References

[1] K. Ball and A. Pajor, *Convex bodies with few faces*, Proc. Amer. Math. Soc. 110 (1990), 225–231.
[2] W. Banaszczyk, *Balancing vectors and convex bodies*, Studia Math. 106 (1993), 93–100.
[3] —, *A Beck–Fiala-type theorem for Euclidean norms*, European J. Combin. 11 (1990), 497–500.
[4] W. Banaszczyk and S. J. Szarek, *Lattice coverings and Gaussian measures of n-dimensional convex bodies*, Discrete Comput. Geom., to appear.
[5] I. Bárány and V. S. Grinberg, *On some combinatorial questions in finite-dimensional spaces*, Linear Algebra Appl. 41 (1981), 1–9.
[6] J. Beck and T. Fiala, *Integer-making theorems*, Discrete Appl. Math. 3 (1981), 1–8.
[7] E. D. Gluskin, *Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces*, Math. USSR-Sb. 64 (1989), 85–96.
[8] B. S. Kashin, *An analogue of Menshov's "correction" theorem for discrete orthonormal systems*, Mat. Zametki 46 (6) (1989), 67–74 (in Russian); English transl.: Math. Notes 46 (5-6) (1989), 934–939.
[9] D. Kleitman, *On a combinatorial conjecture of Erdős*, J. Combin. Theory Ser. A 1 (1966), 209–214.
[10] Z. Sidák, *On multivariate normal probabilities of rectangles: their dependence on correlation*, Ann. Math. Statist. 39 (1968), 1425–1434.
[11] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. 289 (1985), 679–706.
[12] —, *Balancing vectors in the max norm*, Combinatorica 6 (1) (1986), 55–65.
[13] —, *Ten Lectures on the Probabilistic Method*, SIAM, Philadelphia, Penn., 1994.
[14] S. J. Szarek and E. Werner, *Confidence regions for means of multivariate normal distributions and a non-symmetric correlation inequality for Gaussian measure*, preprint.

Department of Mathematics
University of Crete
Iraklion, Crete
Greece
E-mail: deligia@talos.cc.uch.gr

Received October 25, 1995
Revised version June 19, 1996 and September 18, 1996 (3553)

STUDIA MATHEMATICA 122 (3) (1997)

BMO_ψ-spaces and applications to extrapolation theory

by

STEFAN GEISS (Jena)

Abstract. We investigate a scale of BMO_ψ-spaces defined with the help of certain Lorentz norms. The results are applied to extrapolation techniques concerning operators defined on adapted sequences. Our extrapolation works simultaneously with two operators, starts with BMO_ψ-L_∞-estimates, and arrives at L_p-L_p-estimates, or more generally, at estimates between K-functionals from interpolation theory.

Introduction. Extrapolation techniques are an important tool to compare L_p-norms of operators defined on martingales. Basic results were proved by D. L. Burkholder, B. J. Davis, and R. F. Gundy ([11], [10]).

Let us consider a basic example. Assume $f = (d_k)_{k=0}^n \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ to be a sequence of martingale differences with respect to some filtration $(\mathcal{F}_k)_{k=0}^n$ such that $d_0 = 0$ and $|d_k|$ is $\mathcal{F}_{k-1}$-measurable for $k = 1, \ldots, n$. An extension of the Azuma inequality proved by P. Hitczenko [16] (Lemma 4.3) (see also [12], [25], the comments in [16], and the proof of Proposition 1 of this paper) says that for $\lambda > 0$,

$$\mathbb{P}\Big(\Big|\sum_{k=1}^n d_k\Big| > \lambda \|S_2 f\|_\infty\Big) \le 2e^{-\lambda^2/2} \tag{1}$$

where $S_2 f := (\sum_{k=1}^n |d_k|^2)^{1/2}$ is the usual square function operator. The above inequality is of importance for several reasons. For example, in [16] (Corollary 4.2) this inequality is used to prove the Burkholder–Davis–Gundy type inequality

$$\Big\|\sup_{1\le k\le n}\Big|\sum_{i=1}^k d_i\Big|\Big\|_p \le c\sqrt{p}\,\Big\|\Big(\sum_{i=1}^n |d_i|^2\Big)^{1/2}\Big\|_p \quad \text{for } 1 \le p < \infty, \tag{2}$$

which extends the corresponding one for dyadic martingales ([12], [6], [25]). In order to deduce (2) from (1) one has to modify (1) in two steps. First we observe that for $B \in \mathcal{F}_k$ we get a martingale difference sequence $(d_i)_{i=k+1}^l \subset L_1(B, \mathbb{P}_B)$, where $\mathbb{P}_B$ is the normalized restriction of $\mathbb{P}$ to B. Applying (1) to

1991 *Mathematics Subject Classification*: 46B70, 46E30, 47H99, 60G99.

this restricted sequence and taking into account $\|S_2((d_i)_{i=k+1}^l)\|_{L_\infty(B,\mathbb{P}_B)} \le \|S_2 f\|_\infty$ we arrive at the apparently stronger estimate

$$\sup_{0\le k\le l\le n}\ \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \mathbb{P}_B\Big(\Big|\sum_{i=k+1}^{l} d_i\Big| > \lambda\|S_2f\|_\infty\Big) \le 2e^{-\lambda^2/2}. \tag{3}$$

Secondly, analyzing [16] more carefully, we see that we additionally have to replace $\sum_{i=k+1}^l d_i$ by $\sum_{i=k}^l d_i$, which is possible by a change of some constants. This new inequality can be written in the language of BMO-spaces, introduced in Section 1, as

$$\Big\|\Big(\sum_{l=0}^{k} d_l\Big)_{k=0}^{n}\Big\|_{\mathrm{BMO}_{\psi_2}} \le c_0\|S_2f\|_\infty \tag{4}$$

where $\psi_2(t) = t^2$. The step from (4) to (2) is usually called *extrapolation*.

There is another reason to consider inequality (1). Namely, we get a sub-Gaussian tail behaviour of $f_n = \sum_{k=1}^n d_k$ whenever we can control the L_∞-norm of S_2f. It is obviously easier to check $\|S_2f\|_\infty$ than the sub-Gaussian behaviour of f_n itself. But there are situations in which (1) does not yield the asymptotically exact result. For example, if we take independent Rademacher variables $\varepsilon_1,\dots,\varepsilon_n$ such that $\mathbb{P}(\varepsilon_k = 1) = \mathbb{P}(\varepsilon_k = -1) = 1/2$ and define

$$d_k := \varepsilon_k\frac{\varepsilon_{k-1}-1}{2}\cdots\frac{\varepsilon_1-1}{2}\quad \text{for } k = 2,\dots,n,$$

then $\|S_2f\|_\infty = \sqrt{n-1}$ and for $\lambda > 0$ inequality (1) yields

$$\mathbb{P}\Big(\Big|\sum_{k=2}^n d_k\Big| > \lambda\Big) \le 2e^{-\lambda^2/(2(n-1))}$$

whereas

$$\frac{1}{16}e^{-\lambda} \le \mathbb{P}\Big(\Big|\sum_{k=2}^n d_k\Big| > \lambda\Big) \le \frac14 e^{-(\log 2)\lambda} \tag{5}$$

follows by a direct computation for $2 \le \lambda \le n-4$. For instance, if $\lambda = \sqrt{n}$, then $2e^{-\lambda^2/(2(n-1))} \sim 1$ but $\frac14 e^{-(\log 2)\lambda} \sim 2^{-\sqrt{n}}$.

In view of the above example one can ask for possible improvements of (1). One possibility is to use (2) as described in Remark 6.4(1) below. We will go another, and from our point of view more natural, way. To this end let us recall that for a compatible couple (X_0, X_1) of Banach spaces, $x \in X_0 + X_1$, and $t \ge 0$ the *K-functional* is defined by

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

In the case $X_0 = L_q(\Omega)$ and $X_1 = L_p(\Omega)$ we will shorten $K_{q,p}[f,t] := K(f,t;L_q,L_p)$. Now looking at the proof of (1) presented in [16], or at [20] (p. 31), we realize that one can deduce more.

PROPOSITION 1. *There is an absolute constant $c \ge 1$ such that for all martingale difference sequences $f = (d_k)_{k=0}^n \subset L_1(\Omega,\mathcal{F},\mathbb{P})$ such that $d_0 = 0$ and $|d_k|$ is $\mathcal{F}_{k-1}$-measurable for $k = 1,\dots,n$ one has for $t \ge 1$,*

$$K_{\infty,1}[f_n,t] \le c\sqrt{1+\log t}\,K_{\infty,2}[S_2f,\sqrt{t}] \tag{6}$$

where $f_k := \sum_{i=1}^k d_i$. In particular, for all $\lambda \ge 1$ one has

$$\mathbb{P}(|f_n| > c\lambda K_{\infty,2}[S_2f, e^{(\lambda^2-1)/2}]) \le e^{1-\lambda^2}.$$

Proof. First observe that for $1 \le k \le n$ and $\mu > 0$ one has for $d_k(\omega) \ne 0$ (cf. [20] (p. 31))

$$e^{\mu d_k(\omega)} \le e^{\mu^2 d_k(\omega)^2/2} + \frac{d_k(\omega)}{|d_k(\omega)|}\sinh(\mu|d_k(\omega)|)$$

so that

$$\mathbb{E}_{k-1}e^{\mu d_k} \le e^{\mu^2 d_k^2/2}.$$

Iteration with respect to k gives $\mathbb{E}e^{\mu f_n - (\mu^2/2)(S_2f)^2} \le 1$. Since this inequality remains true for $-f_n$ instead of f_n we can continue to get $\mathbb{E}e^g \le 3$ where $g := (\mu|f_n| - (\mu^2/2)(S_2f)^2) \vee 0$, and by standard arguments we obtain $K_{\infty,1}[g,t] \le c(1+\log t)$ for $t \ge 1$. Hence

$$\mu K_{\infty,1}[|f_n|,t] - \frac{\mu^2}{2}K_{\infty,1}[(S_2f)^2,t] \le K_{\infty,1}[g,t] \le c(1+\log t).$$

If $K_{\infty,1}[(S_2f)^2,t] > 0$ and $\mu^2 := 2c(1+\log t)/K_{\infty,1}[(S_2f)^2,t]$, then

$$\begin{aligned}K_{\infty,1}[|f_n|,t] &\le \sqrt{2c}\sqrt{1+\log t}\sqrt{K_{\infty,1}[(S_2f)^2,t]}\\ &\le \sqrt{2c}\sqrt{1+\log t}\,K_{\infty,2}[S_2f,\sqrt{t}]\end{aligned}$$

where one can use (18) in the last inequality. ∎

For the example considered in (5) one has $K_{\infty,2}[S_2f, e^{(\lambda^2-1)/2}] \le c_1\lambda$ for $\lambda \ge 1$ and some $c_1 \ge 1$. Hence the above proposition and the lower estimate of (5) imply for $2 \vee (cc_1)^2 \le \lambda \le n-4$,

$$\begin{aligned}\frac{1}{16}e^{-\lambda} &\le \mathbb{P}\Big(\Big|\sum_{k=2}^n d_k\Big| > \lambda\Big)\\ &\le \mathbb{P}\Big(\Big|\sum_{k=2}^n d_k\Big| > c\frac{\sqrt{\lambda}}{cc_1}K_{\infty,2}[S_2f, e^{(\lambda/(cc_1)^2-1)/2}]\Big) \le ee^{-\lambda/(cc_1)^2}.\end{aligned}$$

In this paper we will show in Theorem 1.7 that, using some general assumptions, one can deduce from (4) inequalities of the form (6). For example, it turns out that by a combination of Theorem 1.7, Proposition 7.3 with $C = 0$, and (4) one gets

$$K_{\infty,p}[\sup_k |f_k|, t^{1/p}] \leq c\sqrt{p}\sqrt{1+\log t}\, K_{\infty,p}[S_2 f, t^{1/p}] \quad (1 \leq p < \infty), \tag{7}$$

which is slightly stronger than (6) in the case $p = 1$. In this way we obtain inequalities which simultaneously give (2) (by setting $t = 1$) and improve (1). Our starting point will be an extrapolation result of P. Hitczenko recalled in Theorem 1.1 of this paper. Reformulating the assumptions of this statement we introduce a scale of BMO_ψ-spaces. A separate investigation of these spaces is one of the main subjects of this paper and leads to further development of Hitczenko's result.

1. Notation and main results. Throughout this paper $[\Omega, \mathcal{F}, \mathbb{P}]$ stands for a probability space. If X is a real Banach space equipped with the Borel σ-algebra $\mathcal{B}(X)$, then $L_0^X(\Omega, \mathcal{F}, \mathbb{P})$ is the linear space of all measurable $f : [\Omega, \mathcal{F}, \mathbb{P}] \to [X, \mathcal{B}(X)]$ which are Radon (cf. [20]). Moreover, $L_0^+(\Omega, \mathcal{F}, \mathbb{P}) := \{f \in L_0^{\mathbb{R}}(\Omega, \mathcal{F}, \mathbb{P}) \mid f \geq 0 \text{ a.s.}\}$ and

$$L_p^X(\Omega, \mathcal{F}, \mathbb{P}) := \{f \in L_0^X(\Omega, \mathcal{F}, \mathbb{P}) \mid \|f\|_{L_p^X} := \|\|f(\omega)\|_X\|_{L_p(\Omega)} < \infty\}$$

($1 \leq p \leq \infty$). If $[\Omega, \mathcal{F}, \mathbb{P}]$ is equipped with a filtration $(\mathcal{F}_k)_{k\in I}$, where $I = \{0, \ldots, n\}$ or $I = \mathbb{N} = \{0, 1, 2, \ldots\}$, and $\bigvee_{k\in I} \mathcal{F}_k = \mathcal{F}$, then we use the linear spaces (under the coordinatewise operations) of *adapted sequences*

$$\mathcal{A}^X((\mathcal{F}_k)_{k\in I}) := \{(d_k)_{k\in I} \mid d_k \in L_0^X(\Omega, \mathcal{F}_k, \mathbb{P}) \text{ for all } k \in I\}$$

and $\mathcal{A}((\mathcal{F}_k)_{k\in I}) := \mathcal{A}^{\mathbb{R}}((\mathcal{F}_k)_{k\in I})$. For stopping times σ, τ and $f = (d_k)_{k\in I} \in \mathcal{A}^X((\mathcal{F}_k)_{k\in I})$ one usually sets

$${}^\sigma f^\tau := (d_k \chi_{\{\sigma<k\leq\tau\}})_{k\in I}, \quad f^\tau := (d_k \chi_{\{k\leq\tau\}})_{k\in I}, \quad {}^\sigma f := (d_k \chi_{\{\sigma<k\}})_{k\in I}.$$

We will consider operators

$$T : \mathcal{A}^X((\mathcal{F}_k)_{k\in I}) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P}),$$

where we make the following *conventions* valid in the whole paper:

(1) $d_0 = 0$, ${}^k f, f^k, -f \in E$ for all $f = (d_k)_{k\in I} \in E$ and $k \in I$.
(2) $T((0,0,\ldots)) = 0$ a.s. and $T(f) = T(-f)$ a.s. for all $f \in E$.

The operator T is called *quasilinear* if there is some $\gamma_T > 0$ such that for all $f, g \in E$ with $f + g \in E$ one has $T(f+g) \leq \gamma_T[Tf + Tg]$ a.s. If $\gamma_T = 1$, then T is called *sublinear*. The operator T is *monotone* if $Tf^k \leq Tf$ a.s., and *measurable* if Tf^k is $\mathcal{F}_k$-measurable for all $f \in E$ and $k \in I$. The operator T is called *local* if for all $f = (d_k)_{k\in I} \in E$ one has $Tf = 0$ a.s. on the set

$$\{0 = \mathbb{E}(\|d_1\| \mid \mathcal{F}_0) = \ldots = \mathbb{E}(\|d_k\| \mid \mathcal{F}_{k-1}) = \ldots\}.$$

The maximal operator T^* of T is given by $T^*f := \sup_{k\in I} Tf^k$. Finally, we will denote the set of all increasing bijections $\psi : [1, \infty) \to [1, \infty)$ by $\mathcal{D}$.

Now let us describe (in a slightly different form) the result of P. Hitczenko we are starting from. Recall that a sequence $(v_k)_{k=1}^\infty \subset L_0(\Omega, \mathcal{F}, \mathbb{P})$ is *predictable* whenever v_k is $\mathcal{F}_{k-1}$-measurable.

THEOREM 1.1 (Hitczenko [16]). *Let $\psi \in \mathcal{D}$ and $E \subseteq \mathcal{A}((\mathcal{F}_k)_{k\in\mathbb{N}})$ be the subset of the martingale difference sequences with $d_0 = 0$. Assume $T : E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$ to be quasilinear, local, measurable, and monotone. Suppose $f = (d_k)_{k=0}^\infty \in E$ and $(v_k)_{k=1}^\infty \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ to be predictable with $|d_k| \vee T({}^{k-1}f^k) \leq v_k$ a.s. for $k = 1, 2, \ldots$ If for all $0 \leq k \leq l$, $B \in \mathcal{F}_k$, all stopping times σ, τ, and all $\lambda \geq 1$ one has*

$$\mathbb{P}\Big(\Big\{\Big|\sum_{k\vee\sigma<i\leq\tau\wedge l} d_i\Big| > \lambda\|\sup_{\sigma<i\leq\tau} v_i \vee T({}^\sigma f^\tau)\|_\infty\Big\} \cap B\Big) \leq e^{1-\psi(\lambda)}\mathbb{P}(B), \tag{8}$$

then one gets for $1 \leq p < \infty$ and some constant $c > 0$, depending on T and ψ only,

$$\Big\|\sup_n \Big|\sum_{i=0}^n d_i\Big|\Big\|_p \leq c\psi^{-1}(p)\|\sup_k v_k \vee Tf\|_p. \tag{9}$$

As in (3)$\Rightarrow$(4) of the introduction we translate (8) into the language of BMO-spaces. To this end we introduce

$$\mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I}) := \{(f_k)_{k\in I} \in \mathcal{A}((\mathcal{F}_k)_{k\in I}) \mid \|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} < \infty\}$$

with

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} := \inf\{c > 0 \mid \sup_{\substack{0\leq k\leq l\\ k,l\in I}} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \mathbb{P}_B(|f_l - f_{k-1}| > \lambda) \leq e^{1-\psi(\lambda/c)} \text{ for } \lambda \geq c\},$$

where $\mathbb{P}_B = \mathbb{P}/\mathbb{P}(B)$ is the normalized restriction of $\mathbb{P}$ to B and $f_{-1} = 0$. In view of $d_k \leq v_k$ the relation (8) implies

$$\Big\|\Big(\sum_{i=0}^k d_i \chi_{\{\sigma<i\leq\tau\}}\Big)_{k=0}^\infty\Big\|_{\mathrm{BMO}_\psi} \leq 2\|\sup_{\sigma<i\leq\tau} v_i \vee T({}^\sigma f^\tau)\|_\infty. \tag{10}$$

Conversely, (10) implies (8) if we replace $\sup_{\sigma<i\leq\tau} v_i \vee T({}^\sigma f^\tau)$ by $3\sup_{\sigma<i\leq\tau} v_i \vee T({}^\sigma f^\tau)$ in (8).

Using the properties of BMO_ψ-spaces we will extend the extrapolation principle (10)$\Rightarrow$(9) in Theorem 1.7, Corollary 6.3, and Corollary 7.12.

The BMO_ψ-spaces are introduced and investigated in Sections 4 and 5 whereas the required material about the Lorentz spaces M_φ, which are the basic modules of BMO_ψ-spaces, is summarized in Section 2. Since we carry out the extrapolation in Theorem 1.7 with two operators A and B

we define BMO_ψ-norms for adapted sequences, instead of for martingale sequences only. Here we follow A. M. Garsia [14] (p. 66), who extended the classical BMO-norm for functions to the case of adapted sequences. First one should ask which weight functions ψ generate the same BMO_ψ-spaces. We completely solve this problem by a regularization $\overline{\psi}$ of ψ defined as follows.

DEFINITION 1.2. For $\psi \in \mathcal{D}$ the function $\overline{\psi} : [1,\infty) \to [1,\infty)$ is given by

$$\overline{\psi}(\mu) := \sup\Big\{ \sum_{i=1}^{N}(\psi(\mu_i)-1)+1 \;\Big|\; \mu = \sum_{i=1}^{N}\mu_i,\ \mu_i \ge 1,\ N=1,2,\ldots\Big\}.$$

It is clear that $\psi \le \overline{\psi}$. In Lemma 4.4 we see $\overline{\psi} \in \mathcal{D}$ (we show the stronger $\overline{\varphi} \in \mathcal{C}_\Delta$, see (20) and Definition 2.1) and that, for some $c_\psi \ge 1$, one has

$$\mu\overline{\psi}(\lambda) \le c_\psi\overline{\psi}(\mu\lambda) \quad \text{for } \mu,\lambda \ge 1.$$

In particular, $\overline{\psi}(a_\psi\mu) \ge \mu$ for $\mu \ge 1$ if we choose λ to be $a_\psi := \overline{\psi}^{-1}(c_\psi)$, which yields the John–Nirenberg Theorem in a version of A. M. Garsia (see Example 4.3 and Corollary 4.8). In Theorem 4.6 we will show

$$(11) \qquad \|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} \le \|(f_k)_{k\in I}\|_{\mathrm{BMO}_{\overline{\psi}}} \le 6\psi^{-1}(3)\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi}.$$

Moreover, in Section 4 we prove

THEOREM 1.3. *For $\psi,\psi' \in \mathcal{D}$ the following assertions are equivalent.*

(1) *There exists $c_1 \ge 1$ such that $\overline{\psi}(\mu) \le \overline{\psi'}(c_1\mu)$ for all $1 \le \mu < \infty$.*

(2) *There exists $c_2 > 0$ such that for all $[\Omega,\mathcal{F},\mathbb{P};(\mathcal{F}_k)_{k\in I}]$ and all adapted sequences $(f_k)_{k\in I} \in \mathrm{BMO}_{\psi'}((\mathcal{F}_k)_{k\in I})$,*

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} \le c_2\|(f_k)_{k\in I}\|_{\mathrm{BMO}_{\psi'}}.$$

We have seen that the regularization $\overline{\psi}$ of ψ is the right tool to investigate the BMO_ψ-spaces themselves. The consequences for the extrapolation (10)⇒(9) are as follows: Starting with (10) we can switch via (11) from ψ to $\overline{\psi}$. Extrapolation gives (9) with $\overline{\psi}^{-1}(p)$ instead of $\psi^{-1}(p)$. In the case $\psi(\mu) = \mu^q$ $(1 \le q < \infty)$ this does not yield an improvement (see Example 4.3). But, for instance, if $\liminf_{\mu\to\infty}\psi(\mu)/\mu = 0$ then $\overline{\psi}^{-1}(p)$ with $\overline{\psi}^{-1}(p) \le a_\psi p$ (which follows from $\overline{\psi}(a_\psi\mu) \ge \mu$) is better than $\psi^{-1}(p)$.

In Section 5 we consider convexity properties of BMO_ψ-spaces. We modify the upper p-estimates, known for Banach lattices, to the following weak upper estimates.

DEFINITION 1.4. If $\psi \in \mathcal{D}$, then the BMO_ψ-spaces satisfy a *weak upper estimate* with respect to the sequence $a = (a_i)_{i\ge1}$ with $1 \ge a_1 \ge a_2 \ge \ldots \ge 0$ provided that there is a constant $c > 0$ such that for all $[\Omega,\mathcal{F},\mathbb{P};(\mathcal{F}_k)_{k\in I}]$ and $(f_k^{(i)})_{k\in I} \in \mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I})$,

$$\|(\sup_{i\ge1} a_i|f_k^{(i)}|)_{k\in I}\|_{\mathrm{BMO}_\psi} \le c\sup_{i\ge1}\|(f_k^{(i)})_{k\in I}\|_{\mathrm{BMO}_\psi}.$$

We set $\mathbf{U}_a(\mathrm{BMO}_\psi) := \inf c$ where $\mathbf{U}_a(\mathrm{BMO}_\psi) = \infty$ if such a $c > 0$ does not exist.

We will prove

THEOREM 1.5. *For $a_i := 1/\overline{\psi}^{-1}(1+\log i)$, $a := (a_i)_{i=1}^\infty$, and $b^N := (a_N,\ldots,a_N,0,0,\ldots)$, where a_N is repeated N times, one has*

$$0 < \inf_N \mathbf{U}_{b^N}(\mathrm{BMO}_\psi) \le \mathbf{U}_a(\mathrm{BMO}_\psi) < \infty.$$

In this way we have shown that $(1/\overline{\psi}^{-1}(1+\log i))_{i=1}^\infty$ is (up to multiplicative constants) the optimal sequence such that the BMO_ψ-spaces satisfy a weak upper estimate. Our interest in Theorem 1.5 is motivated by Theorem 1.7. In order to formulate this extrapolation principle we first replace the assumptions made in Theorem 1.1 by more abstract ones.

DEFINITION 1.6 Let X be a Banach space and let

$$A, B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega,\mathcal{F},\mathbb{P}).$$

We say that (E,A,B) has *property* (EP) [1] with constant $c \ge 1$ provided that A is measurable, and that for all $\lambda > 0$ and all $f \in E$ there is a $g \in E$ with

$$\frac{1}{c}\chi_{\{Bf\le\lambda\}}A^*f \le A^*g \le cA^*f \text{ a.s.} \quad\text{and}\quad Bg \le c\lambda \text{ a.s.}$$

Moreover, for every $\psi \in \mathcal{D}$ and $1 \le t < \infty$ we define the weight function $w_t^\psi \in L_1[0,1]$, where $[0,1]$ is equipped with the Lebesgue measure, by

$$w_t^\psi(s) := \begin{cases} 1/\psi^{-1}(1+\log(st)), & 1/t \le s \le 1,\\ 1, & 0 < s < 1/t,\end{cases}$$

so that

$$\frac{1}{\psi^{-1}(1+\log t)} \le w_t^\psi \le 1.$$

Furthermore, for a weight $w \in L_0^+[0,1]$, $1 \le p,q \le \infty$, $t \ge 0$, and $f \in L_q(\Omega) + L_p(\Omega)$ we use the following weighted K-functional:

$$K_{q,p}^w[f,t] := \inf\{\|g\|_q + t\|h\|_p \mid f(\omega)w(s) = g(\omega,s) + h(\omega,s) \text{ in } L_1(\Omega\times[0,1]),\ g \in L_q(\Omega\times[0,1]),\ h \in L_p(\Omega\times[0,1])\}.$$

In Section 6 we will prove

THEOREM 1.7 *Let $\psi \in D$ and $A, B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$ where X is a Banach space. Assume that (E,A,B) satisfies* (EP) *with constant $c \ge 1$ and that*

$$\|(Af^k)_{k=0}^n\|_{\mathrm{BMO}_\psi} \le \|Bf\|_\infty$$

---
(1) (EP) stands for "extrapolation property".

for all $f \in E$. *Then there is some* $d > 0$, *depending on* $\psi \in \mathcal{D}$ *and* $c \geq 1$ *only, such that for* $1 \leq p < \infty$, $t \geq 1$, $w_t := w_t^{\overline{\psi}}$, *and* $f \in E$,

$$K_{\infty,p}^{w_t}[A^*f, t^{1/p}] \leq d\overline{\psi}^{-1}(p) K_{\infty,p}[Bf, t^{1/p}]. \tag{12}$$

In particular, for $t = 1$ *one has* $\|A^*f\|_p \leq d\overline{\psi}^{-1}(p)\|Bf\|_p$.

Since $K_{\infty,p}[h_1, s] \leq K_{\infty,p}[h_2, s]$ for $0 \leq h_1 \leq h_2$, from (12) we get

$$K_{\infty,p}[A^*f, t^{1/p}] \leq d\overline{\psi}^{-1}(p)\overline{\psi}^{-1}(1 + \log t) K_{\infty,p}[Bf, t^{1/p}]. \tag{13}$$

In Remark 6.4(2) we give an example such that (12) is strictly better than (13). To prove (12) we show, with the help of the weak upper estimates, for $a_i = 1/\overline{\psi}^{-1}(1 + \log i)$,

$$\|\sup_{1 \leq i \leq N} a_i A^*f(\omega_i)\|_{L_p(\Omega^N)} \leq d'\overline{\psi}^{-1}(p)\|\sup_{1 \leq i \leq N} Bf(\omega_i)\|_{L_p(\Omega^N)}. \tag{14}$$

The step from (14) to (12) is carried out with the following result which will be deduced from a more general one in Section 3.

THEOREM 1.8. *Let* $1 \leq p < \infty$ *and* $f \in L_p(\Omega, \mathcal{F}, \mathbb{P})$. *Assume* $N \geq 1$ *and* $w \in L_0^+[0,1]$ *with* $w = \sum_{i=1}^N \alpha_i \chi_{[(i-1)/N, i/N)}$. *Then*

$$\tfrac{1}{2} K_{\infty,p}^w[f, N^{1/p}] \leq \|\sup_{1 \leq i \leq N} \alpha_i |f(\omega_i)|\|_{L_p(\Omega^N)} \leq K_{\infty,p}^w[f, N^{1/p}].$$

In particular, if $\alpha_1 = \ldots = \alpha_N = 1$, *then*

$$K_{\infty,p}[f, N^{1/p}] \sim \|\sup_{1 \leq i \leq N} |f(\omega_i)|\|_{L_p(\Omega^N)}.$$

In the last two sections we will consider the assumptions of Theorem 1.7, property (EP) and the $\|\cdot\|_{\mathrm{BMO}_\psi}$-$\|\cdot\|_\infty$-estimate more precisely.

In Section 7 we give two examples satisfying condition (EP). In the second one we demonstrate that one can reduce Theorem 1.1 to Theorem 1.7 with the help of a construction which starts with a quasilinear operator and produces an "equivalent" operator with properties similar to those of the p-norm with some $0 < p \leq 1$.

In Section 8 we consider self-similar operators which allow us to write the $\|\cdot\|_{\mathrm{BMO}_\psi}$-$\|\cdot\|_\infty$-estimate of Theorem 1.7 in a much more applicable form. As a consequence, we sharpen in Corollary 8.6 an inequality of G. Pisier concerning "martingale-type constants" of Banach spaces.

2. The Lorentz spaces M_φ and M_φ^0.

The spaces M_φ and M_φ^0 we define below are the "basic modules" of BMO_ψ-spaces.

DEFINITION 2.1. For $f \in L_p(\Omega, \mathcal{F}, \mathbb{P})$, $1 \leq p < \infty$, and $0 < t \leq 1$ let

$$f_p^{**}(t) := \left(\frac{1}{t}\int_0^t \widetilde{f}(s)^p\,ds\right)^{1/p} \quad \text{and} \quad f^{**}(t) := f_1^{**}(t),$$

where $\widetilde{f}(s) := \inf\{c > 0 \mid \mathbb{P}(|f| > c) \leq s\}$ is the non-increasing rearrangement of f. The set of all increasing bijections $\varphi : [0,1] \to [0,1]$ is denoted by $\mathcal{C}$, and $\mathcal{C}_\Delta$ is the subset of all $\varphi \in \mathcal{C}$ such that $\Delta(\varphi) := \sup_{0<t\leq 1/2} \varphi(2t)/\varphi(t) < \infty$. Finally, for $\varphi \in \mathcal{C}$ let

$$M_\varphi(\Omega, \mathcal{F}, \mathbb{P}) := \{f \in L_1(\Omega, \mathcal{F}, \mathbb{P}) \mid \|f\|_{M_\varphi} := \sup_{0<t\leq 1} f^{**}(t)\varphi(t) < \infty\},$$

$$M_\varphi^0(\Omega, \mathcal{F}, \mathbb{P}) := \{f \in L_0(\Omega, \mathcal{F}, \mathbb{P}) \mid |f|_{M_\varphi^0} := \sup_{0<t\leq 1} \widetilde{f}(t)\varphi(t) < \infty\}.$$

In Section 1 we have already introduced the class $\mathcal{D}$ of increasing bijections $\psi : [1,\infty) \to [1,\infty)$. If there is no risk of confusion we will freely switch between $\mathcal{C}$ and $\mathcal{D}$ (φ and ψ) with

$$\varphi(t) = \frac{1}{\psi^{-1}(1 + \log(1/t))} \quad \text{and} \quad \varphi^{-1}\left(\frac{1}{\lambda}\right) = e^{1-\psi(\lambda)}. \tag{15}$$

Basic information about the spaces M_φ can be found in [4] (p. 69ff). Moreover, it can be easily seen that $M_\varphi^0(\Omega, \mathcal{F}, \mathbb{P})$ is a linear space for all $[\Omega, \mathcal{F}, \mathbb{P}]$ if and only if $\Delta(\varphi) < \infty$. In particular,

$$|f + g|_{M_\varphi^0} \leq \Delta(\varphi)[|f|_{M_\varphi^0} + |g|_{M_\varphi^0}] \quad \text{for all } f, g \in M_\varphi^0(\Omega, \mathcal{F}, \mathbb{P}). \tag{16}$$

Furthermore, if one uses the relation (15) then one can write

$$|f|_{M_\varphi^0} = \inf\{c > 0 \mid \mathbb{P}(|f| > \lambda) \leq e^{1-\psi(\lambda/c)} \text{ for all } \lambda > c\}. \tag{17}$$

In order to formulate some basic properties of the M_φ-spaces we recall some further standard notation.

DEFINITION 2.2. A Banach space $[X, \|\cdot\|] \subseteq L_1(\Omega, \mathcal{F}, \mathbb{P})$ is called a *rearrangement invariant* (*r.i.*) *Banach function space* if

(1) $\chi_A \in X$ for all $A \in \mathcal{F}$,

(2) $f \in X$, $g \in L_0(\Omega, \mathcal{F}, \mathbb{P})$, $\widetilde{g} \leq \widetilde{f}$ imply $g \in X$ and $\|g\| \leq \|f\|$,

(3) (Fatou's property) for $(f_n)_{n=1}^\infty \subseteq X$ with $f_n \geq 0$ a.s., $\sup_n \|f_n\| < \infty$, and $f_n \uparrow f \in L_0(\Omega, \mathcal{F}, \mathbb{P})$ a.s. one has $f \in X$ and $\|f\| = \lim_n \|f_n\|$.

Identifying f and g if $f = g$ a.s. the following theorem is well known (cf. [4] (Proposition 2.5.8)).

THEOREM 2.3. *If* $\varphi \in \mathcal{C}$, *then* $[M_\varphi(\Omega, \mathcal{F}, \mathbb{P}), \|\cdot\|_{M_\varphi}]$ *is a r.i. Banach function space.*

From the viewpoint of Banach space theory the spaces M_φ are more convenient than the spaces M_φ^0. On the other hand, in view of (17) we use the spaces M_φ^0 to construct our BMO_ψ-spaces. Hence we recall a characterization, which is folklore (see for example [3] (Lemma 2)), of those $\varphi \in \mathcal{C}$

such that $M_\varphi(\Omega,\mathcal{F},\mathbb{P}) = M_\varphi^0(\Omega,\mathcal{F},\mathbb{P})$ for all $[\Omega,\mathcal{F},\mathbb{P}]$. Using this characterization we deduce in particular that the BMO_ψ-spaces always have an equivalent norm.

THEOREM 2.4. *For $\varphi \in \mathcal{C}$ and $c > 0$ the following are equivalent.*

(1) $\|f\|_{M_\varphi} \le c|f|_{M_\varphi^0}$ *for all* $[\Omega,\mathcal{F},\mathbb{P}]$ *and* $f \in M_\varphi^0(\Omega,\mathcal{F},\mathbb{P})$.

(2) $\displaystyle \frac{1}{t}\int_0^t \frac{ds}{\varphi(s)} \le \frac{c}{\varphi(t)}$ *for all* $0 < t \le 1$.

3. The K-functional and the supremum of independent random variables. The following lemma can be found in [19] (Lemma 3) and goes back to [15] (proof of Lemma 3.2).

LEMMA 3.1. *For $f_1,\ldots,f_N \in L_0(\Omega,\mathcal{F},\mathbb{P})$ and $\lambda > 0$ one has*

$$\mathbb{P}^N(\{(\omega_i)_{i=1}^N \mid \sup_{1\le i\le N} |f_i(\omega_i)| > \lambda\}) \ge \frac{\sum_{i=1}^N \mathbb{P}(|f_i| > \lambda)}{1+\sum_{i=1}^N \mathbb{P}(|f_i| > \lambda)}.$$

If $f_1,\ldots,f_N$ have the same distribution, then the above lemma together with a converse inequality is also contained in [1] (Lemmas 2.1, 2.2) (see also [13] (Lemma 3.1)). An immediate consequence is

LEMMA 3.2. *For $f_1,\ldots,f_N \in L_0(\Omega,\mathcal{F},\mathbb{P})$ let $h \in L_0([0,1]\times\Omega)$ and $g \in L_0(\Omega^N)$ be given by*

$$h(s,\omega) := \sum_{i=1}^N \chi_{[(i-1)/N,i/N)}(s) f_i(\omega) \quad \text{and} \quad g(\omega_1,\ldots,\omega_N) := \sup_{1\le i\le N} |f_i(\omega_i)|,$$

where $[0,1]$ is equipped with the Lebesgue measure $|\cdot|$. Then for $0 < t \le 1$ one has

$$\widetilde{h}\Big(\frac{t}{N}\Big) \le \widetilde{g}\Big(\frac{t}{t+1}\Big) \le \widetilde{g}\Big(\frac{t}{2}\Big).$$

Proof. For $0 < s < 1$ it follows from Lemma 3.1 that

$$\widetilde{g}(s) \ge \inf\Big\{c > 0 \,\Big|\, \frac{\sum_{i=1}^N \mathbb{P}(|f_i| > c)}{1+\sum_{i=1}^N \mathbb{P}(|f_i| > c)} \le s\Big\}$$
$$= \inf\Big\{c > 0 \,\Big|\, \frac{N(|\cdot|\times\mathbb{P})(|h| > c)}{1+N(|\cdot|\times\mathbb{P})(|h| > c)} \le s\Big\} = \widetilde{h}\Big(\frac{s}{N(1-s)}\Big).$$

Setting $s = t/(1+t)$ yields the desired result. ∎

Remark 3.3. In Lemma 3.2 one can improve $\widetilde{h}(t/N) \le \widetilde{g}(t/2)$ for some values of t and one gets also a converse inequality. More precisely, it can be shown for $0 < t \le 1 - 1/e =: 1/\gamma$ that

$$\widetilde{h}\Big(\frac{1}{N}\gamma t\Big) \le \widetilde{g}(t) \le \widetilde{h}\Big(\frac{1}{N\gamma}t\Big).$$

For later use let us recall (see [5] (Theorem 5.2.1)) that for $0 < t \le 1$ and $1 \le p < \infty$,

$$f_p^{**}(t) \le K_{\infty,p}[f, t^{-1/p}] \le 2^{1-1/p} f_p^{**}(t). \tag{18}$$

THEOREM 3.4. *Let $1 \le p < \infty$, $f_1,\ldots,f_N \in L_p(\Omega,\mathcal{F},\mathbb{P})$, and let $h \in L_p([0,1]\times\Omega)$ be defined by $h(s,\omega) := \sum_{i=1}^N \chi_{[(i-1)/N,i/N)}(s) f_i(\omega)$. Then for $M \ge 1$ one has*

$$\tfrac{1}{2}K_{\infty,p}[h,(MN)^{1/p}] \le \|\sup_{\substack{1\le i\le N\\ 1\le j\le M}} |f_i(\omega_{ij})|\|_{L_p(\Omega^{MN})} \le K_{\infty,p}[h,(MN)^{1/p}].$$

Proof. Let $\{v_1,\ldots,v_{MN}\} = \{f_1,\ldots,f_1,f_2,\ldots,f_2,\ldots,f_N,\ldots,f_N\}$, where each f_i is repeated M times, and define $H \in L_p([0,1]\times\Omega)$ and $G \in L_0(\Omega^{MN})$ by

$$H(s,\omega) := \sum_{i=1}^{MN} \chi_{[(i-1)/(MN),i/(MN))}(s) v_i(\omega)$$

and

$$G(\omega_1,\ldots,\omega_{MN}) := \sup_{1\le i\le MN} |v_i(\omega_i)|.$$

From Lemma 3.2 we know that $\widetilde{H}(t/(MN)) \le \widetilde{G}(t/2)$ for $0 < t \le 1$ so that $H_p^{**}(1/(MN)) \le \sqrt[p]{2}\|G\|_p$. Since $\widetilde{H} = \widetilde{h}$ we arrive at the left-hand side of our assertion with the help of (18). To show the right-hand side let $h = a+b$ where $a \in L_\infty([0,1]\times\Omega)$ and $b \in L_p([0,1]\times\Omega)$. Then

$$\begin{aligned}
&\|\sup_{\substack{1\le i\le N\\ 1\le j\le M}} |f_i(\omega_{ij})|\|_{L_p(\Omega^{MN})}\\
&\quad = \Big\|\sup_{\substack{1\le i\le N\\ 1\le j\le M}} \Big(N\int_{(i-1)/N}^{i/N} |h(t,\omega_{ij})|^p\,dt\Big)^{1/p}\Big\|_{L_p(\Omega^{MN})}\\
&\quad \le \|a\|_\infty + \Big\|\sup_{\substack{1\le i\le N\\ 1\le j\le M}} \Big(N\int_{(i-1)/N}^{i/N} |b(t,\omega_{ij})|^p\,dt\Big)^{1/p}\Big\|_{L_p(\Omega^{MN})}\\
&\quad \le \|a\|_\infty + M^{1/p}\Big\|\sup_{1\le i\le N} \Big(N\int_{(i-1)/N}^{i/N} |b(t,\omega_i)|^p\,dt\Big)^{1/p}\Big\|_{L_p(\Omega^N)}\\
&\quad \le \|a\|_\infty + M^{1/p}N^{1/p}\Big\|\Big(\sum_{i=1}^N\int_{(i-1)/N}^{i/N} |b(t,\omega_i)|^p\,dt\Big)^{1/p}\Big\|_{L_p(\Omega^N)}
\end{aligned}$$

$$= \|a\|_\infty + (MN)^{1/p}\Big(\sum_{i=1}^{N}\int_{(i-1)/N}^{i/N}\Big[\int_\Omega |b(t,\omega_i)|^p\,d\mathbb{P}(\omega_i)\Big]dt\Big)^{1/p}$$

$$= \|a\|_\infty + (MN)^{1/p}\|b\|_p.$$

Since this holds for all decompositions $h = a + b$ as above we are done. ■

Proof of Theorem 1.8. We have to use the above theorem with $M = 1$ and $f_i := \alpha_i f$. ■

4. BMO_ψ-spaces. We investigate a scale of BMO-norms for adapted sequences of random variables. To define these BMO-norms also for random variables $f \in L_1(\Omega,\mathcal{F},\mathbb{P})$ we denote the conditional expectation of f with respect to the sub-σ-algebra $\mathcal{F}_k \subseteq \mathcal{F}$ by $\mathbb{E}(f \mid \mathcal{F}_k) = \mathbb{E}_k f$. Moreover, $\Omega := [0,1]$ is always equipped with the Lebesgue measure $|\cdot|$ defined on the Borel σ-algebra and is equipped with the filtration $(\mathcal{F}_k)_{k=0}^\infty$ generated by $\mathcal{F}_0 := \{[0,1],\emptyset\}$ and, for $k \ge 1$, by

$$\mathcal{F}_k := \sigma\Big\{\Big[0,\frac{1}{2^k}\Big),\ldots,\Big[\frac{2^k-2}{2^k},\frac{2^k-1}{2^k}\Big),\Big[\frac{2^k-1}{2^k},1\Big]\Big\}.$$

In that situation we simply write $\mathrm{BMO}_\psi[0,1]$ instead of $\mathrm{BMO}_\psi((\mathcal{F}_k)_{k=0}^\infty)$. The intervals $[(i-1)/2^k, i/2^k)$ for $i = 1,\ldots,2^k-1$ and $[(2^k-1)/2^k,1]$ are called the *atoms* of $\mathcal{F}_k$. Moreover, if $B \in \mathcal{F}_k$ is an atom, where $k \ge 1$, then the atom $\widetilde{B} \in \mathcal{F}_{k-1}$ containing B is called the *dyadic predecessor* of B.

The BMO-norm for an adapted sequence $(f_k)_{k\in I} \in \mathcal{A}((\mathcal{F}_k)_{k\in I})$ was defined in [14] (p. 66) by

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}} := \sup_{\substack{0\le k\le l\\ k,l\in I}} \|\mathbb{E}(|f_l - f_{k-1}| \mid \mathcal{F}_k)\|_\infty$$
$$= \sup_{\substack{0\le k\le l\\ k,l\in I}} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \|f_l - f_{k-1}\|_{L_1(B,\mathbb{P}_B)},$$

where $f_{-1} = 0$ and $\mathbb{P}_B = \mathbb{P}/\mathbb{P}(B)$ is the normalized restriction of $\mathbb{P}$ to B. For $f \in L_1(\Omega,\mathcal{F},\mathbb{P})$ we obtain the classical BMO-norm by setting

$$\|f\|_{\mathrm{BMO}} := \|(\mathbb{E}(f \mid \mathcal{F}_k))_{k\in I}\|_{\mathrm{BMO}} = \sup_{k\in I} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \|f - f_{k-1}\|_{L_1(B,\mathbb{P}_B)}. \tag{19}$$

Let us recall a well-known BMO-function defined on $[0,1]$ used later.

EXAMPLE 4.1. For $f(t) := \log(1/t) \in L_1[0,1]$ one has $\|f\|_{\mathrm{BMO}} < \infty$.

In Section 1 we have already defined the BMO_ψ-norm which can be written for an adapted sequence $(f_k)_{k\in I} \in \mathcal{A}((\mathcal{F}_k)_{k\in I})$ with the help of (17) as

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} = \sup_{\substack{0\le k\le l\\ k,l\in I}} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} |f_l - f_{k-1}|_{M^0_\varphi(B,\mathbb{P}_B)}$$

where

$$\varphi(t) = \frac{1}{\psi^{-1}(1+\log(1/t))}.$$

It is useful to define some similar variants of the above quantity.

DEFINITION 4.2. Let $\psi \in \mathcal{D}$, $0 < s < 1$, and $(f_k)_{k\in I} \in \mathcal{A}((\mathcal{F}_k)_{k\in I})$. Then

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_\psi} := \sup_{k\in I} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \Big|\sup_{k\le l\in I} |f_l - f_{k-1}|\Big|_{M^0_\varphi(B,\mathbb{P}_B)},$$

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}_{0,s}} := \sup_{\substack{0\le k\le l\\ k,l\in I}} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \inf\{c > 0 \mid \mathbb{P}_B(|f_l - f_{k-1}| > c) \le s\},$$

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{0,s}} := \sup_{k\in I} \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \inf\{c > 0 \mid \mathbb{P}_B(\sup_{k\le l\in I} |f_l - f_{k-1}| > c) \le s\},$$

where $\mathbb{P}_B := \mathbb{P}/\mathbb{P}(B)$ is the normalized restriction of $\mathbb{P}$ to B and $f_{-1} = 0$.

In the literature several BMO-norms have been considered. The approach which is relevant here, although it looks quite different at first glance, can be found in [2]. In Remark 4.14(1) we will outline the relations between the approach of [2] and ours. Recalling that for $\psi \in \mathcal{D}$ the function $\overline{\psi} : [1,\infty) \to [1,\infty)$ was introduced as

$$\overline{\psi}(\mu) = \sup\Big\{\sum_{i=1}^N (\psi(\mu_i) - 1) + 1 \;\Big|\; \mu = \sum_{i=1}^N \mu_i,\ \mu_i \ge 1,\ N = 1,2,\ldots\Big\}$$

we set, by analogy with (15),

$$\overline{\varphi}(t) := \frac{1}{\overline{\psi}^{-1}(1+\log(1/t))}. \tag{20}$$

Let us summarize some properties of this construction.

EXAMPLE 4.3. (1) For $\psi(\mu) := 1 + \log\mu$ one has $\overline{\psi}(\mu) \ge \mu/e$.
(2) For $1 \le p < \infty$ and $\psi_p(\mu) := \mu^p$ one has $\overline{\psi}_p(\mu) = \mu^p$.

Proof. (1) For some "general" $\psi \in \mathcal{D}$, $k = 1,2,\ldots$ and $k\psi^{-1}(2) \le \mu < (k+1)\psi^{-1}(2)$ we get

$$\overline{\psi}(\mu) - 1 \ge \overline{\psi}(k\psi^{-1}(2)) - 1 \ge k(\psi(\psi^{-1}(2)) - 1) \ge \frac{\mu}{\psi^{-1}(2)} - 1$$

so that

$$\inf_{\mu\ge 1} \overline{\psi}(\mu)/\mu \ge 1/\psi^{-1}(2).$$

(2) Here we exploit $\sum_{i=1}^N (\mu_i^p - 1) + 1 \le (\sum_{i=1}^N \mu_i)^p$ for $\mu_i \ge 0$. ■

LEMMA 4.4. *For $\psi \in \mathcal{D}$ the following holds.*

(1) *If $\alpha, \overline{\alpha} : [1,\infty) \to [0,\infty)$ are given by $\alpha(\mu) := \psi(\mu) - 1$ and $\overline{\alpha}(\mu) := \overline{\psi}(\mu) - 1$, then $\overline{\alpha}$ is the least majorant β of α satisfying for all $\mu, \nu \geq 1$,*

$$\beta(\mu) + \beta(\nu) \leq \beta(\mu + \nu).$$

(2) *$\overline{\varphi} \in \mathcal{C}_\Delta$ and there is some $c_\psi \geq 1$, depending on ψ only, such that for all $\mu, \lambda \geq 1$,*

$$\mu\overline{\psi}(\lambda) \leq c_\psi \overline{\psi}(\mu\lambda).$$

(3) $\|\cdot\|_{M_{\overline{\varphi}}} \leq 4\overline{\psi}^{-1}(1+\log 4)|\cdot|_{M^0_{\overline{\varphi}}}$.

For the proof and for later use we need the following lemma due to D. L. Burkholder.

LEMMA 4.5 [8] (Lemma 7.1). *Let $f, g \in L_0^+(\Omega, \mathcal{F}, \mathbb{P})$ and $0 < p < \infty$ be such that for some $\beta > 1$ and $\delta, \varepsilon > 0$ with $\beta^p\varepsilon < 1$ one has*

$$\mathbb{P}(f > \beta\lambda,\ g \leq \delta\lambda) \leq \varepsilon\mathbb{P}(f > \lambda)$$

for all $\lambda > 0$. Then

$$\|f\|_p \leq \frac{\beta}{\delta} \cdot \frac{1}{\sqrt[p]{1-\beta^p\varepsilon}}\|g\|_p.$$

Proof of Lemma 4.4. (1) From the definition it follows that

$$\overline{\alpha}(\mu) := \sup\Big\{\sum_{i=1}^N \alpha(\mu_i) \Bigm| \mu = \sum_{i=1}^N \mu_i,\ \mu_i \geq 1,\ N = 1,2,\ldots\Big\}.$$

Now $\overline{\alpha}(\mu) \geq \alpha(\mu)$ and $\overline{\alpha}(\mu+\nu) \geq \overline{\alpha}(\mu) + \overline{\alpha}(\nu)$ are clear from the definition. For a majorant β satisfying $\beta(\mu) + \beta(\nu) \leq \beta(\mu+\nu)$ we get

$$\begin{aligned}\beta(\mu) &\geq \sup\Big\{\sum_{i=1}^N \beta(\mu_i) \Bigm| \mu = \sum_{i=1}^N \mu_i,\ \mu_i \geq 1,\ N = 1,2,\ldots\Big\}\\ &\geq \sup\Big\{\sum_{i=1}^N \alpha(\mu_i) \Bigm| \mu = \sum_{i=1}^N \mu_i,\ \mu_i \geq 1,\ N = 1,2,\ldots\Big\} = \overline{\alpha}(\mu).\end{aligned}$$

(2) Clearly $\overline{\alpha}(1) = 0$ and $\lim_{\mu\to\infty} \overline{\alpha}(\mu) = \infty$. For $N \geq 1$ we define $\alpha_N : [1,\infty) \to [0,\infty)$ by

$$\alpha_N(\mu) := \begin{cases} 0 & \text{if } 1 \leq \mu \leq N,\\ \sup\Big\{\sum_{i=1}^N \alpha(\mu_i) \Bigm| \mu = \sum_{i=1}^N \mu_i,\ \mu_i \geq 1\Big\} & \text{if } \mu > N,\end{cases}$$

so that $\overline{\alpha}(\mu) = \sup_{N\leq\mu} \alpha_N(\mu)$. If $\mu \geq N$, then the continuity of α and the compactness of the set $\{(\mu_1,\ldots,\mu_N) \mid \sum_{i=1}^N \mu_i = \mu,\ \mu_i \geq 1\} \subseteq [1,\infty)^N$

imply that there are some $\mu_i \geq 1$ such that $\alpha_N(\mu) = \sum_{i=1}^N \alpha(\mu_i)$. Consequently, for all $\mu \geq 1$ there is an $N \geq 1$ and there are $\mu_1,\ldots,\mu_N \geq 1$ with $\overline{\alpha}(\mu) = \sum_{i=1}^N \alpha(\mu_i)$.

The function $\overline{\alpha}$ is strictly increasing. Indeed, assuming $1 \leq \mu < \nu < \infty$ we write $\nu = \sum_{i=1}^N \nu_i$ with $\nu_1 > \mu_1$, $\nu_2 = \mu_2, \ldots, \nu_N = \mu_N$ such that

$$\overline{\alpha}(\mu) = \sum_{i=1}^N \alpha(\mu_i) < \sum_{i=1}^N \alpha(\nu_i) \leq \overline{\alpha}(\nu).$$

The function $\overline{\alpha}$ is continuous. Since $\overline{\alpha}(\mu) = \sup_{N\leq\tau} \alpha_N(\mu)$ for $\tau \geq \mu$ it suffices to show that α_N is continuous, and this follows from

$$\sup_{\substack{\nu-\mu\leq\delta\\ N\leq\mu<\nu\leq C}} |\alpha_N(\nu) - \alpha_N(\mu)| \leq N \sup_{\substack{\nu-\mu\leq\delta\\ 1\leq\mu<\nu\leq C}} |\alpha(\nu) - \alpha(\mu)|$$

for all $C > N$ and $\delta > 0$ (for $N \leq \mu < \nu = \sum_{i=1}^N \nu_i \leq C$ with $\alpha_N(\nu) = \sum_{i=1}^N \alpha(\nu_i)$ choose $1 \leq \mu_i \leq \nu_i \leq \mu_i + \delta$ with $\mu = \sum_{i=1}^N \mu_i$).

To show $\mu\overline{\psi}(\lambda) \leq c_\psi\overline{\psi}(\mu\lambda)$ we use $\overline{\psi}(k) \geq k/\psi^{-1}(2)$ from the proof of Example 4.3 and obtain for $1 \leq k \leq \mu < k+1$,

$$\begin{aligned}(\psi^{-1}(2)+1)\overline{\psi}(\mu\lambda) &\geq \psi^{-1}(2)\overline{\psi}(k) + \overline{\psi}(k\lambda) \geq k + \overline{\psi}(k\lambda)\\ &\geq k + k(\overline{\psi}(\lambda) - 1) = k\overline{\psi}(\lambda) \geq \frac{\mu}{2}\overline{\psi}(\lambda).\end{aligned}$$

Now, let $\mu_0 := 2c_\psi$ be such that $2c_\psi\overline{\psi}(\lambda) = \mu_0\overline{\psi}(\lambda) \leq c_\psi\overline{\psi}(\mu_0\lambda)$ and $\overline{\psi}^{-1}(2t) \leq \mu_0\overline{\psi}^{-1}(t)$ for all $t \geq 1$. Then

$$\begin{aligned}\sup_{0<t\leq 1}\frac{\overline{\varphi}(t)}{\overline{\varphi}(t/2)} &= \sup_{0<t\leq 1}\frac{\overline{\psi}^{-1}(1+\log 2+\log(1/t))}{\overline{\psi}^{-1}(1+\log(1/t))}\\ &\leq \sup_{0<t\leq 1}\frac{\overline{\psi}^{-1}(2[1+\log(1/t)])}{\overline{\psi}^{-1}(1+\log(1/t))} \leq \mu_0.\end{aligned}$$

(3) For $\lambda_0 \geq 1$ and $f_t(s) := \overline{\varphi}(t)/\overline{\varphi}(st)$ $(0 < s,t \leq 1)$ we have

$$\begin{aligned}\sup_{\lambda\geq\lambda_0}\frac{|\{f_t > 2\lambda\}|}{|\{f_t > \lambda\}|} &= \sup_{\lambda\geq\lambda_0/\overline{\varphi}(t)}\frac{|\{s : 1/\overline{\varphi}(st) > 2\lambda\}|}{|\{s : 1/\overline{\varphi}(st) > \lambda\}|} \leq \sup_{\lambda\geq\lambda_0/\overline{\varphi}(t)}\frac{|\{1/\overline{\varphi} > 2\lambda\}|}{|\{1/\overline{\varphi} > \lambda\}|}\\ &\leq \sup_{\lambda\geq\lambda_0}\frac{\overline{\varphi}^{-1}(1/(2\lambda))}{\overline{\varphi}^{-1}(1/\lambda)}.\end{aligned}$$

Setting $\lambda_0 := \overline{\psi}^{-1}(1+\log 4)$ we continue with the help of (20) and $\overline{\alpha}(\mu+\nu) \geq \overline{\alpha}(\mu) + \overline{\alpha}(\nu)$ to get

$$\sup_{\lambda\geq\lambda_0}\frac{|\{f_t > 2\lambda\}|}{|\{f_t > \lambda\}|} \leq \sup_{\lambda\geq\lambda_0} e^{(\overline{\psi}(\lambda)-1)-(\overline{\psi}(2\lambda)-1)} \leq \sup_{\lambda\geq\lambda_0} e^{1-\overline{\psi}(\lambda)} = \frac{1}{4}.$$

Hence we have proved for $g = 1$ and $\lambda > 0$ that

$$|\{f_t > 2\lambda,\ g \leq \lambda/\lambda_0\}| \leq \tfrac{1}{4}|\{f_t > \lambda\}|.$$

Applying Lemma 4.5 gives $\|f_t\|_1 \leq 4\overline{\psi}^{-1}(1+\log 4)$, which is exactly the assertion of Theorem 2.4(2) with $c = 4\overline{\psi}^{-1}(1+\log 4)$. ∎

The next theorem relates the different BMO-norms introduced above to each other. The point of this theorem is that we start with some weight function $\psi \in \mathcal{D}$ and arrive at $\overline{\psi} \in \mathcal{D}$ which describes the same BMO-space as ψ does. The weight function $\overline{\psi}$ has certain regularity properties as shown in Lemma 4.4, which will be used several times. For example, these regularity properties allow us to classify the BMO_ψ-spaces (see Theorem 1.3). Moreover, the passage $\psi \to \overline{\psi}$ is a natural generalization of an iteration procedure used sometimes to prove the John–Nirenberg Theorem (see for example [23] (p. 154)). In fact, $\psi(\mu) = 1 + \log\mu$ implies $\overline{\psi}(\mu) \geq \mu/e$ (see Example 4.3), which is the John–Nirenberg Theorem as demonstrated in Corollary 4.8.

THEOREM 4.6. (a) *For* $\psi \in \mathcal{D}$, $0 < s < 1$, *and* $\psi_1(t) = t$ *one has*

$$\|\cdot\|_{\mathrm{BMO}_\psi} \leq \|\cdot\|_{\mathrm{BMO}^*_{\overline{\psi}}} = \|\cdot\|_{\mathrm{BMO}^{\pm}_{\overline{\psi}}} \leq 6\psi^{-1}(3)\|\cdot\|_{\mathrm{BMO}_\psi}, \tag{21}$$

$$\frac{1}{1+\log(1/s)}\|\cdot\|_{\mathrm{BMO}^*_{0,s}} \leq \|\cdot\|_{\mathrm{BMO}^*_{\psi_1}} \leq \max\left(1, \frac{1}{\log(1/s)}\right)\|\cdot\|_{\mathrm{BMO}^*_{0,s}}, \tag{22}$$

$$\|\cdot\|_{\mathrm{BMO}^*_{\psi_1}} \leq 3\max\left(1, \frac{1}{\log(1/s)}\right)\|\cdot\|_{\mathrm{BMO}_{0,s/2}}. \tag{23}$$

(b) *In particular*, $\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\overline{\psi}}} \leq 1$ *implies for* $\lambda > 0$ *and* $\mu \geq 1$,

$$\mathbb{P}(\sup_{l\in I}|f_l| > \lambda + \mu) \leq e^{1-\overline{\psi}(\mu)}\mathbb{P}(\sup_{l\in I}|f_l| > \lambda). \tag{24}$$

Proof. First fix $0 \leq k \leq n \in I$ and $B \in \mathcal{F}_k$ with $\mathbb{P}(B) > 0$. For $\varrho > 0$ we define

$$\tau_\varrho := \inf\{i \geq k \mid \chi_B|f_i - f_{k-1}| > \varrho\} \wedge (n+1).$$

The following steps (i) and (ii) are standard and are obtained in several papers.

(i) For $\lambda, \mu > 0$ we get

$$\begin{aligned}
&\mathbb{P}_B(\sup_{k\leq l\leq n}|f_l - f_{k-1}| > \lambda + \mu)\\
&\quad = \sum_{i=k}^n \mathbb{P}_B(\sup_{i\leq l\leq n}|f_l - f_{k-1}| > \lambda + \mu,\ \tau_\lambda = i)\\
&\quad \leq \sum_{i=k}^n \mathbb{P}_B(\sup_{i\leq l\leq n}|f_l - f_{k-1}| > |f_{i-1} - f_{k-1}| + \mu,\ \tau_\lambda = i)\\
&\quad \leq \sum_{i=k}^n \mathbb{P}_B(\sup_{i\leq l\leq n}|f_l - f_{i-1}| > \mu,\ \tau_\lambda = i).
\end{aligned}$$

(ii) Following an idea contained in [14] (p. 75ff) we get for $\lambda > 0$, $\mu \geq 1$, and $\sup_{l\in I}\|f_l - f_{l-1}\|_\infty \leq 1$,

$$\begin{aligned}
&\mathbb{P}_B(\sup_{k\leq l\leq n}|f_l - f_{k-1}| > \lambda + \mu)\\
&= \sum_{i=k}^n \mathbb{P}_B(\sup_{i\leq l\leq n}|f_l - f_{k-1}| > \lambda + \mu,\ \tau_\lambda = i)\\
&= \sum_{k\leq i\leq l\leq n} \mathbb{P}_B(|f_l - f_{k-1}| > \lambda + \mu,\ \tau_\lambda = i,\ \tau_{\lambda+\mu} = l)\\
&\leq \sum_{k\leq i\leq l\leq n} \mathbb{P}_B(|f_l - f_{k-1}| > |f_{i-1} - f_{k-1}| + \mu,\ \tau_\lambda = i,\ \tau_{\lambda+\mu} = l)\\
&\leq \sum_{k\leq i\leq l\leq n} \mathbb{P}_B(|f_n - f_{l-1}| + |f_n - f_{i-1}| > \mu - |f_l - f_{l-1}|,\ \tau_\lambda = i,\ \tau_{\lambda+\mu} = l)\\
&\leq \sum_{k\leq i\leq l\leq n} [\mathbb{P}_B(|f_n - f_{l-1}| > (\mu-1)/2,\ \tau_\lambda = i,\ \tau_{\lambda+\mu} = l)\\
&\qquad + \mathbb{P}_B(|f_n - f_{i-1}| > (\mu-1)/2,\ \tau_\lambda = i,\ \tau_{\lambda+\mu} = l)].
\end{aligned}$$

(iii) We show $\|\cdot\|_{\mathrm{BMO}^{\pm}_{\overline{\psi}}} \leq \|\cdot\|_{\mathrm{BMO}^*_{\overline{\psi}}}$. If $\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\overline{\psi}}} \leq 1$, then step (i) gives

$$\begin{aligned}
&\mathbb{P}_B(\sup_{k\leq l\leq n}|f_l - f_{k-1}| > \lambda + \mu)\\
&\qquad \leq \sum_{i=k}^n \mathbb{P}(\{\tau_\lambda = i\} \cap B \cap \{\sup_{i\leq l\leq n}|f_l - f_{i-1}| > \mu\})/\mathbb{P}(B)\\
&\qquad \leq \sum_{i=k}^n e^{1-\psi(\mu)}\mathbb{P}(\{\tau_\lambda = i\} \cap B)/\mathbb{P}(B)\\
&\qquad = e^{1-\psi(\mu)}\mathbb{P}_B(\sup_{k\leq l\leq n}|f_l - f_{k-1}| > \lambda).
\end{aligned}$$

Replacing ψ by $\overline{\psi}$ by induction over N from Definition 1.2 (and letting $n \to \infty$ if $I = \mathbb{N}$) implies for $\mu \geq 1$,

$$\mathbb{P}_B(\sup_{k\leq l}|f_l - f_{k-1}| > \lambda + \mu) \leq e^{1-\overline{\psi}(\mu)}\mathbb{P}_B(\sup_{k\leq l}|f_l - f_{k-1}| > \lambda) \tag{25}$$

and (24). Now $\lambda \downarrow 0$ gives $\mathbb{P}_B(\sup_{k\leq l}|f_l - f_{k-1}| > \mu) \leq e^{1-\overline{\psi}(\mu)}$ and $\|(f_k)_{k\in I}\|_{\mathrm{BMO}^{\pm}_{\overline{\psi}}} \leq 1$.

(iv) To show $\|\cdot\|_{\mathrm{BMO}^*_{\overline{\psi}}} \leq 6\psi^{-1}(3)\|\cdot\|_{\mathrm{BMO}_\psi}$ let $\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} \leq 1$, which implies

$$\|f_l - f_{l-1}\|_\infty \le \sup_{\substack{B\in\mathcal{F}_l\\ \mathbb{P}(B)>0}} |f_l - f_{l-1}|_{M^0_\varphi(B,\mathbb{P}_B)} \le 1$$

for $l \in I$. Since $B\cap\bigcup_{l=k}^n\{\tau_\lambda = l \vee \tau_{\lambda+\mu} = l\} = B\cap\{\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda\}$ step (ii) gives for $\mu \ge 3$, as in step (iii),

$$(26)\quad \mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+\mu) \le 2e^{1-\psi((\mu-1)/2)}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda).$$

For $\mu_0 := 3\psi^{-1}(1+\log 4) \le 3\psi^{-1}(3)$ and $\mu/2 \ge \mu_0$, using $(\mu/2-1)/2 \ge \mu/6$ we conclude

$$\begin{aligned}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+\mu) &= \mathbb{P}_B\Big(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+\frac{\mu}{2}+\frac{\mu}{2}\Big)\\ &\le (e^{\log 2+1-\psi(\mu/6)})^2\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)\\ &\le e^{1-\psi(\mu/6)}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)\\ &\le e^{1-\psi(\mu/(2\mu_0))}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda),\end{aligned}$$

which as in (iii) implies the inequality $\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_\psi} \le 2\mu_0$.

(v) The relations $\|\cdot\|_{\mathrm{BMO}_\psi} \le \|\cdot\|_{\mathrm{BMO}^*_\psi} \le \|\cdot\|_{\mathrm{BMO}^*_{\overline{\psi}}}$ are obvious.

(vi) Assume that $\|(f_k)_{k\in I}\|_{\mathrm{BMO}_{0,s/2}} \le 1$. Since for all $l \in I$,

$$\|f_l - f_{l-1}\|_\infty \le \|(f_k)_{k=0}^n\|_{\mathrm{BMO}_{0,s/2}} \le 1,$$

from (ii) applied to $\mu = 3$ we get the inequality

$$\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+3) \le 2\frac{s}{2}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)$$

in the same way as (26) in (iv). For $p = 1,2,\ldots$ and $3p \le \mu < 3(p+1)$ we deduce

$$\begin{aligned}\mathbb{P}_B(\sup_{k\le l\le n}&|f_l - f_{k-1}| > \lambda+\mu)\\ &\le \mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+3p)\\ &\le s^p\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)\\ &\le s^{\mu/3-1}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)\\ &\le e^{1-(\mu/3)\min(1,\log(1/s))}\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda)\end{aligned}$$

and get (23) as $\lambda \downarrow 0$ (and $n \to \infty$ if $I = \mathbb{N}$). To verify the right-hand side of (22) assume $\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{0,s}} \le 1$. From (i) we get for $\mu = 1$, as in (iii),

$$\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda+1) \le s\mathbb{P}_B(\sup_{k\le l\le n}|f_l - f_{k-1}| > \lambda).$$

The rest is similar to the proof of (23). The left-hand side of (22) follows from $\widetilde{h}(s) \le (1+\log(1/s))|h|_{M^0_{\varphi_1}}$ if $\varphi_1(t) := 1/(1+\log(1/t))$ and $h \in M^0_{\varphi_1}$. ∎

Remark 4.7. (1) In [24] there is defined a quantity $\|f\|_{\mathrm{BMO}_{0,s}}$ for functions $f \in L_0(\mathbb{R}^n)$ in a slightly different form. Lemma 3.1 of that paper, referred to in [17], is a counterpart to (23).

(2) The proof of (23) gives, on the right-hand side, $\|\cdot\|_{\mathrm{BMO}_{0,s/2}}$ instead of the smaller quantity $\|\cdot\|_{\mathrm{BMO}_{0,s}}$ which is sufficient for the purpose of this paper. Nevertheless it could be of interest to investigate $\|\cdot\|_{\mathrm{BMO}_{0,s}}$ whenever $1/2 \le s < 1$, which is not done here.

Theorem 4.6 includes A. M. Garsia's [14] (Theorems III.1.2, III.2.1) version of the original John–Nirenberg Theorem [18].

COROLLARY 4.8. *$\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\psi_1}} \le 6e^3\|(f_k)_{k\in I}\|_{\mathrm{BMO}}$ for $(f_k)_{k\in I} \in \mathcal{A}((\mathcal{F}_k)_{k\in I})$.*

Proof. For $\psi(\mu) = 1+\log\mu$ and $\varphi(t) = t$, from Example 4.3 and Theorem 4.6 we get

$$\begin{aligned}\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\psi_1}} &\le e\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\overline{\psi}}} \le e6\psi^{-1}(3)\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi}\\ &\le e6\psi^{-1}(3)\|(f_k)_{k\in I}\|_{\mathrm{BMO}} = 6e^3\|(f_k)_{k\in I}\|_{\mathrm{BMO}}.\ ∎\end{aligned}$$

Note that (23) is stronger than the above corollary. Moreover, it can be used to get a better constant in this corollary. In fact, for $s = 1/e$ inequality (23) and the property $\widetilde{h}(s) \le (1/s)\|h\|_{L_1}$ yield

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}^*_{\psi_1}} \le 3\|(f_k)_{k\in I}\|_{\mathrm{BMO}_{0,s/2}} \le 6e\|(f_k)_{k\in I}\|_{\mathrm{BMO}}.$$

As expected, BMO$_\psi$-spaces are Banach spaces if we identify two adapted sequences $(f_k)_{k\in I}$ and $(g_k)_{k\in I}$ whenever $f_k = g_k$ a.s. for all $k \in I$.

THEOREM 4.9. *If $\psi \in \mathcal{D}$, then $[\mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I}), |||\cdot|||_{\overline{\psi}}]$ is a Banach space where*

$$|||(f_k)_{k\in I}|||_{\overline{\psi}} := \sup_{0\le k\le l\in I}\ \sup_{B\in\mathcal{F}_k,\mathbb{P}(B)>0}\|f_l - f_{k-1}\|_{M_{\overline{\psi}}(B,\mathbb{P}_B)}$$

and $f_{-1} = 0$. Moreover,

$$\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} \le |||(f_k)_{k\in I}|||_{\overline{\psi}} \le 24\overline{\psi}^{-1}(3)\psi^{-1}(3)\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi}.$$

Proof. The inequality $\|\cdot\|_{\mathrm{BMO}_\psi} \le |||\cdot|||_{\overline{\psi}}$ is trivial. The other one is a consequence of Theorem 4.6(21) and Lemma 4.4(3). To show the completeness let $(x_i)_{i=1}^\infty = ((f^i_k)_{k\in I})_{i=1}^\infty \subset \mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I})$ be a Cauchy sequence with respect to $|||\cdot|||_{\overline{\psi}}$ such that $f_k := \lim_i f_k^i \in M_{\overline{\varphi}}(\Omega,\mathcal{F}_k,\mathbb{P})$ exists for all $k \in I$. Assume $0 \le k \le l \in I$ and $B \in \mathcal{F}_k$ with $\mathbb{P}(B) > 0$. Taking into account

$$\|g\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)} \le \sup_{0<t\le 1} \frac{\overline{\varphi}(t)}{\overline{\varphi}(t\mathbb{P}(B))}\|g\|_{M_{\overline{\varphi}}(\Omega,\mathbb{P})} < \infty \quad \text{for } g \in M_{\overline{\varphi}}(\Omega,\mathcal{F},\mathbb{P})$$

(note that $\Delta(\overline{\varphi}) < \infty$), we get $|||(f_s)_{s\in I} - x_j|||_{\overline{\psi}} \le \limsup_i |||x_i - x_j|||_{\overline{\psi}}$ from

$$\|(f_l - f_l^j) - (f_{k-1} - f_{k-1}^j)\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)} = \lim_i \|(f_l^i - f_l^j) - (f_{k-1}^i - f_{k-1}^j)\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)}. \ \blacksquare$$

Now let us introduce the BMO_ψ-norm of $f \in L_1(\Omega,\mathcal{F},\mathbb{P})$.

Definition 4.10. For $f \in L_1(\Omega,\mathcal{F},\mathbb{P})$ let $f \in \mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I})$ provided that

$$\|f\|_{\mathrm{BMO}_\psi} := \|(\mathbb{E}_k f)_{k\in I}\|_{\mathrm{BMO}_\psi} < \infty.$$

Remark 4.11. If $f \in L_1[0,1]$ then

$$\|f\|_{\mathrm{BMO}_\psi} \sim \sup_{k\ge 0} \sup_{\substack{B\in\mathcal{F}_k\\ \text{atom}}} |f - \mathbb{E}_{k-1}f|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)} \sim \sup_{k\ge 0} \sup_{\substack{B\in\mathcal{F}_k\\ \text{atom}}} \|f - \mathbb{E}_{k-1}f\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)}$$

where the multiplicative constants involved in these inequalities depend on ψ only. In fact, let $B \in \mathcal{F}_k$ with $|B| > 0$ be the disjoint union of $\mathcal{F}_k$-atoms B^i. A standard computation shows the estimate $|h|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)} \le \sup_i |h|_{M^0_{\overline{\varphi}}(B^i,\mathbb{P}_{B^i})}$. Hence

$$\|f\|_{\mathrm{BMO}_{\overline{\psi}}} = \sup_{0\le k\le l} \sup_{\substack{B\in\mathcal{F}_k\\ \text{atom}}} |\mathbb{E}_l f - \mathbb{E}_{k-1}f|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)}.$$

Since $\sup_{l\ge k} \|\mathbb{E}_l f - \mathbb{E}_{k-1}f\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)} = \|f - \mathbb{E}_{k-1}f\|_{M_{\overline{\varphi}}(B,\mathbb{P}_B)}$ as a consequence of [4] (Theorems 1.1.7, 2.4.8) and $\mathbb{E}_k f \to f$ a.s. we are done by Lemma 4.4 and Theorem 4.6.

The basic example of a BMO_ψ-function defined on the unit interval $[0,1]$ is given by

Theorem 4.12. *For $\psi \in \mathcal{D}$ one has*

$$f(t) := \overline{\psi}^{-1}(1 + \log(1/t)) \in \mathrm{BMO}_\psi[0,1].$$

The theorem follows immediately from Example 4.1 and

Lemma 4.13. *Let $\psi \in \mathcal{D}$ and $f \in \mathrm{BMO}[0,1]$ be non-negative. Then*

$$\|\overline{\psi}^{-1}(1+f)\|_{\mathrm{BMO}_\psi[0,1]} \le c\|1+f\|_{\mathrm{BMO}[0,1]},$$

where $c > 0$ depends on ψ only.

Proof. (1) Let $\psi_1(\mu) = \mu$, corresponding to $\varphi_1(t) = 1/(1+\log(1/t))$, and let $\psi \in \mathcal{D}$ so that one has $\mu\psi(\lambda) \le c\psi(\mu\lambda)$ for $\lambda,\mu \ge 1$. Assume $h \in M^0_{\varphi_1}(\Omega,\mathcal{F},\mathbb{P})$ to be non-negative. Then

$$(27) \qquad |\psi^{-1}(1+h)|_{M^0_\varphi} \le c|1+h|_{M^0_{\varphi_1}}.$$

Indeed, if $a := |1+h|_{M^0_{\varphi_1}} > 0$, then for $\mu \ge ac$ we derive

$$\mathbb{P}(\psi^{-1}(1+h) > \mu) = \mathbb{P}(1+h > \psi(\mu)) \le e^{1-\psi(\mu)c/(ac)} \le e^{1-\psi(\mu/(ac))}.$$

(2) Let $g(t) := \overline{\psi}^{-1}(1+f(t))$. Lemma 4.4(2) and our first step give

$$|g|_{M^0_{\overline{\varphi}}[0,1]} \le c_\psi|1+f|_{M^0_{\varphi_1}[0,1]} \le c_\psi\|1+f\|_{\mathrm{BMO}_{\psi_1}[0,1]}.$$

(3) Now let $k \ge 1$, $B \in \mathcal{F}_k$ be an atom and $\widetilde{B} \in \mathcal{F}_{k-1}$ be the dyadic predecessor. If $g_l := \mathbb{E}(g \mid \mathcal{F}_l)$ and $f_l := \mathbb{E}(f \mid \mathcal{F}_l)$ $(l = 0,1,2,\ldots)$, then for $t \in \widetilde{B}$ one has

$$|g(t) - g_{k-1}(t)| \le \frac{1}{|\widetilde{B}|}\int_{\widetilde{B}} |g(t) - g(s)|\,ds.$$

For $r,s \ge 0$ we use $\overline{\psi}^{-1}(r+s+1) \le \overline{\psi}^{-1}(r+1) + \overline{\psi}^{-1}(s+1)$, which follows from Lemma 4.4(1), and $\overline{\psi}^{-1}(r+1) \le (r+1)\overline{\psi}^{-1}(2)$ (consider $k \le r \le k+1$ for $k \in \{0,1,2,\ldots\}$ and $\overline{\psi}^{-1}(r+1) \le \overline{\psi}^{-1}(k+2) \le (k+1)\overline{\psi}^{-1}(2) \le (r+1)\overline{\psi}^{-1}(2)$) and continue for $t \in \widetilde{B}$ to get

$$\begin{aligned}
|g(t) - g_{k-1}(t)| &\le \frac{1}{|\widetilde{B}|}\int_{\widetilde{B}} \overline{\psi}^{-1}(|f(t) - f(s)| + 1)\,ds\\
&\le \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + \frac{1}{|\widetilde{B}|}\int_{\widetilde{B}} \overline{\psi}^{-1}(|f_{k-1}(t) - f(s)| + 1)\,ds\\
&\le \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + \overline{\psi}^{-1}(2)\frac{1}{|\widetilde{B}|}\int_{\widetilde{B}} (|f_{k-1}(t) - f(s)| + 1)\,ds\\
&\le \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + \overline{\psi}^{-1}(2)[\|f\|_{\mathrm{BMO}} + 1]\\
&\le \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + 2\overline{\psi}^{-1}(2)\|1+f\|_{\mathrm{BMO}}.
\end{aligned}$$

From (27) and $|\cdot|_{M^0_{\varphi_1}(B,\mathbb{P}_B)} \le \|\cdot\|_{M_{\varphi_1}(B,\mathbb{P}_B)}$ it follows that

$$|\overline{\psi}^{-1}(|f - f_{k-1}| + 1)|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)} \le c_\psi\|f - f_{k-1}\|_{M_{\varphi_1}(B,\mathbb{P}_B)} + c_\psi.$$

Applying [4] (Theorems 1.1.7, 2.4.8), Theorem 4.9, and Corollary 4.8 we obtain

$$\begin{aligned}
\|f - f_{k-1}\|_{M_{\varphi_1}(B,\mathbb{P}_B)} &= \lim_{l\to\infty} \|f_l - f_{k-1}\|_{M_{\varphi_1}(B,\mathbb{P}_B)}\\
&\le |||(f_l)_{l=0}^\infty|||_{\psi_1} \le c'_\psi\|f\|_{\mathrm{BMO}}.
\end{aligned}$$

Summarizing all the estimates (and using $\Delta(\overline{\varphi}) < \infty$ and (16)) we arrive at $|g - g_{k-1}|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)} \le d_\psi \|1 + f\|_{\mathrm{BMO}}$ for some $d_\psi > 0$ depending on ψ only. Now steps (2) and (3) combined with Remark 4.11 give the assertion of the lemma. ■

Proof of Theorem 1.3. In view of Theorem 4.6(21) we get (1)⇒(2). We show (2)⇒(1). For this purpose let $f(t) = \overline{\psi'}^{-1}(1 + \log(1/t)) \in L_1[0,1]$ and $f_k = \mathbb{E}_k f$. Then it follows from Theorem 4.6, our assumption, and Theorem 4.12 that

$$|f|_{M^0_{\overline{\varphi}}} \le \|f\|_{\mathrm{BMO}_{\overline{\psi}}} \le 6\psi^{-1}(3)\|f\|_{\mathrm{BMO}_\psi} \le 6c_2\psi^{-1}(3)\|f\|_{\mathrm{BMO}_{\psi'}} \le b$$

so that $\overline{\varphi}(t) \le b\overline{\varphi'}(t)$ for $0 \le t \le 1$. Via (20) we arrive at the desired result. ■

The above proof shows in particular that even martingale sequences $(f_k)_{k\in\mathbb{N}}$ can be used to distinguish BMO_ψ and $\mathrm{BMO}_{\psi'}$.

Remark 4.14. (1) BMO-norms for random variables are defined in [2] with the help of Orlicz norms. To compare this approach with the approach of this paper let $\psi \in \mathcal{D}$ and let $\widetilde{\psi} : [0,\infty) \to [0,\infty)$ be its extension with $\widetilde{\psi}(t) = t$ for $0 \le t \le 1$. We assume $e^{\widetilde{\psi}(t)}$ to be convex and get a Young function $\Phi(t) := e^{\widetilde{\psi}(t)} - 1$ with the corresponding Orlicz norm $\|f\|_{L_\Phi} := \inf\{c > 0 \mid \mathbb{E}\Phi(|f|/c) \le 1\}$ for $f \in L_0(\Omega, \mathcal{F}, \mathbb{P})$. Using the condition $\sup_{a>1}\inf_{\lambda\ge1}\psi(a\lambda)/\psi(\lambda) > 1$ one can verify $\|\cdot\|_{L_\Phi} \sim \|\cdot\|_{M_\varphi} \sim |\cdot|_{M^0_\varphi}$. Now consider for simplicity $\Omega = [0,1]$ and $f \in L_1[0,1]$. For $f_k := \mathbb{E}(f \mid \mathcal{F}_k)$ we obtain, as in Remark 4.11,

$$\begin{aligned}\|(f_k)_{k\ge0}\|_{\mathrm{BMO}_\psi} &\sim \sup_{k\ge0} \sup_{\substack{B\in\mathcal{F}_k\\ \text{atom}}} \|f - f_{k-1}\|_{M_\varphi(B,\mathbb{P}_B)}\\ &\sim \sup_{k\ge0} \sup_{\substack{B\in\mathcal{F}_k\\ \text{atom}}} \|f - f_{k-1}\|_{L_\Phi(B,\mathbb{P}_B)} = \sup_{k\ge0} \|\operatorname{ess\,inf} \mathcal{M}\|_\infty\end{aligned}$$

where $\mathcal{M}$ consists of all positive $\mathcal{F}_k$-measurable $\gamma \in L_1[0,1]$ such that

$$\mathbb{E}(\Phi(|f - f_{k-1}|/\gamma) \mid \mathcal{F}_k) \le 1 \quad \text{a.s.}$$

Hence we are in the situation of [2]. Theorem 6 of [2] contains a sufficient condition for the inclusion of two BMO-spaces generated by different Young functions. But a criterion like Theorem 1.3 of the present paper is not available. It seems that one can use (at least in many relevant cases) the results of this paper, especially the regularization $\overline{\psi}$ of ψ, which is the key of our approach, to extend [2] in some directions.

(2) Let $(f_k)_{k\in I} \in \mathrm{BMO}_\psi((\mathcal{F}_k)_{k\in I})$. If $[\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}] := [\Omega, \mathcal{F}, \mathbb{P}] \times [\Omega', \mathcal{F}', \mathbb{P}']$, $\widetilde{\mathcal{F}}_k := \mathcal{F}_k \times \mathcal{F}'_k$, and $g_k(\omega, \omega') := f_k(\omega)$, then $\|(f_k)_{k\in I}\|_{\mathrm{BMO}_\psi} = \|(g_k)_{k\in I}\|_{\mathrm{BMO}_\psi}$ because

$$\sup_{\widetilde{B}\in\widetilde{\mathcal{F}}_k} \widetilde{\mathbb{P}}_{\widetilde{B}}(|g_l - g_{k-1}| > \lambda) = \sup_{B\in\mathcal{F}_k} \mathbb{P}_B(|f_l - f_{k-1}| > \lambda)$$

for $\lambda > 0$ and $k \le l \in I$.

5. Weak upper estimates of BMO$_\psi$-spaces. Now we will prove Theorem 1.5 and deduce in Corollary 5.3 that the BMO$_\psi$-spaces can also be classified by their weak upper estimates.

LEMMA 5.1. *For $1 \ge a_1 \ge \ldots \ge a_N > 0$ and $\varphi \in \mathcal{C}$ the following are equivalent.*

(1) $\sum_{i=1}^N \varphi^{-1}(a_i/\lambda) \le \varphi^{-1}(1/\lambda)$ *for all* $\lambda \ge 1$.

(2) $|\sup_{1\le i\le N} a_i|f_i||_{M^0_\varphi} \le \sup_{1\le i\le N} |f_i|_{M^0_\varphi}$ *for all* $[\Omega, \mathcal{F}, \mathbb{P}]$ *and* $(f_i)_{i=1}^N \subseteq M^0_\varphi(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. (1)⇒(2). Assume that $\sup_{1\le i\le N} |f_i|_{M^0_\varphi} = 1$. Then for $\lambda \ge 1$,

$$\begin{aligned}\mathbb{P}(\sup_i a_i|f_i| > \lambda) &\le \sum_i \mathbb{P}(|f_i| > \lambda/a_i) \le \sum_i e^{1-\psi(\lambda/a_i)}\\ &= \sum_i \varphi^{-1}(a_i/\lambda) \le \varphi^{-1}(1/\lambda) = e^{1-\psi(\lambda)}.\end{aligned}$$

(2)⇒(1). We can assume $\lambda > 1$ and set $t_i := \varphi^{-1}(a_i/\lambda)$. Choose $A_i \subseteq [0,1]$ such that $|A_i| = t_i$ and define $f_i := (1/a_i)\chi_{A_i}$. Then

$$\begin{aligned}|\sup_i \chi_{A_i}|_{M^0_\varphi[0,1]} &= \Big|\sup_i a_i \frac{1}{a_i}\chi_{A_i}\Big|_{M^0_\varphi[0,1]}\\ &\le \sup_i \frac{|\chi_{A_i}|_{M^0_\varphi[0,1]}}{a_i} = \sup_i \frac{\varphi(t_i)}{a_i} = \frac{1}{\lambda} < 1.\end{aligned}$$

Since $|\chi_{[0,1]}|_{M^0_\varphi} = 1$ one can find disjoint A_i with $\mathbb{P}(A_i) = t_i$ and get for $\lambda > 1$, by applying φ^{-1} to the inequality above,

$$\sum_{i=1}^N t_i = \varphi^{-1}\Big(\varphi\Big(\sum_{i=1}^N t_i\Big)\Big) = \varphi^{-1}(|\sup_i \chi_{A_i}|_{M^0_\varphi[0,1]}) \le \varphi^{-1}(1/\lambda). \ ■$$

LEMMA 5.2. *For $\psi \in \mathcal{D}$ with $\sup_{a>1}\inf_{\lambda\ge1}\psi(a\lambda)/\psi(\lambda) > 1$ there is some $c > 0$ such that for all $[\Omega, \mathcal{F}, \mathbb{P}]$ and all $(f_i)_{i=1}^N \subseteq M^0_\varphi(\Omega, \mathcal{F}, \mathbb{P})$,*

$$|\sup_{1\le i\le N} \varphi(1/i)|f_i||_{M^0_\varphi} \le c \sup_{1\le i\le N} |f_i|_{M^0_\varphi}.$$

Proof. Assuming $\psi(a\lambda) \ge (1+\varepsilon)\psi(\lambda)$ for some $a > 1$ and $\varepsilon > 0$ we get $\psi(a^n\lambda) \ge (1+\varepsilon)^n\psi(\lambda)$. Consequently, we find some $b > 1$ with

$\inf_{\lambda\geq 1}\psi(b\lambda)/\psi(\lambda)\geq 3$. This gives $2\psi(\mu)+\psi(\lambda)\leq 3\psi(\mu\lambda)\leq\psi(b\mu\lambda)$ and $2\log i+\psi(\lambda)=2\psi[\psi^{-1}(1+\log i)]+\psi(\lambda)-2\leq\psi[b\psi^{-1}(1+\log i)\lambda]=\psi(\lambda/a_i)$ where

$$a_i:=\frac{1}{b\psi^{-1}(1+\log i)}=\frac{1}{b}\varphi\left(\frac{1}{i}\right).$$

Hence $e^{1-\psi(\lambda/a_i)}\leq\frac{1}{i^2}e^{1-\psi(\lambda)}$ and

$$\sum_{i=1}^{N}\varphi^{-1}\left(\frac{a_i}{\lambda}\right)=\sum_{i=1}^{N}e^{1-\psi(\lambda/a_i)}\leq\left(\sum_{i=1}^{\infty}\frac{1}{i^2}\right)e^{1-\psi(\lambda)}=e^{\log(\sum_{i=1}^{\infty}1/i^2)+1-\psi(\lambda)}.$$

If we choose $c\geq 1$ such that

$$\log\left(\sum_{i=1}^{\infty}\frac{1}{i^2}\right)\leq\psi(c),$$

then for $\lambda>bc$ we get

$$\log\left(\sum_{i=1}^{\infty}\frac{1}{i^2}\right)+\psi\left(\frac{\lambda}{bc}\right)\leq\psi(c)+\psi\left(\frac{\lambda}{b}\right)\leq 2\psi\left(\frac{\lambda}{b}\right)\leq\psi(\lambda)$$

and

$$\sum_{i=1}^{N}\varphi^{-1}\left(\frac{a_i}{\lambda}\right)\leq e^{\log(\sum_{i=1}^{\infty}1/i^2)+1-\psi(\lambda)}\leq e^{1-\psi(\lambda/(bc))}=\varphi^{-1}\left(\frac{bc}{\lambda}\right)$$

or

$$\sum_{i=1}^{N}\varphi^{-1}\left(\frac{a_i}{bc\lambda}\right)\leq\varphi^{-1}\left(\frac{1}{\lambda}\right)\qquad\text{for }\lambda\geq 1.$$

Now we can apply Lemma 5.1. ■

Proof of Theorem 1.5. Since for $a_i=1/\overline{\psi}^{-1}(1+\log i)=\overline{\varphi}(1/i)$,

$$\begin{aligned}\|(\sup_{i\geq 1}a_i|f_k^{(i)}|)_{k\in I}\|_{\mathrm{BMO}_\psi}&\leq\sup_{0\leq k\leq l\in I}\ \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}}\|\sup_{i\geq 1}a_i|f_l^{(i)}|-\sup_{i\geq 1}a_i|f_{k-1}^{(i)}|\|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)}\\&\leq\sup_{0\leq k\leq l\in I}\ \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}}\|\sup_{i\geq 1}a_i|f_l^{(i)}-f_{k-1}^{(i)}|\|_{M^0_{\overline{\varphi}}(B,\mathbb{P}_B)}\end{aligned}$$

we can deduce from Lemmas 5.2 and 4.4(2) that

$$\|(\sup_{i\geq 1}a_i|f_k^{(i)}|)_{k\in I}\|_{\mathrm{BMO}_\psi}\leq c\sup_{i\geq 1}\|(f_k^{(i)})_{k\in I}\|_{\mathrm{BMO}_{\overline{\psi}}}.$$

Applying Theorem 4.6(21) gives $\mathbf{U}_a(\mathrm{BMO}_\psi)\leq c6\psi^{-1}(3)$. On the other hand, we take the function

$$f(t):=\overline{\psi}^{-1}\left(1+\log\frac{1}{t}\right)=\frac{1}{\overline{\varphi}(t)}\in L_1[0,1]\tag{28}$$

exploited in Theorem 4.12. Letting $f_k:=\mathbb{E}(f\mid\mathcal{F}_k)$ we define, for fixed $N\geq 1$,

$$f_k^{(i)}(t_1,\ldots,t_N):=f_k(t_i)\in L_1([0,1]^N)\qquad(1\leq i\leq N).$$

On $[0,1]^N$ we will use the canonical product filtration $(\mathcal{F}_k^N)_{k=0}^{\infty}$ with $\mathcal{F}_k^N:=\bigtimes_{k=1}^{N}\mathcal{F}_k$, where $(\mathcal{F}_k)_{k=0}^{\infty}$ is the dyadic filtration on $[0,1]$ taken in Section 4. Theorem 1.8 and (18) imply

$$\frac{1}{2}\cdot\frac{1}{\overline{\varphi}(1/N)}=\frac{1}{2}\widetilde{f}\left(\frac{1}{N}\right)\leq\|\sup_{1\leq i\leq N}f(t_i)\|_{L_1([0,1]^N)}.$$

By Fatou's lemma and Theorem 4.9 we can continue with

$$\begin{aligned}&\|\sup_{1\leq i\leq N}f(t_i)\|_{L_1([0,1]^N)}\\&\quad\leq\liminf_{l\to\infty}\|\sup_{1\leq i\leq N}f_l^{(i)}\|_{L_1([0,1]^N)}\leq|||(\sup_{1\leq i\leq N}f_k^{(i)})_{k=0}^{\infty}|||_{\overline{\psi}}\\&\quad\leq 24\overline{\psi}^{-1}(3)\psi^{-1}(3)\|(\sup_{1\leq i\leq N}f_k^{(i)})_{k=0}^{\infty}\|_{\mathrm{BMO}_\psi((\mathcal{F}_k^N)_{k=0}^{\infty})}.\end{aligned}$$

Summarizing all estimates yields for $1/\theta:=48\overline{\psi}^{-1}(3)\psi^{-1}(3)$,

$$\begin{aligned}\theta&\leq\overline{\varphi}(1/N)\|(\sup_{1\leq i\leq N}f_k^{(i)})_{k=0}^{\infty}\|_{\mathrm{BMO}_\psi((\mathcal{F}_k^N)_{k=0}^{\infty})}\\&\leq\mathbf{U}_{b^N}(\mathrm{BMO}_\psi)\sup_{1\leq i\leq N}\|(f_k^{(i)})_{k=0}^{\infty}\|_{\mathrm{BMO}_\psi((\mathcal{F}_k^N)_{k=0}^{\infty})}.\end{aligned}$$

As by Remark 4.14(2) and Theorem 4.12, $\|(f_k^{(i)})_{k=0}^{\infty}\|_{\mathrm{BMO}_\psi}=\|f\|_{\mathrm{BMO}_\psi}<\infty$, we are done. ■

The next corollary adds a further criterion to Theorem 1.3.

Corollary 5.3. *For $\psi,\psi'\in\mathcal{D}$ the following assertions are equivalent.*

(1) *There exists $c_1\geq 1$ such that $\overline{\psi}(\mu)\leq\overline{\psi'}(c_1\mu)$ for all $1\leq\mu<\infty$.*

(2) *There exists $c_2>0$ such that for all $a=(a_i)_{i\geq 1}$ satisfying $1\geq a_1\geq a_2\geq\ldots\geq 0$ one has*

$$\mathbf{U}_a(\mathrm{BMO}_{\psi'})\leq c_2\mathbf{U}_a(\mathrm{BMO}_\psi).$$

Proof. First for $\varphi\in\mathcal{C}$ and $N=1,2,\ldots$ set

$$a_\varphi:=(\varphi(1/i))_{i=1}^{\infty}\quad\text{and}\quad e_\varphi^N:=(\varphi(1/N),\ldots,\varphi(1/N),0,0,\ldots)$$

where $\varphi(1/N)$ is repeated N times $(\varphi(1/i)=1/\psi^{-1}(1+\log i))$.

(1)⇒(2). If $\mathbf{U}_a(\mathrm{BMO}_\psi)<\infty$, then it follows from Theorem 1.5 and (20) that

$$0<\theta:=\inf_N \mathbf{U}_{e^N_{\overline{\varphi}}}(\mathrm{BMO}_\psi)\le \inf_N \frac{\overline{\varphi}(1/N)}{a_N}\mathbf{U}_a(\mathrm{BMO}_\psi)\le c_1\inf_N\frac{\overline{\varphi'}(1/N)}{a_N}\mathbf{U}_a(\mathrm{BMO}_\psi)$$

or $a_N\le d\overline{\varphi'}(1/N)$ with $d:=(c_1/\theta)\mathbf{U}_a(\mathrm{BMO}_\psi)$. Hence

$$\mathbf{U}_a(\mathrm{BMO}_{\psi'})\le d\mathbf{U}_{a_{\overline{\varphi'}}}(\mathrm{BMO}_{\psi'})=\left[\frac{c_1}{\theta}\mathbf{U}_{a_{\overline{\varphi'}}}(\mathrm{BMO}_{\psi'})\right]\mathbf{U}_a(\mathrm{BMO}_\psi)=c_2\mathbf{U}_a(\mathrm{BMO}_\psi)$$

where $c_2<\infty$ according to Theorem 1.5.

(2)⇒(1). Since for $\theta':=\inf_N\mathbf{U}_{e^N_{\overline{\varphi'}}}(\mathrm{BMO}_{\psi'})>0$,

$$\theta'\frac{\overline{\varphi}(1/N)}{\overline{\varphi'}(1/N)}\le \mathbf{U}_{e^N_{\overline{\varphi'}}}(\mathrm{BMO}_{\psi'})\frac{\overline{\varphi}(1/N)}{\overline{\varphi'}(1/N)}=\mathbf{U}_{e^N_{\overline{\varphi}}}(\mathrm{BMO}_{\psi'})\le c_2\mathbf{U}_{e^N_{\overline{\varphi}}}(\mathrm{BMO}_\psi)\le c_2\mathbf{U}_{a_{\overline{\varphi}}}(\mathrm{BMO}_\psi),$$

for $c_2':=(c_2/\theta')\mathbf{U}_{a_{\overline{\varphi}}}(\mathrm{BMO}_\psi)$ and $N=1,2,\ldots$ one has the estimate $\overline{\varphi}(1/N)\le c_2'\overline{\varphi'}(1/N)$. Since $\Delta(\overline{\varphi'})<\infty$ according to Lemma 4.4(2), a standard computation gives $\overline{\varphi}(t)\le c_2'\Delta(\overline{\varphi'})\overline{\varphi'}(t)$ for $0\le t\le 1$. Using (20) we can conclude the proof. ■

6. A general extrapolation principle.

We want to prove Theorem 1.7. For simplicity we assume that $I=\{0,\ldots,n\}$ to avoid difficulties with the existence of the maximal operator A^*. Let us start with some lemmata.

Lemma 6.1. *Let $\psi\in\mathcal{D}$, and $S,T:\mathcal{A}^X((\Sigma_k)_{k=0}^n)\supseteq F\to L_0^+(M,\Sigma,m)$, where X is a Banach space. If (F,S,T) satisfies* (EP) *with constant $c\ge1$, then*

$$\|(Sf^k)_{k=0}^n\|_{\mathrm{BMO}^*_\psi}\le\|Tf\|_\infty\quad(f\in F)$$

implies, for $0<\delta\le1$ and $\lambda>0$,

$$m(S^*f>(c+c^2)\lambda,\ Tf\le\delta\lambda)\le e^{1-\overline{\psi}(1/\delta)}m(cS^*f>\lambda)\quad(f\in F).$$

Proof. Fix $f\in F$, δ, and λ. We find some $g\in F$ satisfying $\|Tg\|_\infty\le c\delta\lambda$,

$$\frac1c\chi_{\{Tf\le\delta\lambda\}}S^*f\le S^*g\le cS^*f\quad\text{a.s.}$$

and

$$\left\|\left(\frac{Sg^k}{c\delta\lambda}\right)_{k=0}^n\right\|_{\mathrm{BMO}^*_\psi}\le\left\|\frac{Tg}{c\delta\lambda}\right\|_\infty\le1.$$

Theorem 4.6(b) gives, for $\lambda':=1/(c\delta)$ and $\mu':=1/\delta$,

$$\begin{aligned}m(S^*f>(c+c^2)\lambda,\ Tf\le\delta\lambda)&\le m(S^*g>(1+c)\lambda)=m\left(\frac{S^*g}{c\delta\lambda}>\lambda'+\mu'\right)\\&\le e^{1-\overline{\psi}(1/\delta)}m\left(\frac{S^*g}{c\delta\lambda}>\lambda'\right)=e^{1-\overline{\psi}(1/\delta)}m(S^*g>\lambda)\\&\le e^{1-\overline{\psi}(1/\delta)}m(cS^*f>\lambda).\ ■\end{aligned}$$

In the following we will sometimes use the fact that $(E,\alpha A,\beta B)$ has (EP) with constant $c>0$ whenever $\alpha,\beta\in\mathbb{R}$ are positive and (E,A,B) has (EP) with constant $c>0$. Furthermore, given Banach spaces $X_1,\ldots,X_N$ we make use of the Banach space $\ell_\infty^N(X_i)$ consisting of all N-tuples $(x_1,\ldots,x_N)$ of $x_i\in X_i$ equipped with the norm $\|(x_i)_{i=1}^N\|_{\ell_\infty^N(X_i)}:=\max_i\|x_i\|_{X_i}$.

Lemma 6.2. *Assume that $A_i,B_i:\mathcal{A}^{X_i}((\mathcal{F}_k^i)_{k=0}^n)\supseteq E_i\to L_0^+(\Omega_i,\mathcal{F}^i,\mathbb{P}_i)$ satisfy* (EP) *with constant $c\ge1$, $[M,\Sigma,m]:=\times_{i=1}^N[\Omega_i,\mathcal{F}^i,\mathbb{P}_i]$, and $\Sigma_k:=\times_{i=1}^N\mathcal{F}_k^i$. Let $F\subseteq\mathcal{A}^{\ell_\infty^N(X_i)}((\Sigma_k)_{k=0}^n)$ be given by*

$$(f^{(1)},\ldots,f^{(N)}):=((d_k^{(1)},\ldots,d_k^{(N)}))_{k=0}^n\in F\quad\textit{iff}\quad f^{(i)}=(d_k^{(i)})_{k=0}^n\in E_i.$$

Define $S,T:F\to L_0^+(M,\Sigma,m)$ by

$$S(f^{(1)},\ldots,f^{(N)}):=\sup_i A_if^{(i)}\quad\textit{and}\quad T(f^{(1)},\ldots,f^{(N)}):=\sup_i B_if^{(i)}.$$

Then (F,S,T) has (EP) *with constant $c\ge1$.*

Proof. It is clear that the operator S is measurable. Now let $\lambda>0$ and $f\in F$ with $f:=(f^{(1)},\ldots,f^{(N)})$ and $f^{(i)}\in E_i$. We find $g^{(i)}\in E_i$ such that

$$\frac1c\chi_{\{B_if^{(i)}\le\lambda\}}A_i^*f^{(i)}\le A_i^*g^{(i)}\le cA_i^*f^{(i)}\text{ a.s.}\quad\text{and}\quad B_ig^{(i)}\le c\lambda\text{ a.s.}$$

For $g:=(g^{(1)},\ldots,g^{(N)})\in F$ one gets a.s. $S^*g=\sup_iA_i^*g^{(i)}\le c\sup_iA_i^*f^{(i)}=cS^*f$, $Tg=\sup_iB_ig^{(i)}\le c\lambda$, and

$$\frac1c\chi_{\{Tf\le\lambda\}}S^*f\le\frac1c\sup_{1\le i\le N}[\chi_{\{B_if^{(i)}\le\lambda\}}A_i^*f^{(i)}]\le S^*g.\ ■$$

Proof of Theorem 1.7. (1) Let $a:=(1/\overline{\psi}^{-1}(1+\log i))_{i=1}^\infty=(\overline{\varphi}(1/i))_{i=1}^\infty$ and $c':=6\psi^{-1}(3)\mathbf{U}_a(\mathrm{BMO}_\psi)$. Using Lemma 6.2 for

$$[\Omega_i,\mathcal{F}^i,\mathbb{P}_i]:=[\Omega,\mathcal{F},\mathbb{P}],\quad E_i:=E,\quad A_i:=\overline{\varphi}(1/i)A,\quad\text{and}\quad B_i:=c'B$$

implies that (F,S,T) defined as in Lemma 6.2 has (EP) with $c\ge1$. For $f=(f^{(1)},\ldots,f^{(N)})\in F$ and $h_k^{(i)}(\omega_1,\ldots,\omega_N):=A((f^{(i)})^k)(\omega_i)\in L_0(M)$, from Theorems 4.6 and 1.5 we get

$$\frac{1}{6\psi^{-1}(3)}\|(Sf^k)_{k=0}^n\|_{\mathrm{BMO}_\psi^*} \le \|(Sf^k)_{k=0}^n\|_{\mathrm{BMO}_\psi}$$
$$\le \mathbf{U}_a(\mathrm{BMO}_\psi) \sup_{1\le i\le N} \|(h_k^{(i)})_{k=0}^n\|_{\mathrm{BMO}_\psi}.$$

Applying Remark 4.14(2) gives

$$\|(h_k^{(i)})_{k=0}^n\|_{\mathrm{BMO}_\psi} = \|(A(f^{(i)})^k)_{k=0}^n\|_{\mathrm{BMO}_\psi}$$

so that our assumption yields

$$\|(Sf^k)_{k=0}^n\|_{\mathrm{BMO}_\psi^*} \le c' \sup_{1\le i\le N} \|(A(f^{(i)})^k)_{k=0}^n\|_{\mathrm{BMO}_\psi}$$
$$\le c' \sup_{1\le i\le N} \|Bf^{(i)}\|_\infty = \|Tf\|_\infty.$$

Lemma 6.1 implies

$$m(cS^*f > c(c+c^2)\lambda,\ Tf \le \delta\lambda) \le e^{1-\overline{\psi}(1/\delta)} m(cS^*f > \lambda)$$

for $0<\delta\le 1$ and $\lambda>0$. We choose $0<\delta\le 1$ with $e^{1-\overline{\psi}(1/\delta)} = [2c(c+c^2)]^{-p}$ and deduce for $\beta = c(c+c^2)$ and $\varepsilon = e^{1-\overline{\psi}(1/\delta)}$ that

$$\frac{\beta}{\delta}\cdot\frac{1}{\sqrt[p]{1-\beta^p\varepsilon}} \le \frac{2c(c+c^2)}{\delta} = 2c(c+c^2)\overline{\psi}^{-1}(1+p\log[2c(c+c^2)]).$$

Lemma 4.4(2) applied to $\mu = ac_\psi$, where $a>1$, yields $\overline{\psi}^{-1}(ap) \le ac_\psi\overline{\psi}^{-1}(p)$ for $p\ge 1$. Hence Lemma 4.5 gives for some $d>0$, depending on c and ψ only,

$$\|c \sup_{1\le i\le N} \overline{\varphi}(1/i)A^*f(\omega_i)\|_{L_p(\Omega^N)} = \|cS^*f\|_p \le \frac{\beta}{\delta}\cdot\frac{1}{\sqrt[p]{1-\beta^p\varepsilon}}\|Tf\|_p$$
$$\le d\overline{\psi}^{-1}(p)\|Tf\|_p$$
$$= d\overline{\psi}^{-1}(p)c'\| \sup_{1\le i\le N} Bf(\omega_i)\|_{L_p(\Omega^N)}.$$

(2) Let $1\le t<\infty$ with $N\le t<N+1$ and

$$w := \sum_{i=1}^N \overline{\varphi}(1/i)\chi_{[(i-1)/N,i/N)} \in L_0^+[0,1] \quad \text{so that} \quad w_t^{\overline{\psi}} \le \Delta(\overline{\varphi})w \text{ a.s.}$$

Then

$$K_{\infty,p}^{w_t^{\overline{\psi}}}[A^*f, t^{1/p}] \le \Delta(\overline{\varphi})K_{\infty,p}^w[A^*f,t^{1/p}] \le 2^{1/p}\Delta(\overline{\varphi})K_{\infty,p}^w[A^*f,N^{1/p}].$$

Exploiting Theorem 1.8 we can finish with

$$K_{\infty,p}^{w_t^{\overline{\psi}}}[A^*f,t^{1/p}] \le 2^{1/p+1}\Delta(\overline{\varphi})\| \sup_{1\le i\le N} \overline{\varphi}(1/i)A^*f(\omega_i)\|_{L_p(\Omega^N)}$$
$$\le 2^{1/p+1}\Delta(\overline{\varphi})\frac{dc'}{c}\overline{\psi}^{-1}(p)\| \sup_{1\le i\le N} Bf(\omega_i)\|_{L_p(\Omega^N)}$$
$$\le 2^{1/p+1}\Delta(\overline{\varphi})\frac{dc'}{c}\overline{\psi}^{-1}(p)K_{\infty,p}[Bf,N^{1/p}]$$
$$\le 2^{1/p+1}\Delta(\overline{\varphi})\frac{dc'}{c}\overline{\psi}^{-1}(p)K_{\infty,p}[Bf,t^{1/p}]. \ \blacksquare$$

From Theorems 1.7 and 4.6(23) we can immediately deduce

COROLLARY 6.3. *Let $A,B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$ where X is a Banach space, and let $\psi_1(t)=t$. Assume that (E,A,B) satisfies* (EP) *with constant $c\ge 1$ and that*

$$\sup_{f\in E}\ \sup_{0\le k\le l\le n}\ \sup_{\substack{B\in\mathcal{F}_k\\ \mathbb{P}(B)>0}} \mathbb{P}_B(|Af^l - Af^{k-1}| > \|Bf\|_\infty) = s < 1/2.$$

Then there is some $d>0$, depending on s and c only, such that for all $1\le p<\infty$, $t\ge 1$, and $f\in E$,

$$K_{\infty,p}^{w_t^{\psi_1}}[A^*f,t^{1/p}] \le dpK_{\infty,p}[Bf,t^{1/p}]. \tag{29}$$

In particular, for $t=1$ one has $\|A^*f\|_p \le dp\|Bf\|_p$.

Remark 6.4. (1) The inequality $\|A^*f\|_p \le d\overline{\psi}^{-1}(p)\|Bf\|_p$ of Theorem 1.7 implies estimates which are weaker than (12). For example, starting with inequality (2) we get for $\lambda>0$ the estimate

$$\lambda\Big(\mathbb{P}\Big(\sup_{1\le k\le n}\Big|\sum_{i=1}^k d_i\Big| > \lambda\|S_2f\|_p\Big)\Big)^{1/p} \le c\sqrt{p}.$$

Given $\lambda \ge ce =: a \ge 1$ we choose p such that $1/e = c\sqrt{p}/\lambda$ and obtain

$$\mathbb{P}\Big(\sup_{1\le k\le n}\Big|\sum_{i=1}^k d_i\Big| > \lambda\|S_2f\|_{(\lambda/a)^2}\Big) \le e^{-(\lambda/a)^2}. \tag{30}$$

Since

$$\|S_2f\|_{(\lambda/a)^2} \ge \sup_t t^{(a/\lambda)^2}(S_2f)^{**}(t)$$
$$\ge e^{(1-\lambda^2)a^2/\lambda^2}(S_2f)^{**}(e^{1-\lambda^2}) \ge e^{-a^2}(S_2f)^{**}(e^{1-\lambda^2})$$

and (18) holds, the estimate (30) is weaker than (7) which implies

$$\mathbb{P}\Big(\sup_k\Big|\sum_{i=1}^k d_i\Big| > c\lambda K_{\infty,1}[S_2f,e^{\lambda^2-1}]\Big) \le e^{1-\lambda^2}.$$

(2) There are examples showing that (13) is weaker than (12). To see this we first remark that according to (18) and $w_t = 1$ on $[0, 1/t]$ (w_t is defined in Theorem 1.7),

$$K^{w_t}_{\infty,p}[A^*f, t^{1/p}] \geq (A^*f \cdot w_t)^{**}_p(1/t) \geq \|A^*f\|_p.$$

Now, consider $1 \leq p < q < \infty$, $\psi(t) = \overline{\psi}(t) = t^p$ (see Example 4.3), and $\Omega = [0,1]$, where we use the dyadic filtration $(\mathcal{F}_k)_{k=0}^\infty$ from Section 4. Let $A : E_n \to L_0^+[0,1]$, where E_n is the set of all mean-zero martingale differences with respect to $(\mathcal{F}_k)_{k=0}^n$ and $A((d_k)_{k=0}^n) := |\sum_{k=0}^n d_k|$. If $f_n := (df_k)_{k=0}^n \in E_n$ is the sequence of martingale differences generated by

$$f(t) := \sqrt[q]{1+\log(1/t)} - \|\sqrt[q]{1+\log(1/t)}\|_1 \in L_1[0,1],$$

then Theorems 4.12 and 4.6 imply, for $h = \sup_n(A^*f_n)$,

$$\widetilde{h}(t) \leq c\sqrt[q]{1+\log(1/t)} \quad \text{for } 0 < t \leq 1.$$

Now one can show $K_{\infty,p}[\sup_n A^*f_n, t^{1/p}] \leq c'\sqrt[q]{1+\log t}$ for $1 \leq t < \infty$ so that

$$\lim_{\substack{t\to\infty \\ t\geq 1}} K_{\infty,p}[\sup_n A^*f_n, t^{1/p}]/\overline{\psi}^{-1}(1+\log t) = 0$$

but

$$\inf_{\substack{n \\ t\geq 1}} K^{w_t}_{\infty,p}[A^*f_n, t^{1/p}] \geq \|A^*f_1\|_p > 0.$$

7. Examples for condition (EP). Several extrapolation procedures use assumptions as in Proposition 7.3 ($C = 0$) or the weaker assumptions considered in Theorem 1.1. Via condition (EP) we have chosen a more abstract way; that is done for several reasons. For instance, property (EP) is stable with respect to procedures like that in Lemma 6.2. Moreover, in the first example we see that (EP) includes a known classical situation (Proposition 7.3 with $C = 0$). A slight modification leads to (EP) for (E, A, B) where B is a special non-local operator, used in Corollary 7.4. In the second example we demonstrate that (EP) also includes the situation of Theorem 1.1.

Finally, in Proposition 7.6 we show that the assumptions of Proposition 7.3 (where $C = 0$) imply the equivalence of the K-functional $K_{\infty,p}[Bf, t]$ and $K(f, t; H^B_\infty, H^B_p)$, the K-functional with respect to the Hardy spaces generated by B. This observation can be found with more restrictive assumptions in [26] and is prepared in Lemma 7.2 for our purpose.

The first example. We will say that a subset $E \subseteq \mathcal{A}^X((\mathcal{F}_k)_{k\in I})$ is *closed under starting and stopping* provided that $f \in E$ implies ${}^\tau f, f^\tau \in E$ for all stopping times τ. Furthermore, an operator $A : \mathcal{A}^X((\mathcal{F}_k)_{k\in I}) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$ is called *predictable* if Af^k is $\mathcal{F}_{k-1}$-measurable for $f \in E$ and $1 \leq k \in I$.

First let us recall a basic property of local and quasilinear operators (see [11] (Lemma 2.1)).

LEMMA 7.1. *Let $T : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume T to be quasilinear with constant $\gamma_T \geq 1$ and local. Then for every stopping time τ, $f \in E$, and $k \in \{0, \ldots, n\}$ one has a.s.*

$$T(f^\tau) \leq \gamma_T^2 T(f^k) \text{ on } \{\tau = k\} \quad \text{and} \quad T({}^\tau f) \leq 2\gamma_T^3 (T^*f)\chi_{\{\tau<n\}}.$$

Proof. On $\{\tau = k\}$ we get a.s.

$$\begin{aligned} T(f^\tau) &\leq \gamma_T[T(f^\tau - f^{\tau\wedge k}) + T(f^{\tau\wedge k})] = \gamma_T T(f^{\tau\wedge k}) \\ &\leq \gamma_T^2[T(f^{\tau\wedge k} - f^k) + T(f^k)] = \gamma_T^2 T(f^k). \end{aligned}$$

For the second inequality we remark that $T({}^\tau f) = 0$ a.s. on $\{\tau = n\}$. For $0 \leq k < n$ from the first step we obtain $T({}^\tau f) \leq \gamma_T[Tf + T(f^\tau)] \leq \gamma_T[Tf + \gamma_T^2 T(f^k)] \leq 2\gamma_T^3 T^*f$ a.s. on $\{\tau = k\}$. ∎

The next lemma is proved for $0 < p \leq 1$ and for σ-sublinear operators using an atomic decomposition in [26] (Theorem 5). Our approach is more direct. As in [26] we will deduce in Proposition 7.6 a relation between different K-functionals, now available for all $0 < p < \infty$ and for quasilinear operators.

LEMMA 7.2. *Let $T : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume T to be quasilinear with constant $\gamma_T \geq 1$, local, and predictable. Let $\lambda \geq 0$, $f \in E$, and let us define $\tau := \inf\{k \mid T(f^{k+1}) > \lambda\} \wedge n$. Then for all $0 < p < \infty$ one has*

$$T(f^\tau) \leq \gamma_T^2\lambda \text{ a.s.} \quad \text{and} \quad \|T({}^\tau f)\|_p \leq 2\gamma_T^3\Big(\int_{\{T^*f>\lambda\}} (T^*f)^p\,d\mathbb{P}\Big)^{1/p}.$$

Proof. Use Lemma 7.1, $T(f^k) \leq \lambda$ a.s. on $\{\tau = k\}$, and $\{\tau < n\} = \{T^*f > \lambda\}$. ∎

PROPOSITION 7.3. *Let $A, B, C : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume that A is quasilinear with constant $\gamma_A \geq 1$, local, and measurable, that B is quasilinear with constant $\gamma_B \geq 1$, local, predictable, and monotone, and that $Cf = c(f)\chi_\Omega$ where $c(f) \in \mathbb{R}$ with $c(0) = 0$ and $c(f^\tau) \leq c(f)$ for all stopping times τ. Then $(E, A, B + C)$ has (EP) with constant $c = \max(\gamma_A^2, \gamma_B^2 + 1)$.*

Proof. Let $\lambda > 0$ and $f \in E$. If $c(f) > \lambda$, then we choose $g = 0$ such that

$$\chi_{\{B+C\leq\lambda\}}A^*f = A^*g = 0 \leq A^*f \text{ a.s.} \quad \text{and} \quad (B+C)g = 0 \text{ a.s.}$$

If $c(f) \le \lambda$, then we choose $g := f^\tau \in E$ where $\tau := \inf\{k \mid B(f^{k+1}) > \lambda\} \wedge n$. Lemma 7.2 gives $Bg = B(f^\tau) \le \gamma_B^2\lambda$ so that $(B+C)g \le Bg + Cf \le (\gamma_B^2+1)\lambda$ a.s. on Ω. Since $\{Bf \le \lambda\} = \{\tau = n\}$ we have

$$\frac{1}{\gamma_A}\chi_{\{(B+C)f\le\lambda\}}A^*f \le \frac{1}{\gamma_A}\chi_{\{Bf\le\lambda\}}A^*f = \frac{1}{\gamma_A}\chi_{\{\tau=n\}}A^*(f - f^\tau + f^\tau)$$
$$\le \chi_{\{\tau=n\}}[A^*(f-f^\tau) + A^*(f^\tau)] \le A^*(f^\tau).$$

Finally, the left-hand side of the assertion of Lemma 7.1 implies $A^*(f^\tau) \le \gamma_A^2 A^*(f)$. ∎

Note that $B+C$ in the above proposition is not necessarily local. Proposition 7.3 and Theorem 1.7 imply

COROLLARY 7.4. *Let $A, B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume that A is quasilinear with constant $\gamma_A \ge 1$, local, and measurable, and that B is quasilinear with constant $\gamma_B \ge 1$, local, predictable, and monotone. Let $\psi \in \mathcal{D}$ and suppose that*

$$c(f) := \sup\{\|(Af^{k\wedge\tau})_{k=0}^n\|_{\mathrm{BMO}_\psi} - \|B(f^\tau)\|_\infty\} \vee 0,$$

with the supremum taken over all stopping times τ, is finite for all $f \in E$. Then there is some $d > 0$, depending on $\psi \in \mathcal{D}$ and $\gamma_A, \gamma_B \ge 1$ only, such that for $1 \le p < \infty$, $t \ge 1$, $w_t := w_t^{\overline{\psi}}$, and $f \in E$,

$$K_{\infty,p}^{w_t}[A^*f, t^{1/p}] \le d\overline{\psi}^{-1}(p)(K_{\infty,p}[Bf, t^{1/p}] + c(f)).$$

In particular, for $t = 1$ one has $\|A^*f\|_p \le d\overline{\psi}^{-1}(p)(\|Bf\|_p + c(f))$.

DEFINITION 7.5. Let $T : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, $t \ge 0$, $f \in E$, and $1 \le p, q \le \infty$. Then

$$K(f, t, H_q^T, H_p^T) := \inf\{\|Tg\|_q + t\|Th\|_p \mid g, h \in E,\ f = g + h\},$$
$$K(f, t; \mathrm{BMO}_\psi^T, H_p^T) := \inf\{\|(Tg^k)_{k=0}^n\|_{\mathrm{BMO}_\psi} + t\|Th\|_p \mid g, h \in E,\ f = g + h\}.$$

PROPOSITION 7.6. *Let $T : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume T to be quasilinear with constant $\gamma_T \ge 1$, local, predictable, and monotone. Then for all $f \in E$ and $1 \le t, p < \infty$ one has*

$$\frac{1}{3\gamma_T^3}K(f, t, H_\infty^T, H_p^T) \le K_{\infty,p}[Tf, t] \le \gamma_T K(f, t, H_\infty^T, H_p^T).$$

Proof. On the one hand, $Tf \le \gamma_T[Tg + Th]$ a.s. gives $K_{\infty,p}[Tf, t] \le \gamma_T K(f, t, H_\infty^T, H_p^T)$. On the other hand, for $\lambda := \widetilde{Tf}(1/t^p)$ Lemma 7.2 says

$$\|T(f^\tau)\|_\infty \le \gamma_T^2\widetilde{Tf}\left(\frac{1}{t^p}\right) \le \gamma_T^2(Tf)_p^{**}\left(\frac{1}{t^p}\right)$$

and

$$t\|T({}^\tau f)\|_p \le 2\gamma_T^3(Tf)_p^{**}\left(\frac{1}{t^p}\right)$$

where we use a Hardy–Littlewood inequality for the latter relation (see [4] (Theorem 2.2.2)). The rest follows from (18). ∎

The above proposition remains true (with an appropriate change of the multiplicative constants and the same proof) in the quasinormed case $0 < p < 1$.

COROLLARY 7.7. *Let $A, B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$, where E is closed under starting and stopping. Assume that A is quasilinear with constant $\gamma_A \ge 1$, local, and measurable, and that B is quasilinear with constant $\gamma_B \ge 1$, local, predictable, and monotone. Let $\psi \in \mathcal{D}$ and suppose*

$$\|(Af^k)_{k=0}^n\|_{\mathrm{BMO}_\psi} \le \|Bf\|_\infty \quad \text{for all } f \in E.$$

Then there is some $d > 0$, depending on ψ, γ_A, and γ_B only, such that for $1 \le p < \infty$, $t \ge 1$, and $f \in E$,

$$K(f, t; \mathrm{BMO}_\psi^A, H_p^A) \le d\overline{\psi}^{-1}(p)K_{\infty,p}[Bf, t].$$

Proof. Let $f \in E$, $1 \le p < \infty$, and $t \ge 1$. According to Proposition 7.6 we find $g, h \in E$ satisfying $f = g + h$ and $\|Bg\|_\infty + t\|Bh\|_p \le 4\gamma_B^3 K_{\infty,p}[Bf, t]$. Applying Theorem 1.7 to $h \in E$ yields

$$\|(Ag^k)_{k=0}^n\|_{\mathrm{BMO}_\psi} + t\|Ah\|_p \le \|Bg\|_\infty + td\overline{\psi}^{-1}(p)\|Bh\|_p$$
$$\le (d+1)\overline{\psi}^{-1}(p)4\gamma_B^3 K_{\infty,p}[Bf, t]. \ \blacksquare$$

The second example. We show in Proposition 7.11 that the situation of Theorem 1.1 can be reduced to the classical one used in Proposition 7.3 (with $C = 0$). For this purpose we replace a measurable operator, which is majorized by a predictable sequence and defined on random variables with values in a Banach space X, by a predictable operator defined on random variables with values in the Banach space $X \oplus_\infty \mathbb{R}$. This is done with the help of

DEFINITION 7.8. Let $B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$ and $0 < p \le 1$. Then $B_p : E \to L_0^+(\Omega, \mathcal{F}, \mathbb{P})$ is defined for $f = (d_k)_{k=0}^n$ by

$$B_pf(\omega) := \inf\left\{\left(\sum_{l=1}^L [B({}^{n_{l-1}}f^{n_l})(\omega)]^p\right)^{1/p} \;\middle|\; \begin{array}{c} 0 = n_0 < n_1 < \ldots < n_L = n \\ L = 1, 2, \ldots \end{array}\right\}.$$

To summarize the relevant properties of B_p in Lemma 7.10 we need

LEMMA 7.9. *Let M be a non-empty finite set and let $f : 2^M \to [0, \infty)$ be a function satisfying $f(\emptyset) = 0$ and*

$$f(C_1 \cup C_2) \le \gamma[f(C_1) + f(C_2)] \quad \textit{for all disjoint } C_1, C_2 \in 2^M,$$

where $\gamma \ge 1$. *Defining* $0 < p \le 1$ *via* $2^{1/p-1} = \gamma$ *one has*

$$\Big(f\Big(\bigcup_{l=1}^{L} C_l\Big)\Big)^p \le 2\sum_{l=1}^{L}(f(C_l))^p \quad \text{for all pairwise disjoint } C_1,\ldots,C_L \in 2^M.$$

Proof. We will use the arguments given for quasinormed linear spaces in [21] (6.2).

(1) First we show that $(f(C_l))^p \le 2^{-k_l}s$ for $l = 1,\ldots,L$, $k_l \in \{0,1,\ldots\}$, and $\sum_{l=1}^{L} 2^{-k_l} \le 1$ imply $(f(\bigcup_{l=1}^{L} C_l))^p \le s$ whenever the C_l are pairwise disjoint. For $s = 0$ there is nothing to prove, hence we can consider $s > 0$. Furthermore, by adding the empty set we can assume that $\sum_{l=1}^{L} 2^{-k_l} = 1$. We proceed by induction. If $\max\{k_1,\ldots,k_L\} = 0$, then $L = 1$ and we are done. Now let $\max\{k_1,\ldots,k_L\} = h+1$ with $h \ge 0$. If $I := \{l \mid k_l = h+1\}$ we obtain a set of even cardinality since $\sum_{l=1}^{L} 2^{-k_l} = 1$. If $i,j \in I$ then one gets

$$(f(C_i \cup C_j))^p \le \gamma^p[f(C_i) + f(C_j)]^p \le 2^{-h}s$$

so that we can write $\bigcup_{l=1}^{L} C_l$ as disjoint union of D_n $(n = 1,\ldots,N)$ with $(f(D_n))^p \le 2^{-m_n}s$ where $\max\{m_1,\ldots,m_N\} \le h$ and $\sum_{n=1}^{N} 2^{-m_n} = 1$. Hence we have reduced the situation for $h+1$ to the case h and we are done by induction over h.

(2) To prove the statement of our lemma we can assume that $f(C) > 0$ for all $C \ne \emptyset$ (otherwise we consider $f_\varepsilon(C) := f(C) + \varepsilon$ if $C \ne \emptyset$ and $f_\varepsilon(\emptyset) = 0$ with $\varepsilon \downarrow 0$). Let $C = \bigcup_{l=1}^{L} C_l$ be a disjoint union of non-empty sets. If we set $s = 2\sum_{l=1}^{L}(f(C_l))^p$ and choose $k_l \in \{1,2,\ldots\}$ with $2^{-k_l-1}s \le (f(C_l))^p \le 2^{-k_l}s$, then we get $\sum_{l=1}^{L} 2^{-k_l} \le 1$ and $(f(\bigcup_{l=1}^{L} C_l))^p \le s$ from step (1), which proves our assertion. ■

LEMMA 7.10. *Let* $B : \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n) \supseteq E \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$ *be a quasilinear operator with constant* $\gamma_B \ge 1$. *Then for* $0 < p \le 1$ *with* $2^{1/p-1} = \gamma_B$ *one has*:

(1) $B_p f \le Bf \le 2^{1/p}B_p f$ *and* $B_p({}^{k-1}f^k) = B({}^{k-1}f^k)$ *for all* $f \in E$ *and* $1 \le k \le n$.

(2) $[B_p(f)]^p \le [B_p(f^k)]^p + [B_p(f - f^k)]^p$ *for all* $f \in E$ *and* $0 \le k \le n$.

(3) *If* B *is local* (*measurable*, *predictable*, *monotone*) *then so is* B_p.

Proof. Most of the things are evident (for example, $B_p f \le Bf$ follows simply from $d_0 = 0$ for $(d_k)_{k=0}^n \in E$). The point is $Bf \le 2^{1/p}B_p f$, which we get from Lemma 7.9. ■

PROPOSITION 7.11. *Let* $f = (d_k)_{k=0}^n \in \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n)$ *with* $d_0 = 0$. *Assume that*

$$A, B : \{\pm{}^\sigma f^\tau \mid \sigma,\tau \text{ stopping times}\} =: E \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$$

are quasilinear with constants $\gamma_A, \gamma_B \ge 1$, *local, and measurable, and assume that* B *is monotone. Suppose that* $B\,{}^{k-1}f^k \le v_k$ *a.s. for* $k = 1,\ldots,n$, *where* $(v_k)_{k=1}^n \in \mathcal{A}((\mathcal{F}_k)_{k=0}^n)$ *is a fixed predictable sequence and* $v_0 = 0$. *If* $F \subseteq \mathcal{A}^{X\oplus_\infty\mathbb{R}}((\mathcal{F}_k)_{k=0}^n)$ *and* $S,T : F \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$ *are defined by*

$$F := \{((g_k,w_k))_{k=0}^n \mid ((g_k,w_k))_{k=0}^n = \pm(\chi_{\{\sigma<k\le\tau\}}(d_k,v_k))_{k=0}^n, \quad \sigma,\tau \text{ stopping times}\},$$

$$S(((g_k,w_k))_{k=0}^n) := A((g_k)_{k=0}^n),$$

$$T(((g_k,w_k))_{k=0}^n) := (\sup_{0\le k\le n}|w_k|) \vee B((g_k)_{k=0}^n),$$

then (F,S,T) *has* (EP) *with some constant* $c > 0$ *depending on* γ_A *and* γ_B *only*.

Proof. (1) Since B is local, for $k = 1,\ldots,n$ one gets on $\{\sigma < k \le \tau\}$ a.s.

$$B({}^{(k-1)\vee\sigma}f^{\tau\wedge k}) \le \gamma_B[B({}^{\tau\vee(k-1)\vee\sigma}f^k) + B({}^{(k-1)\vee\sigma}f^k)] = \gamma_B B({}^{(k-1)\vee\sigma}f^k)$$
$$\le \gamma_B^2[B({}^{k-1}f^{k\wedge\sigma}) + B({}^{k-1}f^k)] = \gamma_B^2 B({}^{k-1}f^k)$$

so that $B({}^{(k-1)\vee\sigma}f^{\tau\wedge k}) \le \gamma_B^2 v_k\chi_{\{\sigma<k\le\tau\}}$ a.s. on Ω.

(2) We define $U : F \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$ by

$$U(((g_k,w_k))_{k=0}^n) := \sup_{0\le k\le n}\{[B_p((g_l\chi_{\{l\le k\}})_{l=0}^n)]^p + \gamma_B^{2p}|w_{k+1}|^p\}^{1/p},$$

where $w_{n+1} \equiv 0$, $0 < p \le 1$ is chosen so that $2^{1/p-1} = \gamma_B$, and B_p is explained in Definition 7.8. It is clear that S and U are quasilinear and local, and that

$$2^{-1/p}T(((g_k,w_k))_{k=0}^n) \le U(((g_k,w_k))_{k=0}^n) \le \gamma_B^2 2^{1/p}T(((g_k,w_k))_{k=0}^n). \tag{31}$$

S is measurable and U is monotone. The point is that U is predictable since for $1 \le k \le n$,

$$B_p^p({}^\sigma f^{\tau\wedge k}) \le B_p^p({}^\sigma f^{\tau\wedge(k-1)}) + B_p^p({}^{\sigma\vee(k-1)}f^{\tau\wedge k})$$
$$\le B_p^p({}^\sigma f^{\tau\wedge(k-1)}) + \gamma_B^{2p}v_k^p\chi_{\{\sigma<k\le\tau\}}$$

by step (1). Proposition 7.3 ($C = 0$) implies that (F,S,U) has (EP). By (31) the triple (F,S,T) has (EP). ■

Combining Proposition 7.11 with Theorem 1.7 yields an extension of Theorem 1.1.

COROLLARY 7.12. *Let* $\psi \in D$, $f = (d_k)_{k=0}^n \in \mathcal{A}^X((\mathcal{F}_k)_{k=0}^n)$ *with* $d_0 = 0$, *and*

$$A, B : \{\pm{}^\sigma f^\tau \mid \sigma,\tau \text{ stopping times}\} =: E \to L_0^+(\Omega,\mathcal{F},\mathbb{P})$$

be quasilinear operators with constants $\gamma_A, \gamma_B \ge 1$, *local, and measurable, and let* B *be monotone. Suppose that* $B\,{}^{k-1}f^k \le v_k$ *a.s. for* $k = 1,2,\ldots,$

where $(v_k)_{k=1}^n \in \mathcal{A}((\mathcal{F}_k)_{k=0}^n)$ *is a fixed predictable sequence. Assume that*

$$\|(A({}^\sigma f^{\tau\wedge k}))_{k=0}^n\|_{\mathrm{BMO}_\psi} \le \| \sup_{\sigma<k\le\tau} v_k \vee B({}^\sigma f^\tau)\|_\infty$$

for all stopping times σ *and* τ. *Then there is some* $d > 0$, *depending on* $\psi \in \mathcal{D}$, γ_A, *and* γ_B *only, such that for* $1 \le p < \infty$, $t \ge 1$, *and* $w_t := w_t^{\overline{\psi}}$,

$$K_{\infty,p}^{w_t}[A^* f, t^{1/p}] \le d\overline{\psi}^{-1}(p) K_{\infty,p}[\sup_k v_k \vee Bf, t^{1/p}]. \tag{32}$$

In particular, for $t = 1$ *one has* $\|A^* f\|_p \le d\overline{\psi}^{-1}(p)\|\sup_k v_k \vee Bf\|_p$.

8. Extrapolation and self-similar operators. In this last section we will discuss in Proposition 8.2 a possibility to start in Theorem 1.7 and Corollary 6.3 with much weaker assumptions. For simplicity we will restrict ourselves to the case where $\Omega := \mathbb{D}_n = \{(\varepsilon_1,\ldots,\varepsilon_n) \mid \varepsilon_1,\ldots,\varepsilon_n = \pm 1\}$ is the Cantor group equipped with the normalized Haar measure μ_n. As filtration $(\mathcal{F}_k)_{k=0}^n$ we use $\mathcal{F}_0 := \{\emptyset, \mathbb{D}_n\}$, and for $1 \le k \le n$,

$$\mathcal{F}_k := \sigma(\{(\varepsilon_1,\ldots,\varepsilon_n) \mid \varepsilon_{k+1} = \pm 1,\ldots,\varepsilon_n = \pm 1\} \mid \varepsilon_1 = \pm 1,\ldots,\varepsilon_k = \pm 1).$$

As subset E of adapted sequences we take the set $\mathcal{M}_n^X$ of all martingale difference sequences $(d_k)_{k=0}^n \subset L_1^X(\mathbb{D}_n)$ such that $d_0 = 0$. Let us start with the main definition of this section.

DEFINITION 8.1. An operator $A : \mathcal{M}_n^X \to L_0^+(\mathbb{D}_n)$ is called *self-similar* provided that for all $0 < k < n$, all atoms $D \in \mathcal{F}_k$, and all $f = (d_l)_{l=0}^n \in \mathcal{M}_n^X$ with $\mathrm{supp}(d_l) \subseteq D$ for all $l = 0,\ldots,n$ and $d_0 = \ldots = d_k = 0$ one has

$$A(f^{(\theta_1,\ldots,\theta_k)}) = (Af)^{(\theta_1,\ldots,\theta_k)} \quad (\theta_1,\ldots,\theta_k \in \{-1,1)),$$

with $f^{(\theta_1,\ldots,\theta_k)} := (d_l^{(\theta_1,\ldots,\theta_k)})_{l=0}^n$, where

$$h^{(\theta_1,\ldots,\theta_k)}(\varepsilon_1,\ldots,\varepsilon_n) := h(\theta_1\varepsilon_1,\ldots,\theta_k\varepsilon_k,\varepsilon_{k+1},\ldots,\varepsilon_n)$$

for some $h \in L_1^X(\mathbb{D}_n)$.

Basic examples of self-similar operators are operators generated by UMD-transforms or generalized square functions, that is, $(Af)(\omega) := \|\sum_{k=1}^n \alpha_k d_k(\omega)\|_X$ or $(Af)(\omega) := (\sum_{k=1}^n \|d_k(\omega)\|_X^p)^{1/p}$ where $f = (d_k)_{k=0}^n \in \mathcal{M}_n^X$, $(\alpha_k)_{k=1}^n \subset \mathbb{R}$, and $0 < p \le \infty$.

PROPOSITION 8.2. *Let* $A, B : \mathcal{M}_n^X \to L_0^+(\mathbb{D}_n)$ *be self-similar, sublinear, and local, where* A *is assumed to be measurable and* B *is assumed to be monotone.*

(1) *If* $\varphi \in C_\Delta$, *then* $|Af|_{M_\varphi^0(\mathbb{D}_n)} \le \|Bf\|_\infty$ *for all* $f \in \mathcal{M}_n^X$ *implies*

$$\|(Af^k)_{k=0}^n\|_{\mathrm{BMO}_\psi} \le 2\Delta(\varphi)\|Bf\|_\infty \quad \textit{for } f \in \mathcal{M}_n^X.$$

(2) *If* $0 < s < 1/4$, *then* $\mu_n(Af > \|Bf\|_\infty) \le s$ *for all* $f \in \mathcal{M}_n^X$ *implies*

$$\|(Af^k)_{k=0}^n\|_{\mathrm{BMO}_{\psi_1}^*} \le 6\max\left(1, \frac{1}{\log(1/(4s))}\right)\|Bf\|_\infty \quad \textit{for } f \in \mathcal{M}_n^X.$$

Proof. (1) Let $0 < k < n$ and assume $g = (dg_l\chi_C)_{l=0}^n \in \mathcal{M}_n^X$ and $h = (dh_l\chi_D)_{l=0}^n \in \mathcal{M}_n^X$ where $C, D \in \mathcal{F}_k$ are disjoint and $dg_0 = \ldots = dg_k = dh_0 = \ldots = dh_k = 0$. If the operator $T : \mathcal{M}_n^X \to L_0^+(\mathbb{D}_n)$ is sublinear and local, then one easily shows that $T(g+h) = Tg + Th$. If we additionally assume that T is self-similar, then for $0 < k < n$, an atom $D \in \mathcal{F}_k$, and $f = (d_l\chi_D)_{l=0}^n \in \mathcal{M}_n^X$ with $d_0 = \ldots = d_k = 0$ we get by induction

$$T\Big(\sum_{\theta_1,\ldots,\theta_k=\pm1} f^{(\theta_1,\ldots,\theta_k)}\Big) = \sum_{\theta_1,\ldots,\theta_k=\pm1} T(f^{(\theta_1,\ldots,\theta_k)}) = \sum_{\theta_1,\ldots,\theta_k=\pm1} (Tf)^{(\theta_1,\ldots,\theta_k)}.$$

(2) Let $f = (d_l)_{l=0}^n \in \mathcal{M}_n^X$, $1 < k \le l \le n$, and $D \in \mathcal{F}_k$ be an atom. Let $\widetilde{D} \supset D$ be the dyadic predecessor, $\widetilde{D} \in \mathcal{F}_{k-1}$. It follows for $\lambda \ge 0$ and

$$g := \sum_{\theta_1,\ldots,\theta_{k-1}=\pm1} ((f^l - f^{k-1})\chi_{\widetilde{D}})^{(\theta_1,\ldots,\theta_{k-1})}$$

that

$$\begin{aligned}(\mu_n)_D(|Af^l - Af^{k-1}| > \lambda) &\le (\mu_n)_D(A(f^l - f^{k-1}) > \lambda)\\ &\le 2(\mu_n)_{\widetilde{D}}(A(f^l - f^{k-1}) > \lambda)\\ &= 2(\mu_n)_{\widetilde{D}}(A((f^l - f^{k-1})\chi_{\widetilde{D}}) > \lambda)\\ &= 2\mu_n(Ag > \lambda).\end{aligned}$$

(3) To prove the first assertion of our proposition we derive, under the assumptions of step (2),

$$\begin{aligned}|Af^l - Af^{k-1}|_{M_\varphi^0(D,(\mu_n)_D)} &\le \Delta(\varphi)|Ag|_{M_\varphi^0(\mathbb{D}_n)} \le \Delta(\varphi)\|Bg\|_\infty\\ &= \Delta(\varphi)\|B((f^l - f^{k-1})\chi_{\widetilde{D}})\|_{L_\infty(\widetilde{D})}\\ &\le 2\Delta(\varphi)\|Bf\|_{L_\infty(\mathbb{D}_n)}.\end{aligned}$$

The case $k = 0, 1$ leads trivially to the same estimate. Finally, letting $D = \bigcup_i D_i$ be a disjoint union of atoms $D_i \in \mathcal{F}_k$ $(k \ge 1)$ one gets

$$\begin{aligned}|Af^l - Af^{k-1}|_{M_\varphi^0(D,(\mu_n)_D)} &\le \sup_i |Af^l - Af^{k-1}|_{M_\varphi^0(D_i,(\mu_n)_{D_i})}\\ &\le 2\Delta(\varphi)\|Bf\|_\infty.\end{aligned}$$

(4) To prove the second assertion we consider the estimate $2\|Bf\|_\infty \ge \|B((f^l - f^{k-1})\chi_{\widetilde{D}})\|_{L_\infty(\widetilde{D})} = \|Bg\|_\infty$, so that step (2) and our assumption yield

$$(\mu_n)_D(|Af^l - Af^{k-1}| > 2\|Bf\|_\infty) \le 2\mu_n(Ag > \|Bg\|_\infty) \le 2s.$$

Similarly to step (3) we get $(\mu_n)_D(|Af^l - Af^{k-1}| > 2\|Bf\|_\infty) \le 2s$ for all $D \in \mathcal{F}_k$. The case $k = 0, 1$ leads to the same estimate and we can finish with Theorem 4.6(23). ■

If we now combine the second assertion of the proposition above with Proposition 7.3 ($C = 0$) and Theorem 1.7, then we get

COROLLARY 8.3. *Let $A, B : \mathcal{M}_n^X \to L_0^+(\mathbb{D}_n)$ be self-similar, sublinear, and local, where A is assumed to be measurable and B is assumed to be predictable and monotone. Then for all $0 < s < 1/4$ there is some $d_s > 0$, not depending on n, A, and B, such that*

$$\|A^* f\|_p > d_s p \|Bf\|_p$$

for some $p \in [1, \infty)$ and some $f \in \mathcal{M}_n^X$ implies the existence of some $g \in \mathcal{M}_n^X$ such that

$$\mu_n(Ag > \|Bg\|_\infty) > s.$$

Remark 8.4. (1) D. L. Burkholder has shown in [9] (Lemma 3.1) that for the UMD-transforms $A_0, B_0 : \mathcal{M}_N^X \to L_0^+(\mathbb{D}_n)$ with $A_0((d_k)_{k=0}^n) := \|\sum_{k=1}^n \theta_k d_k\|_X$ and $B_0((d_k)_{k=0}^n) := \|\sum_{k=1}^n d_k\|_X$, where $\theta_k \in \{-1, 1\}$ are fixed, the inequality $A_0^* f > 6c\|B_0 f\|_1$ a.s. for some $f \in \mathcal{M}_n^X$ and some $c > 0$ gives some $g \in \mathcal{M}_n^X$ with

$$\mu_n(A_0^* g > c\|B_0 g\|_\infty) \ge 1/2.$$

To apply our result in this situation we have to replace B_0 by the monotone operator

$$^*B_0 f := \sup_{0 \le l < n} \Big[\Big\| \sum_{i=1}^l d_i \Big\|_X + \|d_{l+1}\|_X \Big] \quad \text{where } f = (d_l)_{l=0}^n.$$

$^*B_0$ is predictable since for $1 \le k \le n$ one has $^*B_0 f^k = \sup_{0 \le l < k} [\|\sum_{i=1}^l d_i\|_X + \|d_{l+1}\|_X]$ and that $\|d_k\|$ is $\mathcal{F}_{k-1}$-measurable. Moreover, the estimate $((p-1)/(3p))\|^*B_0 f\|_p \le \|B_0 f\|_p \le \|^*B_0 f\|_p$ holds by Doob's maximal inequality. Now A_0 and $^*B_0$ satisfy the assumptions of the corollary above and we obtain [9] (Lemma 3.1) if we replace $\|B_0 f\|_1$ by $\|B_0 f\|_p$ for some $p > 1$. To get Burkholder's lemma in the case $p = 1$ in our general situation it would be necessary to weaken the conditions for the operator B in Corollary 8.3 to self-similar, sublinear, local, and measurable.

(2) A further result concerning UMD-transforms related to our corollary can be found in a slightly different setting in [7] (Theorem 1.1).

The next application of the concept of self-similar operators concerns the following martingale-type quantities.

DEFINITION 8.5. Let $1 < p \le 2$, $0 < s < 1$, $1 \le \alpha, \beta \le \infty$, $n = 1, 2, \ldots$, and let $T : X \to Y$ be a continuous linear operator between the Banach spaces X and Y. Then $\mathrm{Mt}_{p,n}(T \mid L_\alpha, L_\beta) := \inf c$ and $\mathrm{Mt}_{p,n}(T \mid L_{0,s}, L_\beta) := \inf d$, respectively, where the infima are taken over all $c, d > 0$ such that for all $(d_k)_{k=0}^n \in \mathcal{M}_n^X$,

$$\Big\| \sum_{k=1}^n T d_k \Big\|_{L_\alpha^Y} \le c \Big\| \Big(\sum_{k=1}^n \|d_k\|_X^p \Big)^{1/p} \Big\|_\beta$$

and

$$\mu_n \Big(\Big\| \sum_{k=1}^n T d_k \Big\|_Y > d \Big\| \Big(\sum_{k=1}^n \|d_k\|_X^p \Big)^{1/p} \Big\|_\beta \Big) \le s.$$

G. Pisier proved in [22] (Sublemma 3.3) that

$$\sup_n \mathrm{Mt}_{p,n}(T \mid L_{p,\infty}, L_p) \le c_0 \sup_n \mathrm{Mt}_{p,n}(T \mid L_2, L_\infty).$$

The following extends this inequality in two directions.

COROLLARY 8.6. *Let $1 < p \le 2$, $0 < s < 1/4$, $1 \le \alpha < \infty$, $n = 1, 2, \ldots$, and let $T : X \to Y$ be a continuous linear operator. Then one has, for some $c > 0$ depending on s and α only,*

$$\mathrm{Mt}_{p,n}(T \mid L_\alpha, L_\alpha) \le c\, \mathrm{Mt}_{p,n}(T \mid L_{0,s}, L_\infty) \le \frac{c}{s} \mathrm{Mt}_{p,n}(T \mid L_1, L_\infty).$$

Proof. To apply Proposition 8.2(2) we use $A, B : \mathcal{M}_n^X \to L_0^+(\mathbb{D}_n)$ with

$$A((d_k)_{k=0}^n)(\omega) := \Big\| T \Big(\sum_{k=1}^n d_k(\omega) \Big) \Big\|_Y$$

and

$$B((d_k)_{k=0}^n)(\omega) := \alpha \Big(\sum_{k=1}^n \|d_k(\omega)\|_X^p \Big)^{1/p},$$

where $\alpha := \mathrm{Mt}_{p,n}(T \mid L_{0,s}, L_\infty)$. Then we finish with Theorem 1.7.

Acknowledgements. I would like to thank S. Kwapień for calling my attention to the papers of P. Hitczenko and M. Junge for the joint discussion concerning Theorem 3.4. Finally, I would like to thank the referee for his careful reading of the manuscript.

References

[1] N. H. Asmar and S. J. Montgomery-Smith, *On the distribution of Sidon series*, Ark. Mat. 31 (1993), 13–26.
[2] N. L. Bassily and J. Mogyoródi, *On the BMO$_\Phi$-spaces with general Young function*, Ann. Univ. Sci. Budapest Eötvös Sect. Math. 27 (1984), 215–227.
[3] J. Bastero and F. J. Ruiz, *Interpolation of operators when the extreme spaces are L^∞*, Studia Math. 104 (1993), 133–150.
[4] C. Bennett and R. Sharpley, *Interpolation of Operators*, Academic Press, 1988.
[5] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer, 1976.
[6] J. Bourgain, *On the behaviour of the constant in the Littlewood–Paley inequality*, in: J. Lindenstrauss and V. D. Milman (eds.), Israel Seminar, GAFA 1987/88, Lecture Notes in Math. 1376, Springer, 1989, 202–208.

[7] J. Bourgain and W. J. Davis, *Martingale transforms and complex uniform convexity*, Trans. Amer. Math. Soc. 294 (1986), 501–515.
[8] D. L. Burkholder, *Distribution function inequalities for martingales*, Ann. Probab. 1 (1973), 19–42.
[9] —, *Martingale transforms and the geometry of Banach spaces*, in: Probability in Banach spaces III, 1980, Lecture Notes in Math. 860, Springer, 1981, 35–50.
[10] D. L. Burkholder, B. J. Davis, and R. F. Gundy, *Integral inequalities for convex functions of operators on martingales*, Proc. 6th Berkeley Sympos. Math. Statist. Probab., Vol. 2, Univ. of California Press, 1972, 223–240.
[11] D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasilinear operators on martingales*, Acta Math. 124 (1970), 249–304.
[12] S. Chang, M. Wilson and J. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helv. 60 (1985), 217–246.
[13] V. H. de la Peña, S. J. Montgomery-Smith, and J. Szulga, *Contraction and decoupling inequalities for multilinear forms and U-statistics*, Ann. Probab. 22 (1994), 1745–1765.
[14] A. M. Garsia, *Martingale Inequalities, Seminar Notes on Recent Progress*, Benjamin, Reading, 1973.
[15] E. Giné and J. Zinn, *Central limit theorems and weak laws of large numbers in certain Banach spaces*, Z. Wahrsch. Verw. Gebiete 62 (1983), 323–354.
[16] P. Hitczenko, *Upper bounds for the L_p-norms of martingales*, Probab. Theory Related Fields 86 (1990), 225–238.
[17] F. John, *Quasi-isometric mappings*, in: Seminari 1962–1963 di Analisi, Algebra, Geometria e Topologia, Vol. II, Ed. Cremonese, Rome, 1965, 462–473.
[18] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. 14 (1961), 415–426.
[19] W. B. Johnson and G. Schechtman, *Sums of independent random variables in rearrangement invariant function spaces*, Ann. Probab. 17 (1989), 789–808.
[20] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, 1991.
[21] A. Pietsch, *Operator Ideals*, North-Holland, Amsterdam, 1980.
[22] G. Pisier, *Martingales with values in uniformly convex spaces*, Isreal J. Math. 20 (1975), 326–350.
[23] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, 1993.
[24] J.-O. Strömberg, *Bounded mean oscillation with Orlicz norms and duality of Hardy spaces*, Indiana Univ. Math. J. 28 (1979), 511–544.
[25] G. Wang, *Some sharp inequalities for conditionally symmetric martingales*, PhD thesis, Univ. of Illinois, 1989.
[26] F. Weisz, *Martingale operators and Hardy spaces generated by them*, Studia Math. 114 (1995), 39–70.

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

Received November 7, 1995
Revised version July 15, 1996 (3566)

STUDIA MATHEMATICA 122 (3) (1997)

Cohomology groups, multipliers and factors in ergodic theory

by

M. LEMAŃCZYK (Toruń)

Abstract. The problem of compact factors in ergodic theory and its relationship with the problem of extending a cocycle to a cocycle of a larger action are studied.

Introduction. Given an ergodic automorphism $\tau : (Y, \mathcal{C}, \nu) \to (Y, \mathcal{C}, \nu)$ of a Lebesgue space $(Y, \mathcal{C}, \nu)$ call any of its invariant σ-algebras a *factor*. Denote by

$$C(\tau) = \{S : (Y, \mathcal{C}, \nu) \to (Y, \mathcal{C}, \nu) : S\tau = \tau S,\ S \text{ invertible}\}$$

the *centralizer* of τ. Endowed with the weak topology in which

$$S_n \to S \quad \text{iff} \quad \mu(S_n^{\pm 1}A \bigtriangleup S^{\pm 1}A) \to 0 \quad \text{for each } A \in \mathcal{C},$$

it becomes a Polish group. If $\mathcal{H} \subset C(\tau)$ is a subgroup then it determines a factor $\mathcal{A}(\mathcal{H})$ given by

$$\mathcal{A}(\mathcal{H}) = \{A \in \mathcal{C} : SA = A \text{ for each } S \in \mathcal{H}\}.$$

On the other hand, a factor $\mathcal{A}$ determines a subgroup $\mathcal{H}(\mathcal{A}) \subset C(\tau)$ by

$$\mathcal{H}(\mathcal{A}) = \{S \in C(\tau) : SA = A \text{ for each } A \in \mathcal{A}\}.$$

From this point of view compact subgroups are of special interest as for them

$$\mathcal{H}(\mathcal{A}(\mathcal{H})) = \mathcal{H} \tag{1}$$

(see [5], [17]). Moreover, in this case τ can be represented as a compact group extension T_φ defined on the space $(X \times \mathcal{H}, \widetilde{\mu})$, where X stands for the quotient space corresponding to the factor $\mathcal{A}(\mathcal{H})$, $\widetilde{\mu}$ for the product measure of the corresponding image of ν with Haar measure $m_{\mathcal{H}}$ and T denotes the quotient action of τ; T_φ is defined by

$$T_\varphi(x, S) = (Tx, \varphi(x)S),$$

1991 *Mathematics Subject Classification*: Primary 28D05.

Research partially supported by KBN grant 2 P301 031 07 (1994).