(\omega),$$

where $y_i = (0, \ldots, 0, d_{\tau_{i-1}+1}, \ldots, d_{\tau_i}, 0, \ldots, 0)$ and $d_{\tau_{i-1}+1}$ is the $(\tau_{i-1}+1)$-st coordinate and where $(v_i)_{i=1}^n$ is the unit vector basis of l_2^n. Moreover, the martingale difference sequence generated by $\sum_{i=1}^n \big[{}^{\tau_{i-1}}\!\Delta^{\tau_i} f\big] v_i$ belongs clearly to $\mathscr{P}^{l_2^n}\big((\mathscr{F}_k)_{k=0}^N\big)$. ■

If we approximate the Gaussian variables by

$$g_i^M := \frac{1}{\sqrt{M}}(\varepsilon_{(i-1)M+1} + \dots + \varepsilon_{iM}),$$

then, by using the central limit theorem (see [31], p. 90), (14) turns into the Khintchine–Kahane inequality for the Gaussian variables:

$$(15) \qquad \Big\|\sum_{i=1}^n g_i x_i\Big\|_{L_r^X} = \lim_{M\to\infty} \Big\|\sum_{i=1}^n g_i^M x_i\Big\|_{L_r^X} \leqslant c\sqrt{r}\Big\|\sum_{i=1}^n g_i x_i\Big\|_{L_1^X}.$$

Consequently, a defect like in (12) does not appear. In this sense, $\|\sup_{1\leqslant i\leqslant n} S_2({}^{\tau_{i-1}}f^{\tau_i})\|_{L_r}$ is an optimal factor in (14). The argument for inequality (15) shows more. We cannot replace in (14) the Gaussian variables by the Rademacher variables. If this were possible, then (14) and again the central limit theorem would imply

$$\Big\|\sum_{i=1}^n g_i x_i\Big\|_{L_2^X} \leqslant c\sqrt{2}\Big\|\sum_{i=1}^n \varepsilon_i x_i\Big\|_{L_2^X},$$

which uniformly in n holds for Banach spaces of finite cotype only (see [26]). Considering $X = \mathbf{R}$ and $n = 1$ in (14) gives the following Burkholder–Davis–Gundy type inequality:

$$\|f^*\|_{L_r} \leqslant c\sqrt{r}\|g_1\|_{L_1}\|S_2 f\|_{L_r} \qquad \text{for } f \in \mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$$

(see [10], [4], [15], and [32]). Another consequence of Theorem 4.1 is

COROLLARY 4.3. *Let $(g_{ij})_{1\leqslant i<j\leqslant n}$ be independent standard Gaussian random variables, $f = (d_k)_{k=0}^N \in \mathscr{P}\big((\mathscr{F}_k)_{k=0}^N\big)$, and $0 = \tau_0 \leqslant \tau_1 \leqslant \dots \leqslant \tau_n = N$ be a sequence of stopping times. Then for all Banach spaces X, all $(x_{ij})_{1\leqslant i<j\leqslant n} \subset X$, and all $1 \leqslant r < \infty$ we have*

$$\Big\|\sum_{1\leqslant i<j\leqslant n} [{}^{\tau_{i-1}}\Delta^{\tau_i} f][{}^{\tau_{j-1}}\Delta^{\tau_j} f]\, x_{ij}\Big\|_{L_r^X}$$

$$\leqslant c\sqrt{r}\Big\|\sup_{2\leqslant i\leqslant n}\Big(\sum_{l=1}^{i-1} |{}^{\tau_{l-1}}\Delta^{\tau_l} f|^2\Big)^{1/2} S_2({}^{\tau_{i-1}}f^{\tau_i})\Big\|_{L_r}\Big\|\sum_{1\leqslant i<j\leqslant n} g_{ij} x_{ij}\Big\|_{L_1^X},$$

where $c > 0$ is an absolute constant.

Proof. Define random vectors $V_2, \dots, V_n \in l_2^{n(n-1)/2}$ by

$$V_i := (0, \dots, 0, {}^{\tau_0}\Delta^{\tau_1} f, \dots, {}^{\tau_{i-2}}\Delta^{\tau_{i-1}} f, 0, \dots, 0),$$

where ${}^{\tau_0}\Delta^{\tau_1} f$ is the $\big(1 + [1 + \dots + (i-2)]\big)$-nd coordinate, and determine random vectors $y_2, \dots, y_n \in l_2^N$ by

$$y_i := (0, \dots, 0, d_{\tau_{i-1}+1}, \dots, d_{\tau_i}, 0, \dots, 0),$$

where $d_{\tau_{i-1}+1}$ is the $(\tau_{i-1}+1)$-st coordinate. If we arrange the elements $(x_{ij})_{1\leqslant i<j\leqslant n}$ in the linear order $x_{12}, x_{13}, x_{23}, \dots, x_{1n}, \dots, x_{n-1,n}$, then the

matrix $A(\omega)$ from Theorem 4.1 takes the form $A(\omega)=\sum_{i=2}^{n} V_i(\omega)\otimes y_i(\omega)$. Again, we can easily check that the martingale difference sequence generated by

$$\sum_{1\leqslant i<j\leqslant n} [{}^{\tau_{i-1}}\Delta^{\tau_i}f]\,[{}^{\tau_{j-1}}\Delta^{\tau_j}f]\, v_{ij},$$

where $(v_{ij})_{i<j}$ is the unit vector basis of $l_2^{n(n-1)/2}$ arranged for example in the linear order of the x_{ij}, belongs to $\mathscr{P}^{l_2^{n(n-1)/2}}((\mathscr{F}_k)_{k=0}^{N})$. Hence we can apply Theorem 4.1 and (13). ■

Finally, let us mention the classical setting behind Corollary 4.3.

COROLLARY 4.4. *Let X be a Banach space and $(x_{ij})_{1\leqslant i<j\leqslant n}\subset X$. Then for all $1\leqslant r<\infty$ we have*

$$\Big\|\sum_{1\leqslant i<j\leqslant n} g_i g_j x_{ij}\Big\|_{L_r^X}\leqslant c\,r\sqrt{n}\,\Big\|\sum_{1\leqslant i<j\leqslant n} g_{ij} x_{ij}\Big\|_{L_1^X},$$

where $(g_i)_{i=1}^{n}$ and $(g_{ij})_{1\leqslant i<j\leqslant n}$ are mutually independent standard Gaussian variables and $c>0$ is an absolute constant.

Proof. We apply Corollary 4.3 to the sequence of independent Rademacher variables $f=(\varepsilon_k/\sqrt{s})_{k=1}^{sn}$ and $\tau_i=is$ such that the central limit theorem (cf. [31], p. 90) and the inequality $\|(\sum_{l=1}^{n-1}|g_l|^2)^{1/2}\|_{L_r}\leqslant c\sqrt{r}\sqrt{n-1}$ imply our assertion. ■

Remark 4.5. The factor $\sqrt{n}$ (for fixed r) in Corollary 4.4 is optimal up to a multiplicative factor. To see this consider $x_{ij}:=v_i\otimes v_j\in X:=\mathscr{L}(l_2^n, l_2^n)$, where $(v_i)_{i=1}^{n}$ is the standard basis of l_2^n. Then, on the one hand, we obtain

$$2\Big\|\sum_{1\leqslant i<j\leqslant n} g_i g_j v_i\otimes v_j\Big\|_{L_1^X}\geqslant\Big\|\sum_{i,j=1}^{n} g_i g_j v_i\otimes v_j-\sum_{i=1}^{n} g_i^2 v_i\otimes v_i\Big\|_{L_1^X}$$

$$\geqslant\Big\|\,\Big\|\sum_{i=1}^{n} g_i v_i\Big\|_{l_2^n}^2\Big\|_{L_1}-\Big\|\sup_{1\leqslant i\leqslant n}|g_i|^2\Big\|_{L_1}.$$

Since $E\sum_{i=1}^{n}|g_i|^2=n$ and $\|\sup_{1\leqslant i\leqslant n}|g_i|^2\|_{L_1}\leqslant c_1(1+\log n)$ we continue to

$$\Big\|\sum_{1\leqslant i<j\leqslant n} g_i g_j v_i\otimes v_j\Big\|_{L_1^X}\geqslant c_2\, n,$$

where $c_2>0$ is an absolute constant. On the other hand, we have according to Chevet's inequality [11] (or [23], Theorem 3.20)

$$\Big\|\sum_{1\leqslant i<j\leqslant n} g_{ij} v_i\otimes v_j\Big\|_{L_1^X}\leqslant\Big\|\sum_{i,j=1}^{n} g_{ij} v_i\otimes v_j\Big\|_{L_1^X}\leqslant c_3\sqrt{n}.$$

5. The contraction principle and Rademacher variables. In this last section we prove a version of Corollary 4.2 for the Rademacher variables.

THEOREM 5.1. *Let $(d_k)_{k=0}^N \subset L_1(\Omega, \mathscr{F}, P)$ be a martingale difference sequence with respect to $(\mathscr{F}_k)_{k=0}^N$. Then for all Banach spaces X, all $x_1, \ldots, x_N \in X$, and all $1 \leqslant r < \infty$ we have*

$$\Big(E \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} d_l x_l\Big\|_X^r\Big)^{1/r} \leqslant cr \Big\| \sup_{1\leqslant k\leqslant N} |d_k| \Big\|_{L_r} \Big\|\sum_{k=1}^{N} \varepsilon_k x_k\Big\|_{L_1^X},$$

where $\varepsilon_1, \ldots, \varepsilon_N$ is a sequence of independent Rademacher variables and $c > 0$ is an absolute constant.

For the proof the following direct consequence of [22] (Theorem 5.1.2) is needed:

LEMMA 5.2. *Let X be a Banach space, $x_1, \ldots, x_N \in X$, and $(d_k)_{k=1}^N \subset L_1(\Omega, \mathscr{F}, P)$ be a martingale difference sequence with respect to $(\mathscr{G}_k)_{k=1}^N$. Then for independent Rademacher variables $\varepsilon_1, \ldots, \varepsilon_N$ we have*

$$E\Big\|\sum_{k=1}^{N} d_k x_k\Big\| \leqslant \sup_{1\leqslant k\leqslant N} \|d_k\|_{L_\infty} E\Big\|\sum_{k=1}^{N} \varepsilon_k x_k\Big\|.$$

Proof of Theorem 5.1. We can assume that $d_0 = 0$. First we apply the Davis decomposition ([12]; see also [5], Chapter III) to $(d_k)_{k=0}^N$ and obtain martingale difference sequences $(a_k)_{k=0}^N$ and $(b_k)_{k=0}^N$ with respect to the same filtration satisfying $a_0 = b_0 = 0$,

(1) $d_k = a_k + b_k$ a.s. for $k = 1, \ldots, N$,

(2) $|a_k| \leqslant 4d_{k-1}^*$ a.s. for $k = 1, \ldots, N$,

(3) $\sum_{k=1}^N |b_k| \leqslant \sum_{k=1}^N |z_k| + \sum_{k=1}^N E(|z_k| \,|\, \mathscr{F}_{k-1})$ a.s., where $z_k := d_k \chi_{\{|d_k| > 2d_{k-1}^*\}}$,

(4) $\sum_{k=1}^N |z_k| \leqslant 2d_N^*$ a.s.,

where we make use of the notation $d_k^* := \sup_{0\leqslant l\leqslant k} |d_l|$. We get

$$(16)\qquad \Big(E \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} d_l x_l\Big\|_X^r\Big)^{1/r} \leqslant \Big(E \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} a_l x_l\Big\|_X^r\Big)^{1/r} + \Big(E \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} b_l x_l\Big\|_X^r\Big)^{1/r}.$$

The second term on the right-hand side can be estimated as follows:

$$\begin{aligned}\Big(E \sup_{1\leqslant k\leqslant N} \Big\|\sum_{l=1}^{k} b_l x_l\Big\|_X^r\Big)^{1/r} &\leqslant \sup_{1\leqslant k\leqslant N} \|x_k\|_X \Big\|\sum_{k=1}^{N} |b_k|\Big\|_{L_r} \\ &\leqslant \sup_{1\leqslant k\leqslant N} \|x_k\|_X \Big[\Big\|\sum_{k=1}^{N} |z_k|\Big\|_{L_r} + \Big\|\sum_{k=1}^{N} E(|z_k| \,|\, \mathscr{F}_{k-1})\Big\|_{L_r}\Big] \\ &\leqslant \sup_{1\leqslant k\leqslant N} \|x_k\|_X (1+r) \Big\|\sum_{k=1}^{N} |z_k|\Big\|_{L_r} \leqslant \sup_{1\leqslant k\leqslant N} \|x_k\|_X \, 2(1+r) \|d_N^*\|_{L_r},\end{aligned}$$

where we have used the convexity lemma [8] (cf. [5] (Lemma 16.1), and for the constant e.g. [25] (I.9.6.)). Let us turn to the first term on the right-hand side of (16). Define

$$v_k := 4d_{k-1}^* \text{ for } k = 1, \ldots, N, \quad v_0 := 0,$$

and the set E of sequences adapted to $(\mathscr{F}_k)_{k=0}^N$:

$$E := \{\pm((a_k, v_k)\chi_{\{\sigma<k\leqslant\tau\}})_{k=0}^N \subset L_1^{R\oplus_\infty R}(\Omega, \mathscr{F}, P) \mid \sigma, \tau \text{ stopping times}\}.$$

Moreover, we consider operators $A, B\colon E \to L_0^+(\Omega, \mathscr{F}, P)$ given by

$$A(((\alpha_k, \beta_k))_{k=0}^N) := \|\sum_{k=1}^N \alpha_k x_k\|_X \quad \text{and} \quad B(((\alpha_k, \beta_k))_{k=0}^N) := \sup_{1\leqslant k\leqslant N} |\beta_k|.$$

The triple (E, A, B) satisfies the conditions of [13] (Proposition 7.3, $C = 0$). Now let $0 \leqslant k \leqslant l \leqslant N$ and $C \in \mathscr{F}_k$ with $P(C) > 0$. For $f = ((\alpha_k, \beta_k))_{k=0}^N \in E$ we get

$$\|Af^l - Af^{k-1}\|_{L_1(C,P_C)} \leqslant \|\sum_{i=k}^l \alpha_i x_i\|_{L_1^X(C,P_C)} \leqslant \sup_{k\leqslant i\leqslant l} \|\alpha_i\|_{L_\infty} \|\sum_{i=k}^l \varepsilon_i x_i\|_{L_1^X},$$

where we have used Lemma 5.2. Consequently,

$$\|Af^l - Af^{k-1}\|_{L_1(C,P_C)} \leqslant \sup_{k\leqslant i\leqslant l} \|\beta_i\|_{L_\infty} \|\sum_{i=k}^l \varepsilon_i x_i\|_{L_1^X} \leqslant \|Bf\|_{L_\infty} \|\sum_{i=1}^N \varepsilon_i x_i\|_{L_1^X}.$$

Applying [13] (Theorem 1.7) with $\psi(\lambda) = 1 + \log\lambda$ we obtain

$$\|A^*f\|_{L_r} \leqslant cr \|\sum_{i=1}^N \varepsilon_i x_i\|_{L_1^X} \|Bf\|_{L_r}.$$

Summarizing the estimates of the first and second terms on the right-hand side of (16) we can conclude the proof with

$$\begin{aligned}(E \sup_{1\leqslant k\leqslant N} \|\sum_{l=1}^k d_l x_l\|_X^r)^{1/r} &\leqslant cr \|\sup_{1\leqslant k\leqslant N} |v_k|\,\|_{L_r} \|\sum_{k=1}^N \varepsilon_k x_k\|_{L_1^X} + \sup_{1\leqslant k\leqslant N} \|x_k\|_X 2(1+r) \|\sup_{1\leqslant k\leqslant N} |d_k|\|_{L_r} \\ &\leqslant [4cr + 2(1+r)] \|\sup_{1\leqslant k\leqslant N} |d_k|\|_{L_r} \|\sum_{k=1}^N \varepsilon_k x_k\|_{L_1^X}. \blacksquare\end{aligned}$$

COROLLARY 5.3. *Let $f = (d_k)_{k=0}^N \subset L_1(\Omega, \mathscr{F}, P)$ be a martingale difference sequence, $0 = t_0 < t_1 < \ldots < t_n = N$, and $(\alpha_k)_{k=1}^N$ be a sequence of positive reals such that $\sum_{k=t_{i-1}+1}^{t_i} \alpha_k^2 = 1$ for $i = 1, \ldots, n$. Then for all Banach spaces X, $x_1, \ldots, x_n \in X$, and all $1 \leqslant r < \infty$ we have*

$$(E \sup_{1\leqslant k\leqslant N} \|\sum_{i=1}^n [{}^{t_{i-1}}\Delta^{t_i\wedge k} f] x_i\|_X^r)^{1/r} \leqslant cr \left\|\sup_{1\leqslant k\leqslant N} \frac{|d_k|}{\alpha_k}\right\|_{L_r} \|\sum_{i=1}^n g_i x_i\|_{L_1^X},$$

where $g_1, \ldots, g_n$ is a sequence of independent standard Gaussian variables and $c > 0$ is an absolute constant.

Before proving the corollary let us note that

$$\sup_{1\leqslant i\leqslant n}\Big(\sum_{k=t_{i-1}+1}^{t_i}|d_k|^2\Big)^{1/2}=\sup_{1\leqslant i\leqslant n}\Big(\sum_{k=t_{i-1}+1}^{t_i}\frac{|d_k|^2}{\alpha_k^2}\alpha_k^2\Big)^{1/2}$$
$$\leqslant \sup_{1\leqslant i\leqslant n}\Big[\sup_{t_{i-1}<k\leqslant t_i}\frac{|d_k|}{\alpha_k}\Big]\Big(\sum_{k=t_{i-1}+1}^{t_i}\alpha_k^2\Big)^{1/2}=\sup_{1\leqslant k\leqslant N}\frac{|d_k|}{\alpha_k}.$$

Hence Corollary 5.3 is closely related to Corollary 4.2.

Proof of Corollary 5.3. We apply Theorem 5.1 for the martingale difference sequence $((d_k/\alpha_k)\,y_k)_{k=1}^N$, where $y_k := \alpha_k x_l$ for $t_{l-1}<k\leqslant t_l$ and observe that

$$\Big\|\sum_{k=1}^N \varepsilon_k y_k\Big\|_{L_1^X}\leqslant\sqrt{\frac{\pi}{2}}\Big\|\sum_{k=1}^N g_k y_k\Big\|_{L_1^X}=\sqrt{\frac{\pi}{2}}\Big\|\sum_{l=1}^n g_l x_l\Big\|_{L_1^X}. \blacksquare$$

Remark 5.4. Assume the martingale difference sequence from Lemma 5.2 to be a Walsh–Paley martingale difference sequence $(d_k)_{k=0}^N\subset L_1(D_N)$ so that $d_k=\varepsilon_k v_k$ $(1\leqslant k\leqslant N)$ for some predictable sequence $(v_k)_{k=1}^N\subset L_1(D_N)$, where $D_N=\{-1, 1\}^N$ is the Cantor group equipped with the Haar measure and the filtration generated by the coordinates. Then $\sum_{k=1}^N\varepsilon_k x_k\to\sum_{k=1}^N d_k x_k$ turns into a UMD-transform $\sum_{k=1}^N D_k\to\sum_{k=1}^N v_k D_k$, where $D_k:=\varepsilon_k x_k$. Using this interpretation Lemma 5.2 states the following: Among all transforms $\sum_{k=1}^N D_k\to\sum_{k=1}^N v_k D_k$ with $\sup_{1\leqslant k\leqslant N}\|v_k\|_{L_\infty}=1$ the deterministic transforms $v_k=\theta_k$ $(\theta_k\in\{-1, 1\})$ are the extreme ones. This is closely related to a general fact about UMD-transforms, proved by Burkholder in [7] (Lemma A.1) and [6] (Lemma 2.1).

The example below shows that it is not sufficient to consider a symmetrized inequality

$$\Big(E_\varepsilon\int_\Omega\sup_{1\leqslant k\leqslant N}\Big\|\sum_{l=1}^k\varepsilon_l d_l(\omega)x_l\Big\|_X^r dP(\omega)\Big)^{1/r}\leqslant c_r\Big\|\sup_{1\leqslant k\leqslant N}|d_k|\Big\|_{L_r}\Big\|\sum_{k=1}^N\varepsilon_k x_k\Big\|_{L_1^X}$$

to get the assertion of Theorem 5.1.

EXAMPLE 5.5. *There is a constant* $c>0$ *such that for all* $N=2^m-1$ $(m=1, 2, 3, \ldots)$ *there is a Banach space* X, $(x_k)_{k=1}^N\subset B_X$, *and* $(d_k)_{k=0}^N\in\mathscr{P}((\mathscr{F}_k)_{k=0}^N)$ *with*

(1) $\|d_k\|_{L_\infty}\leqslant 1$,

(2) $\|\sum_{k=1}^N d_k x_k\|_X=m$ *a.s.*,

(3) $(E_\varepsilon\int_\Omega\|\sum_{k=1}^N\varepsilon_k d_k(\omega)x_k\|_X^r dP(\omega))^{1/r}\leqslant c\sqrt{rm}$ *for* $1\leqslant r<\infty$,

(4) $\|\sum_{k=1}^N g_k x_k\|_{L_1^X}\leqslant cm$,

where $(\varepsilon_k)_{k=1}^N$ *is a sequence of independent Rademacher variables. Consequently,*

(17)
$$\frac{\sqrt{m}}{c\sqrt{r}}\Big(E_\varepsilon\int_\Omega\Big\|\sum_{k=1}^N\varepsilon_k d_k(\omega)x_k\Big\|_X^r dP(\omega)\Big)^{1/r}$$
$$\leqslant\Big\|\sum_{k=1}^N d_k x_k\Big\|_{L_r^X}\sim\Big\|\sup_{1\leqslant k\leqslant N}|d_k|\Big\|_{L_r}\Big\|\sum_{k=1}^N\varepsilon_k x_k\Big\|_{L_1^X}.$$

Proof. Let $(H_k)_{k=0}^N \subset l_\infty^{2^m}$ be the sequence of 'discrete' Haar-functions normalized with respect to $l_\infty^{2^m}$ and starting with $H_0 = (1, \ldots, 1)$ and $H_1 = (1, \ldots, 1, -1, \ldots, -1)$. Furthermore, let $X := l_\infty^{2^m}$, $x_k := H_k$, $\Omega := \{1, \ldots, 2^m\}$ equipped with the measure $P(\{\omega\}) := 2^{-m}$, and let $d_k := H_k$. Finally, let $(\mathscr{F}_k)_{k=0}^N$ be the filtration on Ω generated by $(H_k)_{k=0}^N$. Now (1) is evident. (2) follows from

$$\Big\|\sum_{k=1}^N d_k(\omega)\, x_k\Big\|_X = \Big\|\sum_{k=1}^N H_k(\omega)\, H_k\Big\|_{l_\infty^{2^m}} = \Big|\sum_{k=1}^N H_k(\omega)\, H_k(\omega)\Big| = m.$$

To prove (3) let $\sigma_m\colon l_1^{2^m} \to l_\infty^{2^m}$ be the operator of summation and let $x_k^0 \in l_1^{2^m}$ be such that $\sigma_m x_k^0 = x_k$ and $\|x_k^0\|_{l_1^{2^m}} \leqslant 4$. Now it is known that the operator of summation $\sigma\colon l_1 \to l_\infty$ has type 2 (according to [19] and [20], σ even factors through a Banach space which is of type 2), which means that there is a constant $c_1 > 0$ such that for all finite sequences $(y_i)_i \subset l_1$ we have

$$\Big(\boldsymbol{E}_\varepsilon \Big\|\sum_i \varepsilon_i \sigma y_i\Big\|_{l_\infty}^2\Big)^{1/2} \leqslant c_1 \Big(\sum_i \|y_i\|_{l_1}^2\Big)^{1/2}.$$

We get, for all $\omega \in \Omega$, by the Khintchine–Kahane inequality for the Rademacher averages (see [23], Theorem 4.7),

$$\begin{aligned}
\Big(\boldsymbol{E}_\varepsilon \Big\|\sum_{k=1}^N \varepsilon_k d_k(\omega)\, x_k\Big\|_{l_\infty^{2^m}}^r\Big)^{1/r} &= \Big(\boldsymbol{E}_\varepsilon \Big\|\sigma_m\Big(\sum_{k=1}^N \varepsilon_k d_k(\omega)\, x_k^0\Big)\Big\|_{l_\infty^{2^m}}^r\Big)^{1/r} \\
&\leqslant c_0 \sqrt{r} \Big(\boldsymbol{E}_\varepsilon \Big\|\sigma_m\Big(\sum_{k=1}^N \varepsilon_k d_k(\omega)\, x_k^0\Big)\Big\|_{l_\infty^{2^m}}^2\Big)^{1/2} \\
&\leqslant c_0 c_1 \sqrt{r}\, \Big(\sum_{k=1}^N |d_k(\omega)|^2\, \|x_k^0\|_{l_1^{2^m}}^2\Big)^{1/2} \\
&\leqslant 4 c_0 c_1 \sqrt{r}\, \Big(\sum_{k=1}^N |d_k(\omega)|^2\Big)^{1/2} = 4 c_0 c_1 \sqrt{rm}.
\end{aligned}$$

Integrating with respect to ω we obtain assertion (3). Finally, let us show (4). From [24] we get

$$\begin{aligned}
\boldsymbol{E}\Big\|\sum_{k=1}^N g_k x_k\Big\| &\leqslant c_2 \sqrt{m} \sup_{a \in B_{l_1^{2^m}}} \Big(\sum_{k=1}^N |\langle x_k, a\rangle|^2\Big)^{1/2} \\
&= c_2 \sqrt{m} \sup_{i=1,\ldots,2^m} \Big(\sum_{k=1}^N |\langle x_k, e_i\rangle|^2\Big)^{1/2} = c_2 m,
\end{aligned}$$

where $(e_i)_{i=1}^{2^m}$ is the unit vector basis of $l_1^{2^m}$. Concerning the $\sim$ part of (17), the relation $\prec$ follows for example from Theorem 5.1 whereas $\succ$ is a consequence of (1), (2), and (4). ∎

Acknowledgment. I would like to thank S. Kwapień, who proposed the usage of Theorem 5.1.2 from [22] to prove Theorem 5.1.

REFERENCES

[1] P. Assouad, *Applications sommantes et radonifiantes*, Ann. Inst. Fourier 22 (3) (1972), pp. 81–93.

[2] N. L. Bassily and J. Mogyoródi, *On the BMO_Φ-spaces with general Young function*, Ann. Univ. Sci. Budapest Eötvös, Sect. Math. 27 (1984), pp. 215–227.

[3] C. Bennet and R. Sharpley, *Interpolation of Operators*, Academic Press, 1988.

[4] J. Bourgain, *On the behaviour of the constant in the Littlewood–Paley inequality*, in: *Israel Seminar, GAFA 1987/88*, J. Lindenstrauss and V. D. Milman (Eds.), Lecture Notes in Math. 1376, Springer, 1989, pp. 202–208.

[5] D. L. Burkholder, *Distribution function inequalities for martingales*, Ann. Probab. 1 (1973), pp. 19–42.

[6] — *Boundary value problems and sharp inequalities for martingale transforms*, ibidem 12 (1984), pp. 647–702.

[7] — *Explorations in martingale theory and its applications*, in: *Ecole d'Eté de Probabilités de Saint-Flour XIX-1989*, Lecture Notes in Math. 1464, Springer, 1991, pp. 1–66.

[8] — B. J. Davis and R. F. Gundy, *Integral inequalities for convex functions of operators on martingales*, in: *Proceedings of the Sixth Berkeley Symposium*, Math. Statist. Probab. 2 (1972), pp. 223–240.

[9] D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasilinear operators on martingales*, Acta Math. 124 (1970), pp. 249–304.

[10] S. Chang, M. Wilson and J. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helv. 60 (1985), pp. 217–246.

[11] S. Chevet, *Séries de variables aléatoires gaussiens à valeurs dans $E\hat{\otimes}_\varepsilon F$. Applications aux produits d'espaces de Wiener abstraits*, Séminaire Maurey–Schwartz, Exposé XIX, 1977/78.

[12] B. Davis, *On the integrability of the martingale square function*, Israel J. Math. 8 (1970), pp. 187–190.

[13] S. Geiss, *BMO_ψ-spaces and applications to extrapolation theory*, Studia Math. 122 (3) (1997), pp. 235–274.

[14] — and M. Junge, *Type and cotype with respect to arbitrary orthonormal systems*, J. Approx. Theory 82 (3) (1995), pp. 399–433.

[15] P. Hitczenko, *Upper bound for the L_p-norms of martingales*, Probab. Theory Related Fields 86 (1990), pp. 225–238.

[16] — *Domination inequality for martingale transforms of a Rademacher sequence*, Israel J. Math. 84 (1993), pp. 161–178.

[17] J. Hoffmann-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Math. 52 (1974), pp. 159–186.

[18] — *Probability in Banach spaces*, in: *Ecole d'Eté de Probabilités de Saint-Flour VI-1976*, Lecture Notes in Math. 598, Springer, 1977, pp. 1–186.

[19] R. C. James, *Super-reflexive Banach spaces*, Canad. J. Math. 5 (1972), pp. 896–904.

[20] — *Non-reflexive spaces of type* 2, Israel J. Math. 30 (1978), pp. 1–13.

[21] J.-P. Kahane, *Some random series of functions*, in: *Heath Mathematical Monographs*, 2nd edition, Cambridge Univ. Press, 1985.

[22] S. Kwapień and W. A. Woyczyński, *Random series and stochastic integrals: Single and multiple*, in: *Probability and Its Applications*, Birkhäuser, Boston–Basel–Berlin 1992.

[23] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, 1991.

[24] W. Linde and A. Pietsch, *Mappings of Gaussian cylindrical measures in Banach spaces*, Theory Probab. Appl. 19 (1974), pp. 445–460.

[25] R. Sh. Liptser and A. N. Shiryayev, *Theory of Martingales*, Kluwer Academic Publishers, Dordrecht–Boston–London 1989.

[26] B. Maurey et G. Pisier, *Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach*, Studia Math. 58 (1976), pp. 45–90.

[27] A. Pietsch, *Operator Ideals*, North-Holland, Amsterdam 1980.
[28] – and J. Wenzel, *Orthonormal Systems and Banach Space Geometry*, Cambridge Univ. Press (to appear).
[29] G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. 20 (1975), pp. 326–350.
[30] M. Talagrand, *Regularity of Gaussian processes*, Acta Math. 159 (1987), pp. 99–149.
[31] N. Tomczak-Jaegermann, *Banach–Mazur Distances and Finite-dimensional Operator Ideals*, Pitman, New York 1989.
[32] G. Wang, *Some Sharp Inequalities for Conditionally Symmetric Martingales*, PhD Thesis, University of Illinois, 1989.

Mathematisches Institut der Friedrich-Schiller-Universität Jena
Postfach, D-07740 Jena, Germany
E-mail: geiss@minet.uni-jena.de

Received on 20.1.1997;
revised version on 12.8.1997