References


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

Received October 25, 1995
Revised version June 18, 1996 and September 18, 1996

BMO_p-spaces and applications to extrapolation theory

by

STEFAN GEISS (Jena)

Abstract. We investigate a scale of BMO_p-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_0-Lorentz-estimates, and arrives at L_p-Lorentz-estimates, or more generally, at estimates between K-functional 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 \( \mathcal{F}_k \) 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. Hitzlpenko [16] (Lemma 4.3) (see also [12], [25], the comments in [16], and the proof of Proposition 1 of this paper) says that for \( \lambda > 0 \),

\[
\mathbb{P}\left( \left| \sum_{k=1}^n d_k \right| > \lambda \| S_2 f \|_2 \right) \leq 2e^{-\lambda^2/2}
\]

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

\[
\left\| \sup_{1 \leq k \leq n} \left| \sum_{i=1}^k d_i \right| \right\|_p \leq c\sqrt{p} \left\| \left( \sum_{i=1}^n |d_i|^2 \right)^{1/2} \right\|_p \quad \text{for } 1 \leq p < \infty,
\]

which is 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 \( (d_i)_{i=k+1}^n \subset L_1(B, \mathbb{P}_B) \), where \( \mathbb{P}_B \) is the normalized restriction of \( \mathbb{P} \) to \( B \). Applying (1) to

1991 Mathematics Subject Classification: 46B70, 46B30, 47H99, 60G99.
this restricted sequence and taking into account \( \| S_2(d_k) \|_{\ell^k} \leq \| S_2 h \|_{\ell^2} \) we arrive at the apparently stronger estimate

\[
(3) \quad \sup_{0 \leq s \leq t \leq \lambda} \sup_{\beta \in \mathbb{R}^k} \mathbb{P}_B \left( \left| \sum_{i=0}^k d_i \right| > \lambda \| S_2 h \|_{\ell^2} \right) \leq 2e^{-\lambda^2/2}.
\]

Secondly, analyzing [16] more carefully, we see that we additionally have to replace the terms \( \sum_{i=k+1}^k d_i \) by \( \sum_{i=1}^k d_i \), 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

\[
(4) \quad \left\| \left( \sum_{i=0}^k d_i \right)^n \right\|_{\text{BMO}_\psi_{2\lambda}} \leq c_0 \| S_2 h \|_{\ell^2}
\]

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 \( L_\infty \)-norm of \( S_2 h \). It is obviously easier to check \( \| S_2 h \|_{\ell^2} \) 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 \( e_k, \ldots, e_n \) such that \( \mathbb{P}(e_k = 1) = \mathbb{P}(e_k = -1) = \frac{1}{2} \) and define

\[ d_k := \frac{e_k}{2} \quad (k = 2, \ldots, n), \]

then \( \| S_2 h \|_{\ell^2} = \sqrt{n-1} \) and for \( \lambda > 0 \) inequality (1) yields

\[
\mathbb{P}\left( \left| \sum_{k=2}^n d_k \right| > \lambda \right) \leq 2e^{-\lambda^2/(2(n-1))}
\]

whereas

\[
(5) \quad \frac{1}{16} e^{-\lambda} \leq \mathbb{P}\left( \left| \sum_{k=2}^n d_k \right| > \lambda \right) \leq \frac{1}{4} e^{-\lambda (\log 2) \sqrt{\lambda}}
\]

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

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, \( x \in X_0 \times X_1 \), and \( t \geq 0 \) the \( K \)-functional is defined by

\[
K(x, t; X_0, X_1) := \inf \left\{ \| x_0 \|_{X_0} + t \| x_1 \|_{X_1} \mid x = x_0 + x_1, \ x_0 \in X_0, \ x_1 \in X_1 \right\}
\]

In the case \( X_0 = L_0(\Omega) \) and \( X_1 = L_\infty(\Omega) \) we will shorten \( K_{\ref{K}}(f, t) := K(f, t; L_0, L_\infty) \). 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 \( c \geq 1 \) such that for all martingale difference sequences \( f = (d_k)_{k=0}^\infty \subset L_1(\Omega, \mathcal{F}, \mathbb{P}) \) such that \( d_0 = 0 \) and \( \langle d_k \rangle \) is \( \mathcal{F}_{k-1} \)-measurable for \( k = 1, \ldots, n \) one has for \( t \geq 1 \),

\[
K_{\ref{K}}(f, t; L_0, L_\infty) \leq c \sqrt{1 + \log t} \sqrt{K_{\ref{K}}(S_2 f, \sqrt{t})}
\]

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

\[
\mathbb{P}\left( \left| f_n \right| > c \lambda \sqrt{K_{\ref{K}}(S_2 f, \sqrt{t})} \right) \leq e^{-\lambda^2}.
\]

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

\[
eq \mu d_k(\omega) \leq d_k(\omega) + \frac{d_k(\omega)}{|d_k(\omega)|} \sinh(\mu|d_k(\omega)|)
\]

so that

\[
eq d_k(\omega) e^{\mu d_k(\omega)} \leq e^{\mu e d_k(\omega)^2}.
\]

Iteration with respect to \( k \) gives \( E e^{\mu e d_k(\omega)^2} \leq 1 \). Since this inequality remains true for \( -f_n \) instead of \( f_n \) we can continue to get \( E e^{\mu e} \leq 3 \) where

\[
g(\omega) := (\mu | f_n(\omega) | - (\mu^2/2)(S_2 f)^2) > 0, \quad \text{and by standard arguments we obtain}
\]

\[
K_{\ref{K}}(f_n, t) \leq c(1 + \log t) \quad \text{for} \quad t \geq 1.
\]

If \( K_{\ref{K}}(S_2 f, t) > 0 \) and \( \mu^2 := 2c(1 + \log t) / K_{\ref{K}}(S_2 f, t) \), then

\[
K_{\ref{K}}(f_n, t) \leq 2(1 + \log t) K_{\ref{K}}(S_2 f, t) \leq 2(1 + \log t) K_{\ref{K}}(S_2 f, \sqrt{t})
\]

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

For the example considered in (5) one has \( K_{\ref{K}}(S_2 f, e^{(\lambda^2/2)^2}) \leq e c_1 \lambda \) for \( \lambda \geq 1 \) and some \( c_1 \geq 1 \). Hence the above proposition and the lower estimate of (5) imply for \( 2 \sqrt{(c c_0)^2} \leq \lambda \leq n-4 \),

\[
\frac{1}{16} e^{-\lambda} \leq \mathbb{P}\left( \left| \sum_{k=2}^n d_k \right| > \lambda \right) \leq \frac{1}{4} e^{-\lambda (\log 2) \sqrt{\lambda}}
\]

and

\[
eq c \sqrt{K_{\ref{K}}(S_2 f, e^{(\lambda^2/2)^2})} \leq e c_1 \lambda \sqrt{K_{\ref{K}}(S_2 f, e^{(\lambda^2/2)^2})}
\]

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
\[
K_{\infty,p} [\sup_{k \in \mathbb{Z}} |f_k|, t^{1/p}] \leq c \sqrt{p} \sqrt{1 + \log t} K_{\infty,p} [S_{f_k}^p, t^{1/p}] \quad (1 \leq p < \infty),
\]
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. Hitzcnko recalled in Theorem 1.1 of this paper. Reformulating the assumptions of this statement we introduce a scale of \( \text{BMO}_\psi \)-spaces. A separate investigation of these spaces is one of the main subjects of this paper and leads to further development of Hitzcnko’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 \( \sigma \)-algebra \( \mathcal{B}(X) \), then \( L^0_0(\Omega, \mathcal{F}, \mathbb{P}) \) is the linear space of all measurable \( f : [\Omega, \mathcal{F}, \mathbb{P}] \to [X, B(X)] \) which are Radon (cf. [20]). Moreover, \( L^0_0(\Omega, \mathcal{F}, \mathbb{P}) := \{ f \in L_0^0(\Omega, \mathcal{F}, \mathbb{P}) \mid f \geq 0 \text{ a.s.} \} \) and
\[
L^0_0(\Omega, \mathcal{F}, \mathbb{P}) := \{ f \in L_0^0(\Omega, \mathcal{F}, \mathbb{P}) \mid \| f \|_{L^0_0} := \| f(w) \|_{L^0_0(\mathbb{P})} < \infty \}
\]
\((1 \leq p \leq \infty)\). If \([\Omega, \mathcal{F}, \mathbb{P}]\) is equipped with a filtration \( (\mathcal{F}_k)_{k \in \mathbb{N}} \), where \( I = \{0, \ldots, n\} \) or \( I = \mathbb{N} = \{0, 1, 2, \ldots\} \), and \( \mathcal{V} \in \mathcal{F}_k \), then we use the linear spaces (under the coordinatewise operations) of adopted sequences
\[
A^I(\mathcal{F}_k)_{k \in \mathbb{N}} := \{ (d_k)_{k \in \mathbb{N}} \mid d_k \in L_0^0(\Omega, \mathcal{F}_k, \mathbb{P}) \text{ for all } k \in I \}
\]
and \([A(\mathcal{F}_k)_{k \in \mathbb{N}}] := A^\mathbb{N}(\mathcal{F}_k)_{k \in \mathbb{N}} \). For stopping times \( \sigma, \tau \) and \( f = (d_k)_{k \in \mathbb{N}} \in A^I(\mathcal{F}_k)_{k \in \mathbb{N}} \) one usually sets
\[
\sigma^f := (d_k \chi_{\sigma < k \leq \tau})_{k \in \mathbb{N}}, \quad \tau^f := (d_k \chi_{k < \tau})_{k \in \mathbb{N}}, \quad \sigma^f := (d_k \chi_{\sigma < k \leq \tau})_{k \in \mathbb{N}}.
\]
We will consider operators
\[
T : A^I(\mathcal{F}_k)_{k \in \mathbb{N}} \ni E \to L_0^0(\Omega, \mathcal{F}, \mathbb{P}),
\]
where we make the following conventions valid in the whole paper:

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

The operator \( T \) is called quasilinear if there is some \( \gamma > 0 \) such that for all \( f, g \in E \) with \( f + g \in E \) one has \( T(f + g) \leq \gamma T(f + T(g)) \) a.s. If \( \gamma = 1 \), then \( T \) is called sublinear. The operator \( T \) is monotone if \( T f^k \leq T f \) a.s., and measurable if \( T f^k \) is \( \mathcal{F}_k \)-measurable for all \( f \in E \) and \( k \in I \). The operator \( T \) is called local if for all \( f = (d_k)_{k \in \mathbb{N}} \in E \) one has \( T f = 0 \) a.s. on the set
\[
\{ 0 = \mathbb{E}(\|d_k\|_F) = \ldots = \mathbb{E}(\|d_k\|_F^k) = \ldots \}.
\]
The maximal operator \( T^* \) of \( T \) is given by \( T^* f := \sup_{k \in \mathbb{N}} T f^k \). Finally, we will denote the set of all increasing bijections \( \phi : [1, \infty) \to [1, \infty) \) by \( \mathcal{D} \).

Now let us describe (in a slightly different form) the result of P. Hitzcnko 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** (Hitzcnko [16]). Let \( \psi \in \mathcal{D} \) and \( E \subset A(\mathcal{F}_k)_{k \in \mathbb{N}} \) be the subset of the martingale difference sequences with \( d_0 = 0 \). Assume \( T : E \to L^0_0(\Omega, \mathcal{F}, \mathbb{P}) \) 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 \( \sigma, \tau \), and all \( \lambda \geq 1 \) one has
\[
\mathbb{P} \left( \left| \sum_{k=0}^n d_k \right| > \lambda \right| \sup_{\sigma < k < \tau} \mathbb{E}(\psi(v_k \vee T^{(k-1)f^k})) \right) \subset B \right) \leq c^1 - \psi^1(\lambda) \mathbb{P}(B),
\]
then one gets for \( 1 \leq p < \infty \) and some constant \( c > 0 \), depending on \( T \) and \( \psi \) only,
\[
\left( \sup_n \left| \sum_{k=0}^n d_k \right|_p \leq c^1 - \psi^1(p) \sup_{k \in \mathbb{N}} \mathbb{E}(v_k \vee T f) \right)
\]
As in (3) \( \Rightarrow \) (4) of the introduction we translate (8) into the language of \( \text{BMO}-\psi \)-spaces. To this end we introduce
\[
\text{BMO}_\psi((\mathcal{F}_k)_{k \in \mathbb{N}}) := \{ (f_k)_{k \in \mathbb{N}} \in A(\mathcal{F}_k)_{k \in \mathbb{N}} \mid \|(f_k)_{k \in \mathbb{N}}\|_{\text{BMO}_\psi(\infty)} < \infty \}
\]
with
\[
\|(f_k)_{k \in \mathbb{N}}\|_{\text{BMO}_\psi} := \inf \{ c > 0 \mid \sup_{0 \leq k \leq \infty} \sup_{k \leq \infty} \mathbb{P} |f_k - f_{k-1}| > \lambda \} \leq c^1 - \psi^1(\lambda/c) \text{ for } \lambda \geq c\}
\]
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 \( d_0 \leq v_0 \) the relation (8) implies
\[
\left( \sum_{k=0}^n d_k \chi_{(\sigma < k \leq \tau)} \right)_{\sigma < k < \tau} \subseteq \text{BMO}_\psi \quad \text{with}
\]
\[
\left( \sum_{k=0}^n d_k \chi_{(\sigma < k \leq \tau)} \right)_{\sigma < k < \tau} \subseteq \text{BMO}_\psi \quad \text{with}
\]
Conversely, (10) implies (8) if we replace \( \sup_{0 \leq k \leq \infty} \sup_{k \leq \infty} \mathbb{P} (|f_k - f_{k-1}| > \lambda) \leq c^1 - \psi^1(\lambda/c) \text{ for } \lambda \geq c\}
\]
Using the properties of \( \text{BMO}_\psi \)-spaces we will extend the extrapolation principle (10) \( \Rightarrow \) (9) in Theorem 1.7, Corollary 6.5, and Corollary 7.12.

The \( \text{BMO}_\psi \)-spaces are introduced and investigated in Sections 4 and 5 whereas the required material about the Lorentz spaces \( \mathcal{M}_\rho \), which are the basic modules of \( \text{BMO}_\psi \)-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$_{\psi}$-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 $\psi$ generate the same BMO$_{\psi}$-spaces. We completely solve this problem by a regularization $\overline{\psi}$ of $\psi$ defined as follows.

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

$$
\overline{\psi}(\mu) := \sup \left\{ \sum_{i=1}^{N} (\psi(\mu_i) - 1) + 1 \mid \mu = \sum_{i=1}^{N} \mu_i, \mu_i \geq 1, N = 1, 2, \ldots \right\}.
$$

It is clear that $\psi \leq \overline{\psi}$. In Lemma 4.4 we see $\overline{\psi} \in D$ (we show the stronger $\overline{\psi} \in C_\Delta$, see (20) and Definition 2.1) and that, for some $c_\psi \geq 1$, one has

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

In particular, $\overline{\psi}(a_0 \mu) \geq \mu$ for $\mu \geq 1$ if we choose $\lambda$ to be $a_0 := \overline{\psi}^{-1}(c_\psi)$, 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

$$
\|f_k\|_{BMO_\psi} \leq \|f_k\|_{BMO_\psi} \leq 6 \psi^{-1}(3)\|f_k\|_{BMO_\psi}.
$$

Moreover, in Section 4 we prove

**Theorem 1.3.** For $\psi, \psi' \in D$ 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 $(\Omega, F, \mathcal{F})$, $(F_k)_{k \in I}$ and all adapted sequences $(f_k)_{k \in I} \in BMO_{\psi'}((F_k)_{k \in I})$, we have

$$
\|f_k\|_{BMO_\psi} \leq c_2 \|f_k\|_{BMO_{\psi'}}.
$$

We have seen that the regularization $\overline{\psi}$ of $\psi$ is the right tool to investigate the BMO$_{\psi}$-spaces themselves. The consequences for the extrapolation (10) $\Rightarrow$ (9) are as follows: Starting with (10) we can switch via (11) from $\psi$ to $\overline{\psi}$. Extrapolation gives (9) with $\overline{\psi}^{-1}(p)$ instead of $\psi^{-1}(p)$. In the case $\psi(\mu) = \mu^p$ $(1 \leq q < \infty)$ this does not yield an improvement (see Example 4.3). But, for instance, if $\lim_{\mu \to \infty} \psi(\mu)/\mu = 0$ then $\overline{\psi}^{-1}(p)$ with $\psi^{-1}(p) \leq a_0 \mu$ (which follows from $\overline{\psi}(a_0 \mu) \geq \mu$) is better than $\psi^{-1}(p)$.

In Section 5 we consider convexity properties of BMO$_{\psi}$-spaces. We modify the upper $p$-estimates, known for Banach lattices, to the following weak upper estimates.

**Definition 1.4.** If $\psi \in D$, then the BMO$_{\psi}$-spaces satisfy a weak upper estimate with respect to the sequence $a = (a_i)_{i \geq 1}$ with $1 \geq a_1 \geq a_2 \geq \ldots \geq 0$ provided that there is a constant $c > 0$ such that for all $(\Omega, F, \mathcal{F})$, $(F_k)_{k \in I}$ and $(f_k)_{k \in I} \in BMO_{\psi}((F_k)_{k \in I})$ and

$$
\left\| \sup_{i \geq 1} a_i f_k^{(i)} \right\|_{BMO_\psi} \leq c \sup_{i \geq 1} \left\| f_k^{(i)} \right\|_{BMO_\psi}.
$$

We set $U_a(BMO_\psi) := \inf c$ where $U_a(BMO_\psi) = \infty$ if such a $c > 0$ does not exist.

We will prove

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

$$
0 < \inf N \sum_{N} U_b^{N}(BMO_\psi) \leq U_a(BMO_\psi) < \infty.
$$

In this way we have shown that $(1/\psi^{-1}(1 + \log i))_{i \geq 1}$ is (up to multiplicative constants) the optimal sequence such that the BMO$_{\psi}$-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

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

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

$$
\frac{1}{\lambda} \chi_{\{|f| \leq \lambda\}} A^* f \leq A^* g \leq c A^* f \quad \text{a.s. and } B g \leq c \lambda \quad \text{a.s.}
$$

Moreover, for every $\psi \in D$ and $1 \leq t < \infty$ we define the weight function

$$
\psi_\psi(s) := \left\{ \begin{array}{ll}
\frac{1}{\psi^{-1}(1 + \log s)}, & 1 < s \leq 1,
1, & 0 < s < 1/t,
\end{array} \right.
$$

so that

$$
\frac{1}{\psi^{-1}(1 + \log s)} \leq \psi_\psi(s) \leq 1.
$$

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

$$
K_{w,w}^N[f, \psi] := \inf \left\{ \|g\|_q + \|h\|_p : \int_{[0, 1]} f(\omega)w(s) = g(\omega, s) + h(\omega, s) \right\}
$$

in $L_1(\Omega \times [0, 1])$, $g \in L_0(\Omega \times [0, 1])$, $h \in L_0(\Omega \times [0, 1])$.

In Section 6 we will prove

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

$$
\|A^* f\|_{BMO_\psi} \leq \|B f\|_{\infty}.
$$

$(1)$ (EP) stands for "extrapolation property."
for all \( f \in E \). Then there is some \( d > 0 \), depending on \( \psi \in D \) and \( c \geq 1 \) only, such that for \( 1 \leq p < \infty \), \( t \geq 1 \), \( w_t := w_t^f \), and \( f \in E \),

\[
K_{\infty,p}^{\psi \rho}(Af, t^{1/p}) \leq d\psi^{-1}(p)K_{\infty,p}(Bf, t^{1/p}).
\]

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

\[
K_{\infty,p}[Af, t^{1/p}] \leq d\psi^{-1}(p)\psi^{-1}(1 + \log t)K_{\infty,p}[Bf, t^{1/p}].
\]

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 \( a_t = 1/\psi^{-1}(1 + \log t) \)

\[
\sup_{1 \leq i \leq N} \alpha_i Af(\omega_i) \leq d'\psi^{-2}(p) \sup_{1 \leq i \leq N} Bf(\omega_i) \leq K_{\infty,p}[f, N^{1/p}].
\]

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 1.3.** Let \( 1 \leq p < \infty \) and \( f \in L_p(\Omega, \mathcal{F}, \mathbb{P}) \). Assume \( N \geq 1 \) and \( w \in L^0_p[0, 1] \) with \( w = \sum_{i=1}^N \alpha_i X_{(i-1)/N, i/N} \). Then

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

In particular, if \( \alpha_1 = \ldots = \alpha_N = 1 \), then

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

The last two sections we will consider the assumptions of Theorem 1.7, property (EP) and the \( \| \cdot \|_{\text{BMO}_p} \cdot \| \cdot \|_{\text{oc}} \)-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 \( 0 < p < 1 \).

In Section 8 we consider self-similar operators which allow us to write the \( \| \cdot \|_{\text{BMO}_p} \cdot \| \cdot \|_{\text{oc}} \)-estimate of Theorem 1.7 in a much more applicable form. As a consequence, we sharpen in Corollary 8.6 an inequality of C. Pisier concerning “martingale-type constants” of Banach spaces.

### 2. The Lorentz spaces \( M_p^\varphi \) and \( M_p^0 \)

The spaces \( M_p^\varphi \) and \( M_p^0 \) we define below are the “basic modules” of \( \text{BMO}_p \)-spaces.

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

\[ f_p^\tau(t) := \left( \frac{1}{t} \int_0^t \left| f(s) \right|^p \, ds \right)^{1/p} \quad \text{and} \quad f^\tau(t) := f_1^\tau(t), \]

where \( \tilde{f}(s) := \inf\{c > 0 : \mathbb{P}(|f| > c) \leq s \} \) 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 \( \mathcal{C}_d \) is the subset of all \( \varphi \in \mathcal{C} \) such that \( \Delta(\varphi) := \sup_{0 < s < 1/2} \varphi(2s) - \varphi(s) < \infty \). Finally, for \( \varphi \in \mathcal{C} \) let

\[
M_p^\varphi(\Omega, \mathcal{F}, \mathbb{P}) := \{ f \in L_1(\Omega, \mathcal{F}, \mathbb{P}) : \| f \|_{M_p^\varphi} := \sup_{0 < s < 1/2} \tilde{f}(s) \varphi(s) < \infty \}
\]

and

\[
M_p^0(\Omega, \mathcal{F}, \mathbb{P}) := \{ f \in L_0(\Omega, \mathcal{F}, \mathbb{P}) : \| f \|_{M_p^0} := \sup_{0 < s < 1/2} \tilde{f}(s) \varphi(s) < \infty \}.
\]

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

\[
\varphi(t) = \frac{1}{\psi^{-1}(1 + \log(1/t))} \quad \text{and} \quad \varphi^{-1}(\frac{1}{\lambda}) = e^{\psi(\lambda)}.
\]

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

\[
\| f + g \|_{M_p^\varphi} \leq \Delta(\varphi) \| f \|_{M_p^\varphi} + \| g \|_{M_p^\varphi}
\]

for all \( f, g \in M_p^\varphi(\Omega, \mathcal{F}, \mathbb{P}) \). Furthermore, if one uses the relation (15) then one can write

\[
\| f \|_{M_p^\varphi} = \inf\{ c > 0 : \mathbb{P}(|f| > c) \leq e^{\psi(\lambda/c)} \}
\]

for all \( \lambda > c \).

In order to formulate some basic properties of the \( M_p^\varphi \)-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

\[
\begin{align*}
(1) & \, \chi_X \in X \quad \text{for all} \ A \in \mathcal{F}, \\
(2) & \, f \in X, \ g \in L_0(\Omega, \mathcal{F}, \mathbb{P}), \ \tilde{g} \leq \tilde{f} \quad \text{imply} \, g \in X \quad \text{and} \ |g| \leq |f|, \\
(3) & \, (\text{Fatou's property}) \quad \text{for} \ (f_n)_n \subseteq X \quad \text{with} \ f_n \geq 0 \ \text{a.s.,} \sup_n \|f_n\| < \infty, \ \text{and} \ f_n \uparrow f \in L_0(\Omega, \mathcal{F}, \mathbb{P}) \ \text{a.s.} \ \text{one has} \ f \in X \ \text{and} \ |f| = \lim_n |f_n|.
\end{align*}
\]

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 \( [M_p^\varphi(\Omega, \mathcal{F}, \mathbb{P}), \| \cdot \|_{M_p^\varphi}] \) is a r.i. Banach function space.

From the viewpoint of Banach space theory the spaces \( M_p^\varphi \) are more convenient than the spaces \( M_p^0 \). On the other hand, in view of (17) we use the spaces \( M_p^0 \) to construct our \( \text{BMO}_p \)-spaces. Hence we recall a characterization, which is folklore (see for example [3] (Lemma 2)), of those \( \varphi \in \mathcal{C} \).
such that \( M^p_\varphi (\Omega, \mathcal{F}, \mathbb{P}) = M^p_\varphi (\Omega, \mathcal{F}, \mathbb{P}) \) for all \( \Omega, \mathcal{F}, \mathbb{P} \). Using this characterization we deduce in particular that the BMO\(\varphi\)-spaces always have an equivalent norm.

**Theorem 2.4.** For \( \varphi \in C \) and \( c > 0 \) the following are equivalent.

1. \( \sup_{t \leq 1} \varphi (t) \leq c \left< f \right>_{M^2_\varphi (\Omega, \mathcal{F}, \mathbb{P})} \) for all \( \Omega, \mathcal{F}, \mathbb{P} \) and \( f \in M^2_\varphi (\Omega, \mathcal{F}, \mathbb{P}) \).
2. \( \frac{1}{t^\gamma} \int_0^t \varphi (s) \, ds \leq \frac{c}{\varphi (t)} \) for all \( 0 < t \leq 1 \).

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

\[
\mathbb{P} \left( \left\{ \omega : \sup_{1 \leq i \leq N} \left| f_i (\omega) \right| > \lambda \right\} \right) \leq \frac{\sum_{i=1}^N \mathbb{P} \left( \left| f_i (\omega) \right| > \lambda \right)}{1 + \sum_{i=1}^N \mathbb{P} \left( \left| f_i (\omega) \right| > \lambda \right)}.
\]

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 [18] (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

\[
h(s, \omega) := \sum_{i=1}^N \chi_{\left\{ \omega \in \Omega^N : f_i (\omega) > 0 \right\}}(s, \omega) \quad \text{and} \quad g(\omega_1, \ldots, \omega_N) := \sup_{1 \leq i \leq N} \left| f_i (\omega_i) \right|,
\]

where \([0, 1]\) is equipped with the Lebesgue measure \( \cdot \). Then for \( 0 < t \leq 1 \) one has

\[
\tilde{h} \left( \frac{t}{N} \right) \leq \tilde{g} \left( \frac{t}{t + 1} \right) \leq \tilde{g} \left( \frac{t}{2} \right).
\]

**Proof.** For \( 0 < s < 1 \) it follows from Lemma 3.1 that

\[
\tilde{g} (s) \geq \inf \left\{ c > 0 : \frac{\sum_{i=1}^N \mathbb{P} \left( \left| f_i (\omega) \right| > c \right)}{1 + \sum_{i=1}^N \mathbb{P} \left( \left| f_i (\omega) \right| > c \right)} \leq s \right\}
\]

\[
= \inf \left\{ c > 0 : \frac{N \left( 1 - s \right)}{1 + N \left( 1 - s \right)} \leq s \right\} = \tilde{h} \left( \frac{s}{N (1 - s)} \right).
\]

Setting \( s = \frac{t}{1 + t} \) yields the desired result. \( \blacksquare \)

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

\[
\tilde{h} \left( \frac{1}{N \gamma} \right) \leq \tilde{g} \left( \frac{1}{N \gamma} \right).
\]

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

(18) \[
f_p^{**}(t) \leq K_{\infty, p} \left| f, t^{-1/p} \right| \leq 2^{1-1/p} f_p^{**}(t).
\]

**Theorem 3.4.** Let \( 1 \leq p < \infty \), \( f_1, \ldots, f_N \in L^p(\Omega, \mathcal{F}, \mathbb{P}) \), and let \( h \in L^p([0, 1] \times \Omega) \) be defined by \( h(s, \omega) := \sum_{i=1}^N \chi_{\left\{ \omega \in \Omega^N : f_i (\omega) \right\}}(s, \omega) \). Then for \( M \geq 1 \) one has

\[
\frac{1}{2} K_{\infty, p} |h(\cdot), (MN)\rangle^{1/p} \leq \sup_{1 \leq i \leq N} \left< f_i (\omega_i) \right>_{L^p(\Omega^N)} \leq K_{\infty, p} |h(\cdot), (MN)\rangle^{1/p}.
\]

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

\[
H(s, \omega) := \sum_{i=1}^{MN} \chi_{\left\{ \omega \in \Omega^N : f_i (\omega) \right\}}(s, \omega) v_i(\omega)
\]

and

\[
G(\omega_1, \ldots, \omega_{MN}) := \sup_{1 \leq i \leq MN} \left< v_i (\omega_i) \right>.
\]

From Lemma 3.2 we know that \( \tilde{H}(t/MN) \leq \tilde{G}(t/2) \) for \( 0 < t \leq 1 \) so that \( H(t/MN) \leq \gamma \mathbb{E} \left< |G| \right> \). Since \( \tilde{H} = \tilde{h} \) 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

\[
\left< f_i (\omega_i) \right>_{L^p(\Omega^N)} \leq \left< a \right>_{L^\infty(\Omega^N)} \left< b \right>_{L^p(\Omega^N)} \leq \left< a \right>_{L^\infty(\Omega^N)} M^{1/p} \left< b \right>_{L^p(\Omega^N)}.
\]
\[ \|a\|_\infty + (MN)^{1/p} \left( \sum_{t=1}^{N} \frac{1}{t^{1/N}} \left[ \int_{A} |b(t, \omega_t)|^p \, dP(\omega_t) \right] \right)^{1/p} \]

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

Since this holds for all decompositions \( h = a + b \) as above we are done. \( \blacksquare \)

4. BM\( \mathcal{O}_\psi \)-spaces. We investigate a scale of BM\( \mathcal{O}_\psi \)-norms for adapted sequences of random variables. To define these BM\( \mathcal{O}_\psi \)-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-\( \sigma \)-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 \( \cdot \cdot \cdot \) defined on the Borel \( \sigma \)-algebra and is equipped with the filtration \( (\mathcal{F}_k)_{k=0}^\infty \) generated by \( \mathcal{F}_0 := \{\{0\}, \{0,1\}\} \), and, for \( k \geq 1 \), by

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

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

The BM\( \mathcal{O}_\psi \)-norm for an adapted sequence \( (f_k)_{k \in \mathbb{N}} \in \mathcal{A}((\mathcal{F}_k)_{k \in \mathbb{N}}) \) was defined in [14] (p. 66) by

\[ \|(f_k)_{k \in \mathbb{N}}\|_{\text{BM} \mathcal{O}_\psi} := \sup_{0 < k \in \mathbb{N}} \mathbb{E}(\|f_k - f_{k-1}\|_{L_1(\mathcal{F}_k)}). \]

(19) \( \|f\|_{\text{BM} \mathcal{O}_\psi} := \|(\mathbb{E}(f \mid \mathcal{F}_k))_{k \in \mathbb{N}}\|_{\text{BM} \mathcal{O}_\psi} = \sup_{0 < k \in \mathbb{N}} \sup_{B \in \mathcal{F}_k \mathbb{P}(B) > 0} \|f - f_{k-1}\|_{L_1(\mathbb{P}B)}. \)

Let us recall a well-known BM\( \mathcal{O}_\psi \)-function defined on \([0,1]\) used later.

**Example 4.1.** For \( f(t) := \log(1/t) \in L_1[0,1] \) one has \( \|f\|_{\text{BM} \mathcal{O}_\psi} < \infty. \)

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

\[ \|(f_k)_{k \in \mathbb{N}}\|_{\text{BM} \mathcal{O}_\psi} = \sup_{0 \leq k \leq l} \sup_{B \in \mathcal{F}_k \mathbb{P}(B) > 0} \|f_l - f_{k-1}\|_{L_1(\mathcal{F}_k)} \]

where

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

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

**Definition 4.2.** Let \( \psi \in \mathcal{D}, 0 < s < 1, \) and \( (f_k)_{k \in \mathbb{N}} \in \mathcal{A}((\mathcal{F}_k)_{k \in \mathbb{N}}) \). Then

\[ \|(f_k)_{k \in \mathbb{N}}\|_{\text{BM} \mathcal{O}_\psi^s} := \sup_{0 \leq k \leq l} \sup_{B \in \mathcal{F}_k \mathbb{P}(B) > 0} \|f_l - f_{k-1}\|_{L_1(\mathcal{F}_k)} \]

where \( \mathcal{F}_k := \mathcal{F}_{k+1} / \mathbb{P}(B) \) is the normalized restriction of \( \mathbb{P} \) to \( B \) and \( f_{-1} = 0 \).

In the literature several BM\( \mathcal{O}_\psi \)-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 4.1(i) 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

\[ \overline{\psi}(\mu) = \sup \left\{ N \left( \frac{1}{\psi(\mu)} \right) - 1 \mid \mu = \sum_{i=1}^{N} \mu_i, \mu_i \geq 1, N = 1, 2, \ldots \right\} \]

we set, by analogy with (15),

\[ \overline{\psi}(t) := \frac{1}{\psi^{-1}(1 + \log(1/t))}. \]

Let us summarize some properties of this construction.

**Example 4.3.** (1) For \( \psi(\mu) := 1 + \mu/\alpha \) one has \( \overline{\psi}(\mu) \geq \mu/\alpha. \)

(2) For \( 1 \leq p < \infty \) and \( \psi(\mu) := \mu^p \) one has \( \overline{\psi}(\mu) = \mu^p. \)

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

\[ \overline{\psi}(\mu) = \frac{1}{\psi^{-1}(2)} \text{ and } \overline{\psi}(k\psi^{-1}(2)) = \frac{1}{\psi^{-1}(2)} \text{ so that} \]

\[ \frac{\mu}{\psi^{-1}(2)} \geq \frac{1}{\psi^{-1}(2)} - 1 \geq \frac{\mu}{\psi^{-1}(2)} \]

(2) Here we exploit \( \sum_{i=1}^{N} (\mu_i - 1) + 1 \leq (\sum_{i=1}^{N} \mu_i)^p \) for \( \mu_i \geq 0. \) \( \blacksquare \)
LEMMA 4.4. For \( \psi \in D \) the following holds.

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

\[
\beta(\mu) + \beta(\nu) \leq \beta(\mu + \nu).
\]

2. \( \overline{\psi} \in C_\Delta \) and there is some \( c_\psi \geq 1 \), depending on \( \psi \) only, such that for all \( \mu, \lambda \geq 1 \),

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

3. \( \|f\|_p \leq 4 \overline{\psi}^{-1}(1 + \log 4) \|g\|_p \).

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 \( f, g \in L^q_t(\Omega, F, \mathbb{P}) \) and \( 0 < p < \infty \) be such that for some \( \beta > 1 \) and \( \delta, \varepsilon > 0 \) with \( \beta \delta < 1 \) one has

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

for all \( \lambda > 0 \). Then

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

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

\[
\overline{\alpha}(\mu) := \sup \left\{ \sum_{i=1}^N \alpha(\mu_i) \mid \mu = \sum_{i=1}^N \mu_i, \mu_i \geq 1, N = 1, 2, \ldots \right\}.
\]

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 \( \beta \) satisfying \( \beta(\mu) + \beta(\nu) \leq \beta(\mu + \nu) \) we get

\[
\beta(\mu) \geq \sup \left\{ \sum_{i=1}^N \beta(\mu_i) \mid \mu = \sum_{i=1}^N \mu_i, \mu_i \geq 1, N = 1, 2, \ldots \right\}
\]

\[
\geq \sup \left\{ \sum_{i=1}^N \alpha(\mu_i) \mid \mu = \sum_{i=1}^N \mu_i, \mu_i \geq 1, N = 1, 2, \ldots \right\} = \overline{\alpha}(\mu).
\]

(2) Clearly \( \overline{\alpha}(1) = 0 \) and \( \lim_{\mu \to \infty} \overline{\alpha}(\mu) = \infty \). For \( N \geq 1 \) we define

\[
\alpha_N : [1, \infty) \to [0, \infty)
\]

\[
\alpha_N(\mu) := \sup \left\{ \sum_{i=1}^N \alpha(\mu_i) \mid \mu = \sum_{i=1}^N \mu_i, \mu_i \geq 1 \right\} \text{ if } 1 \leq \mu \leq N,
\]

\[
\alpha_N(\mu) := \sum_{i=1}^N \alpha(\mu_i) \text{ if } \mu > N.
\]

so that \( \overline{\alpha}(\mu) = \sup_{N \leq \mu} \alpha_N(\mu) \). If \( \mu \geq N \), then the continuity of \( \alpha \) 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 \( \alpha \) 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

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

The function \( \overline{\alpha} \) is continuous. Since \( \overline{\alpha}(\mu) = \sup_{N \leq \mu} \alpha_N(\mu) \) for \( \tau \geq 1 \) it suffices to show that \( \alpha_N(\mu) \) is continuous, and this follows from

\[
\sup_{N \leq \mu \leq \delta} \alpha_N(\nu) \leq N \sup_{N \leq \mu \leq \delta} |\alpha(\nu) - \alpha(\mu)|
\]

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

\[
(\psi^{-1}(2) + 1) \overline{\psi}(\mu \lambda) \geq \psi^{-1}(2) \overline{\psi}(k \lambda) + \psi(k \lambda) \geq k + \overline{\psi}(k \lambda)
\]

\[
\geq k + \overline{\psi}(\lambda) = k \overline{\psi}(\lambda) \geq k/\psi(\lambda).
\]

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

\[
\sup_{0 \leq t \leq 1} \overline{\psi}(t) = \sup_{0 \leq t \leq 1} \overline{\psi}^{-1}(1 + \log 2 + \log(1/t))
\]

\[
\leq \sup_{0 \leq t \leq 1} \overline{\psi}^{-1}(1 + \log(1/t)) \leq \mu_0.
\]

(3) For \( \lambda \geq 1 \) and \( f_0(s) := \psi(s)/(\overline{\psi}(s)) \) (0 < s, t ≤ 1) we have

\[
\sup_{\lambda \geq \lambda_0} \left\{ \int_{f_0 > 2x} \overline{\psi}(x) \right\} \leq \sup_{\lambda \geq \lambda_0} \left\{ \int_{f_0 > 2\lambda} \overline{\psi}(x) \right\} \leq \sup_{\lambda \geq \lambda_0} \left\{ \int_{f_0 > \lambda} \overline{\psi}(x) \right\} \leq \sup_{\lambda \geq \lambda_0} \overline{\psi}^{-1}(1/2\lambda) \leq \sup_{\lambda \geq \lambda_0} \overline{\psi}^{-1}(1/\lambda).
\]

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

\[
\sup_{\lambda \geq \lambda_0} \left\{ \int_{f_0 > 2\lambda} \overline{\psi}(x) \right\} \leq \sup_{\lambda \geq \lambda_0} e^{\overline{\psi}(\lambda) - (\overline{\psi}(2\lambda) - 1)} \leq \sup_{\lambda \geq \lambda_0} e^{1 - \overline{\psi}(\lambda)} = \frac{1}{4}.
\]

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

\[
\int_{f_0 > 2\lambda} \overline{\psi}(x) \leq \frac{1}{4} \int_{f_0 > \lambda} \overline{\psi}(x).
\]
Applying Lemma 4.5 gives \( \|f_i\|_1 \leq 4\psi^{-1}(1 + \log 4) \), which is exactly the assertion of Theorem 2.4(2) with \( c = 4\psi^{-1}(1 + \log 4) \).

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 \( \psi \in D \) and arrive at \( \psi \in D \) which describes the same BMO-space as \( \psi \) does. The weight function \( \psi \) 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 \( \text{BMO}_\psi \)-spaces (see Theorem 1.3). Moreover, the passage \( \psi \to \psi \) 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 \( \psi(\mu) \geq \mu/e \) (see Example 4.3), which is the John–Nirenberg Theorem as demonstrated in Corollary 4.8.

**Theorem 4.6.** (a) For \( \psi \in D, 0 < s < 1, \) and \( \psi_i(t) = t \) one has

\[
\|\cdot\|_{\text{BMO}_\psi} \leq\|\cdot\|_{\text{BMO}_\psi^s} \leq \|\cdot\|_{\text{BMO}_\psi^s} \leq 6\psi^{-1}(3)\|\cdot\|_{\text{BMO}_\psi^s},
\]

\[
\|\cdot\|_{\text{BMO}_\psi^s} \leq \left(1, \frac{1}{\log(1/s)}\right)\|\cdot\|_{\text{BMO}_s},
\]

\[
\|\cdot\|_{\text{BMO}_\psi^{s+1}} \leq 3\max\left(1, \frac{1}{\log(1/s)}\right)\|\cdot\|_{\text{BMO}_s},
\]

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

\[
P_\psi\left(\sup_{i \leq I} |f_i| > \lambda + \mu\right) \leq e^{-\psi(\mu)}P_\psi\left(\sup_{i \leq I} |f_i| > \lambda\right).
\]

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

\[
\tau_B := \inf\{t \geq k | \chi_B|f_i - f_{k-1}| > \mu\} \wedge (n + 1).
\]

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

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

\[
P_B\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda + \mu\right)
\]

\[
= \sum_{i=k}^n P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > \lambda + \mu, \tau_B = i\right)
\]

\[
\leq \sum_{i=k}^n P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > |f_i - f_{k-1}| + \mu, \tau_B = i\right)
\]

\[
\leq \sum_{i=k}^n P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > \mu, \tau_B = i\right).
\]

(ii) Following an idea contained in [14] (p. 75ff) we get for \( \lambda > 0, \mu \geq 1, \) and \( \sup_{k \leq i \leq n} |f_i - f_{k-1}| \leq \lambda + \mu \),

\[
P_B\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda + \mu\right)
\]

\[
= \sum_{i=k}^n P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > \lambda + \mu, \tau_B = i\right)
\]

\[
= \sum_{k \leq \lambda + \mu} P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > \lambda + \mu, \tau_B = i, \tau_{\lambda + \mu} = \lambda\right)
\]

\[
\leq \sum_{k \leq \lambda + \mu} P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > |f_i - f_{k-1}| + \mu, \tau_B = i, \tau_{\lambda + \mu} = \lambda\right)
\]

\[
\leq \sum_{k \leq \lambda + \mu} P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > (\mu - 1)/2, \tau_B = i, \tau_{\lambda + \mu} = \lambda\right)
\]

\[
+ P_B\left(\sup_{i \leq n} |f_i - f_{k-1}| > (\mu - 1)/2, \tau_B = i, \tau_{\lambda + \mu} = \lambda\right).
\]

(iii) We show \( \|\cdot\|_{\text{BMO}_\psi} \leq \|\cdot\|_{\text{BMO}_\psi} \). If \( \|(f_k)_{k \in I}\|_{\text{BMO}_\psi} \leq 1 \), then step (i) gives

\[
P_B\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda + \mu\right)
\]

\[
\leq \sum_{i=k}^n P\left(\tau_B = i \right) \wedge B \cap \{ \sup_{i \leq n} |f_i - f_{k-1}| > \mu\}/P(B)
\]

\[
\leq e^{-\psi(\mu)}P\left(\tau_B = i \right) \wedge B \cap \{ \sup_{i \leq n} |f_i - f_{k-1}| > \mu\}/P(B)
\]

\[
= e^{-\psi(\mu)}P\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda\right).
\]

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

\[
P_B\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda + \mu\right) \leq e^{-\psi(\mu)}P\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \lambda\right)
\]

and (24). Now \( \lambda \downarrow 0 \) gives \( P_B\left(\sup_{k \leq i \leq n} |f_i - f_{k-1}| > \mu\right) \leq e^{-\psi(\mu)} \) and \( \|(f_k)_{k \in I}\|_{\text{BMO}_\psi} \leq 1 \).

(iv) To show \( \|\cdot\|_{\text{BMO}_\psi} \leq 6\psi^{-1}(3)\|\cdot\|_{\text{BMO}_\psi} \) let \( \|(f_k)_{k \in I}\|_{\text{BMO}_\psi} \leq 1 \), which implies
\[ \| f_{l} - f_{l-1} \|_\infty \leq \sup_{P \in P} \| f_{l} - f_{l-1} \|_{M^q_P(B,P_B)} \leq 1 \]

for \( l \in I \). Since \( B \cap \bigcup_{\lambda=1}^{\infty} \{ \gamma : \gamma \cap \mu = \{ l \} \} = B \cap \{ \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \} \),

step (ii) gives for \( \mu \geq 3 \), as in step (iii),

\[
\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + \mu \right) \leq 2e^{1-\varphi((\mu-1)/2)}\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right).
\]

For \( \varphi_0 := 3 \varphi^{-1}(1+\log 4) \leq 3 \varphi^{-1}(3) \) and \( \mu_2 \geq \varphi_0 \), using \( (\mu_2 - 1)/2 \geq \mu_6/6 \) we conclude

\[
\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + \mu \right) = \mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + \frac{\mu_2}{2} + \frac{\mu_6}{2} \right) \leq e^{1-\varphi(\mu_2/6)}\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right) \leq e^{1-\varphi(\mu_2/6)}\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right).
\]

which as in (iii) implies the inequality \( \| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 2 \varphi_0 \).

(v) The relations \( \| \cdot \|_{\text{BMO}_q} \leq \| \cdot \|_{\text{BMO}_q} \leq \| \cdot \|_{\text{BMO}_q} \) are obvious.

(vi) Assume that \( \| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 3 \). Since for all \( l \in I \),

\[
\| f_{l} - f_{l-1} \|_\infty \leq \| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 3,
\]

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

\[
\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + 3 \right) \leq \frac{2}{3} \mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right)
\]

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

\[
\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + \mu \right) \leq \mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + 3p \right) \leq \mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right) \leq e^{1-\varphi(\mu/3)\min(1,\log(1/\delta))}\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right)
\]

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

\[
\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda + 1 \right) \leq e^{1-\varphi(\mu/3)\min(1,\log(1/\delta))}\mathbb{P}_B \left( \sup_{k \leq \lambda} \| f_{l} - f_{l-1} \| > \lambda \right).
\]

The rest is similar to the proof of (23). The left-hand side of (22) follows from \( \mathbb{h}(s) \leq 1/(1+\log(1/s))\mathbb{h}(1) \) if \( \varphi_1(t) := 1/(1+\log(1/t)) \) and \( \mathbb{h} \in \text{M}^q_{A_1} \).

Remark 4.7. (i) In [24] there is defined a quantity \( \| f \|_{\text{BMO}_q} \) for functions \( f \in L^q(\mathbb{R}^n) \) in a slightly different form. Lemma 3.1 of that paper, referred to in [17], is a counterpart to (23).

(ii) The proof of (23) gives, on the right-hand side, \( \| \cdot \|_{\text{BMO}_{q/2}} \) instead of the smaller quantity \( \| \cdot \|_{\text{BMO}_q} \), which is sufficient for the purpose of this paper. Nevertheless it could be of interest to investigate \( \| \cdot \|_{\text{BMO}_q} \) whenever \( 1/2 \leq s < 1 \), 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. \( \| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 6^3\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \) for \( (f_{k})_{k \in I} \in \mathcal{A}(\{f_{k}\}_{k \in I}) \).

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

\[
\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq e\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq e\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 6^3\| (f_{k})_{k \in I} \|_{\text{BMO}_q}.
\]

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 \( s = 1/3 \) inequality (23) and the property \( \mathbb{h}(s) \leq 1/(1/s)\mathbb{h}(1) \) yield

\[
\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 3\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq 6\| (f_{k})_{k \in I} \|_{\text{BMO}_q}.
\]

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.e. for all \( k \in I \).

Theorem 4.9. If \( \psi \in \mathcal{D} \), then \( \text{BMO}_\psi(\{f_{k}\}_{k \in I}) \), \( \| \cdot \|_{\psi} \) is a Banach space where

\[
\| (f_{k})_{k \in I} \|_{\psi} := \sup_{0 \leq \lambda \leq \lambda} \sup_{B \supseteq \lambda \in I} \mathbb{P}_B \left( \| f_{l} - f_{l-1} \|_{M^q_P(B,P_B)} > \lambda \right)
\]

and \( f_{-1} = 0 \). Moreover,

\[
\| (f_{k})_{k \in I} \|_{\text{BMO}_q} \leq \| (f_{k})_{k \in I} \|_{\psi} \leq 24\psi^{-1}(3)\psi^{-1}(3)\| (f_{k})_{k \in I} \|_{\text{BMO}_q}.
\]

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

$$\|g\|_{M_{R}(B, \mathbb{P}_A)} \leq \sup_{0 < t \leq 1} \frac{\overline{\psi}(t)}{\overline{\psi}(2t)} \|g\|_{M_{\mathcal{G}}(\mathcal{F}_k, \mathbb{P})} \leq \frac{c_1}{\overline{\psi}(2t)} \|g\|_{M_{\mathcal{G}}(\mathcal{F}_k, \mathbb{P})} < \infty \quad \text{for } g \in M_{R}(\Omega, \mathcal{F}_k, \mathbb{P})$$

(note that $\Delta(\overline{\psi}) < \infty$), we get $\|(f_{k-1} - f_k)\|_{M_{R}(B, \mathbb{P}_A)} \leq \lim \sup_{n} \|x_n - x_{n-1}\|_{\overline{\mathcal{F}}}$ from

$$\|(f_{k-1} - f_k)\|_{M_{R}(B, \mathbb{P}_A)} = \lim \sup_{n} \|(f_{k-1} - f_k)\|_{M_{R}(B, \mathbb{P}_A)}.$$

Now let us introduce the $\text{BMO}_{\overline{\psi}}$-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 $f \in \text{BMO}_{\overline{\psi}}(\mathcal{F}_k)$ provided that

$$\|f\|_{\text{BMO}_{\overline{\psi}}} := \|(E_k f)_{k \in I}\|_{\text{BMO}_{\overline{\psi}}} < \infty.$$

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

$$\|f\|_{\text{BMO}_{\overline{\psi}}} \leq \sup_{0 \leq t \leq 1} \frac{\|f - E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}}{\|E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}} \leq \sup_{0 \leq t \leq 1} \frac{\|f - E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}}{\|E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}}$$

where the multiplicative constants involved in these inequalities depend on $\psi$ only. In fact, let $B \in \mathcal{F}_k$ with $|B| > 0$ be the disjoint union of $\mathcal{F}_k$-atoms $B^1$. A standard computation shows the estimate $|h|_{M_{\mathcal{G}}(B, \mathbb{P}_A)} \leq \sup_{I} |h|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}$.

Hence

$$\|f\|_{\text{BMO}_{\overline{\psi}}} = \sup_{0 \leq t \leq 1} \frac{\|E_{\psi} f - E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}}{\|E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}}.$$

Since $\sup_{t \geq k} E_{\psi} f = E_{\psi} f - E_{\psi} f\|_{M_{\mathcal{G}}(B, \mathbb{P}_A)}$ as a consequence of [4] (Theorems 1.1.7, 2.4.8) and $E_{\psi} f \to f$ a.s. we are done by Lemma 4.4 and Theorem 4.6.

The basic example of a $\text{BMO}_{\overline{\psi}}$-function defined on the unit interval $[0, 1]$ is given by

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

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

The theorem follows immediately from Example 4.1 and

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

$$\overline{\psi}^{-1}(1 + f) \in \text{BMO}_{\overline{\psi}}[0, 1] \leq c_1 \|1 + f\| \text{BMO}[0, 1],$$

where $c > 0$ depends on $\psi$ only.

**Proof.** (1) Let $\psi(\mu) = \mu$, corresponding to $\varphi_1(t) = 1/(1 + \log(1/t))$, and let $\psi \in \mathcal{D}$ so that one has $\psi(\lambda) \leq c_\psi(\mu \lambda)$ for $\lambda, \mu \geq 1$. Assume $h \in M_{\mathcal{G}}^{\infty}(\Omega, \mathcal{F}, \mathbb{P})$ to be non-negative. Then

$$|\psi^{-1}(1 + h)|_{M_{\mathcal{G}}^{\infty}} \leq c_\psi |1 + h|_{M_{\mathcal{G}}^{\infty}}.$$

Indeed, if $a := |1 + h|_{M_{\mathcal{G}}^{\infty}} > 0$, then for $\mu \geq 0$ we derive

$$\mathbb{P}(|\psi^{-1}(1 + h)| > \mu) = \mathbb{P}(1 + h > \psi(\mu)) \leq e^{-\psi(\mu)/a} \leq e^{-\psi(\mu)/a}.$$

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

$$|g|_{\text{BMO}_{\overline{\psi}}[0,1]} \leq c_\psi |1 + f|_{\mathcal{G}[0,1]} \leq c_\psi |1 + f|_{\text{BMO}[0,1]}.$$

(3) Now let $k \geq 1$. Let $B$ be an atom and $\overline{\mathcal{F}} \in \mathcal{F}_k$ be the dyadic predecessor. If $g_k \equiv E_k(g | F_k)$ and $f_k \equiv E_k(f | F_k)$ ($k = 0, 1, 2, \ldots$), then for $t \in B$ one has

$$|g(t) - g_{k-1}(t)| \leq \frac{1}{|B|} \int_B |g(t) - g(s)| \, ds.$$

For $r, s \geq 0$ we use

$$\overline{\psi}^{-1}(r + s + 1) \leq \overline{\psi}^{-1}(r + 1) + \overline{\psi}^{-1}(s + 1),$$

which follows from Lemma 4.4(1), and

$$\overline{\psi}^{-1}(r + 1) \leq (r + 1)\overline{\psi}^{-1}(2)$$

(consider $k \leq r \leq k + 1$ for $r \in [0, 1, 2, \ldots]$) and

$$\overline{\psi}^{-1}(r + 1) \leq \overline{\psi}^{-1}(k+2) \leq (k + 1)\overline{\psi}^{-1}(2)$$

and continue for $t \in B$ to get

$$|g(t) - g_{k-1}(t)|$$

$$\leq \frac{1}{|B|} \int_B \overline{\psi}^{-1}(|f(t) - f(s)| + 1) \, ds$$

$$\leq \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + \frac{1}{|B|} \int_B \overline{\psi}^{-1}(|f_{k-1}(t) - f(s)| + 1) \, ds$$

$$\leq \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + \overline{\psi}^{-1}(2)|f|_{\text{BMO} + 1}$$

$$\leq \overline{\psi}^{-1}(|f(t) - f_{k-1}(t)| + 1) + 2\overline{\psi}^{-1}(2)|f|_{\text{BMO} + 1}.$$

From (27) and $|\cdot|_{M_{\mathcal{G}}^{\infty}(B, \mathbb{P})} \leq \|\cdot\|_{\mathcal{G}(B, \mathbb{P})}$ it follows that

$$|\overline{\psi}^{-1}(|f - f_{k-1}| + 1)|_{\mathcal{G}(B, \mathbb{P})} \leq c_\psi |f - f_{k-1}|_{\mathcal{G}(B, \mathbb{P})} + c_\psi.$$

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

$$\|f - f_{k-1}\|_{\mathcal{G}(B, \mathbb{P})} = \lim_{t \to 1} \|f(t) - f_{k-1}(t)\|_{\mathcal{G}(B, \mathbb{P})}$$

$$\leq \|f(t)\|_{L_1} \leq c_\psi \|f\|_{\text{BMO} + 1}.$$
Summarizing all the estimates (and using $\Delta(\varphi) < \infty$ and (16)) we arrive at $|g - g_{k-1}|_{M^p}(\Omega, F_k) \leq d_0(1 + \log(1/\delta))$ for some $d_0 > 0$ depending on $\delta$ only. Now steps (2) and (3) combined with Remark 4.11 give the assertion of the lemma.

\textbf{Proof of Theorem 1.3.} In view of Theorem 4.6(21) we get (1) $\Rightarrow$ (2). We show (2) $\Rightarrow$ (1). For this purpose let $f(t) = \varphi^{-1}(1 + \log(1/\delta)) \in L_1[0, 1]$ and $f_k = E\varphi(f)$. Then it follows from Theorem 4.6, our assumption, and Theorem 4.12 that

$$|f|_{M^p} \leq \|f\|_{BMO_{\varphi}} \leq 6\psi^{-1}(3)|f|_{L_1} \leq 6\psi^{-1}(3)|f|_{M^p} \leq b$$

so that $\varphi(t) \leq b\varphi(t)$ for $0 \leq t \leq 1$. Via (20) we arrive at the desired result. $\blacksquare$

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

\textbf{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 $\psi : [0, \infty) \to [0, \infty)$ be its extension with $\psi(t) = t$ for $0 \leq t \leq 1$. We assume $\psi(0) = 0$ and that $\psi(t)$ is convex and a Young function $\Phi(t) := \psi(t) - t$ with the corresponding Orlicz norm $\|f\|_{\ell_\varphi} := \inf\{c > 0 \mid \mathbb{E}[f^\Phi(c)] \leq 1\}$ for $f \in L_1(\Omega, F, \mathbb{P})$. Using the condition $\sup_{a > 0} a\psi(a)/\psi(\lambda) > 1$, one can verify $\|f\|_{\ell_\varphi} \sim \|f\|_{M^p}$. Now consider for simplicity $\Omega = [0, 1]$ and $f \in L_1[0, 1]$. For $f_k := \mathbb{E}(f | F_k)$ we obtain, as in Remark 4.11,

$$\|(f_k)_{k \geq 0}\|_{BMO_{\varphi}} \sim \sup_{k \geq 0} \sup_{B \in \mathcal{B}_{F_k}} \|f - f_{k-1}\|_{\mathcal{M}_{\varphi}(B, F_{k-1})} \sim \sup_{k \geq 0} \sup_{B \in \mathcal{B}_{F_k}} \|f - f_{k-1}\|_{\mathcal{M}_{\varphi}(B, F_{k-1})} = \sup_{k \geq 0} \inf_{\mathcal{M}} \|\mathcal{M}\|_{\infty}$$

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

$$\mathbb{E}(\Phi(|f - f_{k-1}|/\gamma | F_k)) \leq 1 \quad \text{a.s.}$$

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 $\psi$ of $\psi$, which is the key of our approach, to extend [2] in some directions.

(2) Let $(f_k)_{k \in \mathbb{N}} \in BMO_{\varphi}((\mathcal{F}_k)_{k \in \mathbb{N}})$. If $[\Omega, \mathcal{F}, \mathbb{P}] := (\Omega, F, \mathbb{P}) \times (\Omega', F', \mathbb{P}')$, $\mathcal{F}'_k := \mathcal{F}_k \times \mathcal{F}'_k$, and $g_{\omega'}(\omega, \omega') := f_k(\omega)$, then $\|(f_k)_{k \in \mathbb{N}}\|_{BMO_{\varphi}} = \sup_{B \in \mathcal{F}_k} \mathbb{P}_B(|g_k - g_{k-1}| > \lambda) \leq \lambda$ for $\lambda > 0$ and $k \leq l \in I$.

\textbf{5. Weak upper estimates of $BMO_{\varphi}$-spaces.} Now we will prove Theorem 1.5 and deduce in Corollary 5.3 that the $BMO_{\varphi}$-spaces can also be classified by their weak upper estimates.

\textbf{Lemma 5.1.} For $1 \leq a_1 \leq \ldots \leq a_N > 0$ and $\varphi \in \mathcal{C}$ the following are equivalent.

1. $\sup_{i \leq N} \varphi^{-1}(a_i/\lambda) \leq \varphi^{-1}(1/\lambda)$ for all $\lambda \geq 1$.
2. $\sup_{i \leq N} |a_i|_{M^p} \leq \sup_{i \leq N} |f_i|_{M^p}$ for all $\Omega, F, \mathbb{P}$ and $(f_i)_{i=1}^N \subseteq M^p(\Omega, F, \mathbb{P})$.

\textbf{Proof.} (1) $\Rightarrow$ (2). Assume that $\sup_{i \leq N} |f_i|_{M^p} = 1$. Then for $\lambda \geq 1$,

$$P(\sup_i |a_i|/\lambda < \lambda) \leq \sum_i P(|a_i|/\lambda < \lambda) \leq \sum_i e^{-\varphi^{-1}(\lambda)} \leq \sum_i e^{-\varphi^{-1}(1/\lambda)} = e^{-\varphi^{-1}(1/\lambda)} \leq e^{-\varphi^{-1}(1)} = 1.$$

(2) $\Rightarrow$ (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 $f_i := \mathbb{1}_{[a_i]} \chi_{A_i}$. Then

$$\sup_{i \leq N} |A_i|_{M^p(0, 1)} = \sup_i \frac{|A_i|_{M^p(0, 1)}}{a_i} = \sup_i \frac{\varphi(t_i)}{a_i} = \frac{1}{\lambda} \leq 1.$$

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

$$\sum_{i=1}^N t_i = \varphi^{-1}\left(\sum_{i=1}^N t_i\right) = \varphi^{-1}\left(\sup_{i \leq N} |A_i|_{M^p(0, 1)}\right) \leq \varphi^{-1}(1/\lambda).$$

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

$$\sup_{1 \leq i \leq N} \psi(1/|f_i|_{M^p}) \leq c \sup_{1 \leq i \leq N} |f_i|_{M^p}.$$

\textbf{Proof.} Assuming $\psi(a) \geq (1 + \varepsilon)\psi(\lambda)$ for some $a > 1$ and $\varepsilon > 0$ we get $\psi(a^n) \geq (1 + \varepsilon)^n \psi(\lambda)$. Consequently, we find some $b > 1$ with
Applying Theorem 4.6(21) gives $U_a(BMO_\psi) \leq c_0 \psi^{-1}(3)$. On the other hand, we take the function

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

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

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

On $[0,1]^N$ we will use the canonical product filtration $(\mathcal{F}_k)_{k=0}^\infty$ with $\mathcal{F}_k := \mathcal{F}_k \upharpoonright \mathcal{F}_0$, where $(\mathcal{F}_k)_{k=0}^\infty$ is the dyadic filtration on $[0,1]$ taken in Section 4. Theorem 1.8 and (18) imply

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

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

$$ \| f \|_{L_1(\mathbb{R}^+ \times \mathbb{R}^+)} = \liminf_{t \to \infty} \sup_{1 \leq i \leq N} f(t_i) \leq \| \sup_{1 \leq i \leq N} f^{(i)}(t_i) \|_{L_1([0,1]^N)} \leq \left( \sup_{1 \leq i \leq N} f^{(i)} \right)_{k=0}^\infty \leq \| f \|_{BMO_\psi} \leq 24 \psi^{-1}(3) \psi^{-1}(3) \left( \sup_{1 \leq i \leq N} f^{(i)} \right)_{k=0}^\infty \psi_{\infty} \psi_{\infty} \psi_{\infty} \psi_{\infty} \psi_{\infty}. $$

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

$$ \theta \leq \psi(1/N) \left( \sup_{1 \leq i \leq N} f^{(i)} \right)_{k=0}^\infty \| BMO_\psi \| (\mathcal{F}_k)_{k=0}^\infty \| \psi \|_{\infty} \psi_{\infty} \psi_{\infty} \psi_{\infty} \psi_{\infty} \psi_{\infty} \leq U_{BMO_\psi} \| f \|_{BMO_\psi} < \infty, $$

we are done.

As a consequence of Theorem 4.14(2) and Theorem 4.12, $\| (\mathcal{F}_k)_{k=0}^\infty \|_{BMO_\psi} = \| f \|_{BMO_\psi} < \infty$, 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 $\psi(\mu) \leq \psi(\varphi_1 \mu)$ for all $1 \leq \mu < \infty$.
2. There exists $c_2 > 0$ such that for all $a \in (a_1)c_2$ satisfying $1 \geq a_1 \geq a_2 \geq \ldots \geq 0$ one has

$$ U_a(BMO_\psi) \leq c_2 U_a(BMO_\psi). $$

**Proof.** First for $\psi \in C$ and $N = 1, 2, \ldots$, set

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

where $\varphi(1/N)$ is repeated $N$ times ($\varphi(1/i) = 1/\psi^{-1}(1 + \log i)$).
(1)⇒(2). If \( U_a(BMO_\psi) < \infty \), then it follows from Theorem 1.5 and (20) that
\[
0 < \theta := \inf_N U_{c_\theta} \varphi(N) \leq \inf_N \frac{\varphi(1/N)}{a_N} U_a(BMO_\psi) \leq c_1 \inf_N \frac{\varphi(1/N)}{a_N} U_a(BMO_\psi)
\]
or \( a_N \leq d\varphi(1/N) \) with \( d := (c_1/\theta) U_a(BMO_\psi) \). Hence,
\[
U_a(BMO_\psi) \leq dU_{c_\theta} \varphi(BMO_\psi) = \left[ \frac{c_1}{\theta} U_{c_\theta} \varphi(BMO_\psi) \right] U_a(BMO_\psi) = c_2 U_a(BMO_\psi)
\]
where \( c_2 < \infty \) according to Theorem 1.5.

(2)⇒(1). Since for \( \theta' := \inf_N U_{c_{\theta'}} \varphi(N) > 0 \),
\[
\theta' \frac{\varphi(1/N)}{\varphi(1/N)} \leq U_{c_{\theta'}} \varphi(BMO_\psi) \leq U_{c_{\theta'}} \varphi(BMO_\psi) \leq c_2 U_{c_{\theta'}} \varphi(BMO_\psi),
\]
for \( c_{\theta'} := (c_2/\theta') U_{c_{\theta'}} \varphi(BMO_\psi) \) and \( N = 1,2,\ldots \) one has the estimate \( \varphi(1/N) \leq c_{\theta'} \varphi(1/N) \). Since \( \Delta(\varphi) < \infty \) according to Lemma 4.4(2), a standard computation gives \( \varphi(t) \leq c_2 \Delta(\varphi\varphi)(t) \) for \( 0 \leq t \leq 1 \). Using (20) we can conclude the proof. \( \blacksquare \)

6. A general extrapolation principle. We want to prove Theorem 1.7. For simplicity we assume that \( I = \{0,\ldots,n\} \) to avoid the existence of the maximal operator \( A^\delta \). Let us start with some lemmata.

**Lemma 6.1.** Let \( \psi \in \mathcal{D} \), and \( S, T : A^\mathbb{N}((\Sigma_k)_{k=0}^n) \to L^\mathbb{N}(M, \Sigma, m) \), where \( X \) is a Banach space. If \( (F, S, T) \) satisfies (EP) with constant \( c \geq 1 \), then
\[
\|Sf\|_{M, BMO_\varphi} \leq \|Tf\|_{\infty} \quad (f \in F)
\]
implies, for \( 0 < \delta \leq 1 \) and \( \lambda > 0 \),
\[
m(S^{\lambda} f) > (c + c^2)\lambda, \quad T f \leq \delta \lambda \leq e^1(\varphi(1/\delta)) m(S^{\lambda} f) > \lambda \quad (f \in F).
\]

**Proof.** Fix \( f \in F, \delta, \) and \( \lambda \). We find some \( g \in F \) satisfying \( \|Tg\|_{\infty} \leq \delta \lambda, \)
\[
\frac{1}{c} \chi_{(Tf \leq \delta \lambda)} S_f \leq S^\lambda g \leq c S^{\lambda} f \quad \text{a.s.}
\]
and
\[
\left\| \left( \frac{S^\lambda g}{c \delta \lambda} \right)_{k=0}^n \right\|_{BMO_\varphi} \leq \left\| Tg \right\|_{\infty} \leq 1.
\]

**Theorem 4.6(b) gives**, for \( \lambda' := 1/(c\delta) \) and \( \mu' := 1/\delta \),
\[
m(S^{\lambda'} f) > (c + c^2)\lambda', \quad T f \leq \delta \lambda'
\]
\[
\leq m(S^{\lambda'} g) > (1 + c)\lambda = m \left( \frac{S^\lambda g}{c \delta \lambda} \right) > \lambda' + \mu'
\]
\[
\leq e^1(\varphi(1/\delta))m(S^{\lambda'} g) > \lambda' = e^1(\varphi(1/\delta)) m(S^{\lambda'} g) > \lambda
\]
\[
\leq e^1(\varphi(1/\delta)) m(S^{\lambda'} f) > \lambda. \quad \blacksquare
\]

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

**Lemma 6.2.** Assume that \( A_i, B_i : \mathcal{A}^\mathbb{N}(\Omega_i, \mathcal{F}_i, \mathbb{P}_i) \to L_0^\infty(\Omega_i, \mathcal{F}_i, \mathbb{P}_i) \) satisfy (EP) with constant \( c \geq 1 \), \( \{M_i, \Sigma_i, m_i\} := \bigotimes_{i=1}^N \{\Omega_i, \mathcal{F}_i, \mathbb{P}_i\} \), and \( \mathcal{L}_{\infty}(\{X_i\}_{i=1}^N) \) be given by
\[
(X^{(1)}, \ldots, X^{(N)}) := ((d_k^{i(1)}, \ldots, d_k^{i(n)})_{k=0}^n)_{i=0}^N \quad \text{in } F, \quad f^{(i)} := (d_k^{i(n)})_{k=0}^N \in E_i.
\]
Define \( S, T : F : \to L_0^\infty(M, \Sigma, m) \) by
\[
S(f^{(1)}, \ldots, f^{(N)}) := \sup_i A_i f^{(i)} \quad \text{and} \quad T(f^{(1)}, \ldots, f^{(N)}) := \sup_i B_i f^{(i)}.
\]
Then \( (F, S, T) \) has (EP) with constant \( c \geq 1 \).

**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
\[
\frac{1}{c} \chi_{(f^{(i)} \leq \delta \lambda)} A_i f^{(i)} \leq A_i g^{(i)} \leq c A_i f^{(i)} \quad \text{a.s.} \quad \text{and} \quad B_i g^{(i)} \leq c \lambda \quad \text{a.s.}
\]
For \( g := (g^{(1)}, \ldots, g^{(N)}) \in F \) one gets a.s. \( S^\lambda g = \sup_i A_i g^{(i)} \leq c \sup_i A_i f^{(i)} = c S^\lambda f, \quad T g = \sup_i B_i g^{(i)} \leq c \lambda, \) and
\[
\frac{1}{c} \chi_{(Tf \leq \delta \lambda)} S_f \leq \frac{1}{c} \sup_{1 \leq i \leq N} \chi_{(f^{(i)} \leq \delta \lambda)} A_i f^{(i)} \leq S^\lambda g. \quad \blacksquare
\]

**Proof of Theorem 1.7.** (1) Let \( a := (1/k^\psi)^{\infty}_{k=0} = (\varphi(1/h))^{\infty}_{h=1} \) and \( c := 6\psi^{-1}(3)U_a(BMO_\psi) \). Using Lemma 6.2 for \( \{\Omega_i, \mathcal{F}_i, \mathbb{P}_i\} := \{\Omega, \mathcal{F}, \mathbb{P}\}, \quad E_i := \mathbb{E}, \quad A_i := \varphi(1/h) A_i, \quad \text{and} \quad B_i := c'B_i \)
implies that \( (F, S, T) \) defined as in Lemma 6.2 has (EP) with \( c \geq 1 \). For \( f = (f^{(1)}, \ldots, f^{(N)}) \in F \) and \( h_k^{i(n)}(\omega_1, \ldots, \omega_N) := A((f^{(i)}(k))^{(n)}(\omega_i) \in L_0(M), \)

\[
\text{Theorem 4.6(b) gives, for } \lambda' := 1/(c\delta) \text{ and } \mu' := 1/\delta,
\]
\[
m(S^{\lambda'} f) > (c + c^2)\lambda, \quad T f \leq \delta \lambda
\]
\[
\leq m(S^{\lambda'} g) > (1 + c)\lambda = m \left( \frac{S^\lambda g}{c \delta \lambda} \right) > \lambda' + \mu'
\]
\[
\leq e^1(\varphi(1/\delta))m(S^{\lambda'} g) > \lambda' = e^1(\varphi(1/\delta)) m(S^{\lambda'} g) > \lambda
\]
\[
\leq e^1(\varphi(1/\delta)) m(S^{\lambda'} f) > \lambda. \quad \blacksquare
\]
from Theorems 4.6 and 1.5 we get

\[
\frac{1}{6\psi^{-1}(3)}(\|(Sf^k)\|_{H^0}) \leq \|(Sf^k)\|_{BMO_\phi} \leq \sup_{1 \leq i \leq N} \|(h_i^{(i)})\|_{BMO_\phi}.
\]

Applying Remark 4.14(2) gives

\[
\|(h_i^{(i)})\|_{BMO_\phi} = \|(A(f^{(i)}))\|_{BMO_\phi}.
\]

so that our assumption yields

\[
\|(Sf^k)\|_{BMO_\phi} \leq c \sup_{1 \leq i \leq N} \|(A(f^{(i)}))\|_{BMO_\phi},
\]

\[
\leq c \sup_{1 \leq i \leq N} \|Bf^{(i)}\|_{BMO_\phi} = \|Tf\|_{BMO_\phi}.
\]

Lemma 6.1 implies

\[
m(cS^*f > c + \varepsilon^2, Tf \leq \delta) \leq e^{1 - \psi(1/\delta)}m(cS^*f > \lambda)
\]

for \(0 < \delta \leq 1 \) and \(\lambda > 0\). We choose \(0 < \delta \leq 1 \) with \(e^{1 - \psi(1/\delta)} = \frac{2(c + \varepsilon^2)}{\delta} \) and deduce for \(c = c(c + \varepsilon^2) \) and \(c = c(c + \varepsilon^2) \) that

\[
\frac{\beta}{\delta} \cdot \frac{1}{\sqrt{1 - \beta \varepsilon}} \leq 2(c(c + \varepsilon^2)) = 2(c(c + \varepsilon^2)) \cdot \frac{1}{\log(2(c(c + \varepsilon^2)))}.
\]

Lemma 4.4(2) applied to \(\eta = ac_\phi\), where \(a > 1\), yields \(\psi^{-1}(ap) \leq \psi_{ac_\phi}^{-1}(p)\)

for \(p \geq 1\). Hence Lemma 4.5 gives for some \(d > 0\), depending on \(c\) and \(\psi\), only

\[
\|
\begin{array}{l}
\sup_{1 \leq i \leq N} \varphi((1/i)A^*f(\omega_i)) < \|cS^*f\|_p \leq \frac{\beta}{\delta} \cdot \frac{1}{\sqrt{1 - \beta \varepsilon}} \|Tf\|_p \\
\leq \frac{d \psi^{-1}(p)}{\psi^{-1}(p)} \|Tf\|_p \\
= \frac{d \psi^{-1}(p) \psi^{-1}(p)}{\psi^{-1}(p)} \sup_{1 \leq i \leq N} \|Bf(\omega_i)\|_{L_p(\mathbb{R}^N)}.
\end{array}
\]

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

\[w := \sum_{i=1}^{N} \varphi((1/i)X_{((t-1)/N,i/N)}) \in L_0^0([0,1]) \text{ so that } w \varphi \leq \Delta(\varphi)w \text{ a.s.}
\]

Then

\[
K_{\infty,0}^{e}(A^*f, t^{1/p}) \leq \Delta(\varphi)K_{\infty,0}^{e}(A^*f, t^{1/p}) \leq 2^{1/p} \Delta(\varphi)K_{\infty,0}^{e}(A^*f, t^{1/p})
\]

Exploiting Theorem 1.8 we can finish with

\[
K_{\infty,0}^{e}(A^*f, t^{1/p}) \leq 2^{1/p} \Delta(\varphi)K_{\infty,0}^{e}(A^*f, t^{1/p}) \leq 2^{1/p} \Delta(\varphi)K_{\infty,0}^{e}(A^*f, t^{1/p})
\]

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

COROLLARY 6.3. Let \(A, B : \mathcal{A}(X)(\mathcal{F}_\phi)^{\lambda=0} \rightarrow E \rightarrow L_0^0(\mathbb{R}, \mathcal{F}, \mathbb{P}) \) where \(X \) is a Banach space, and let \(\psi(1) = t\). Assume that \((E, A, B)\) satisfies \((BP)\) with constant \(c > 1\) and that

\[
\sup_{f} \sup_{0 \leq k \leq n} \mathbb{P}(\|A^*f - A^*f^k\|_p > \|Bf\|_p) = s < 1/2.
\]

Then there is some \(d > 0\), depending on \(s\) and \(c\) only, such that for all \(1 \leq p < \infty\), \(\partial \geq 1\), and \(f \in E\),

(29)

\[
K_{\infty,0}^{e}(A^*f, t^{1/p}) \leq dpK_{\infty,0}^{e}(Bf, t^{1/p})
\]

In particular, for \(t = 1\) one has \(\|A^*f\|_p \leq dp\|Bf\|_p\).

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

\[
\lambda \left(\sup_{1 \leq i \leq N} \left| \sum_{i=1}^{N} d_i \right| > \|S_2f\|_p \right)^{1/p} \leq c\sqrt{p}.
\]

Given \(\lambda = c = a \geq 1\) we choose \(p\) such that \(1/a = c\sqrt{p}/\lambda\) and obtain

(30)

\[
\mathbb{P}(\|S_2f\|_{(a/\lambda)^2} \geq \|S_2f\|_{(a/\lambda)^2} > \|S_2f\|_{(a/\lambda)^2}) \leq e^{-\lambda/a^2}.
\]

Since

\[
\|S_2f\|_{(a/\lambda)^2} \geq \sup_t (e^{(a/\lambda)^2} (S_2f)(t)) \geq \sup_t \left( e^{(1-\lambda)^2/\lambda^2} (S_2f)(t) \right) \geq e^{-a^2} (S_2f)(e^{-1-\lambda^2})
\]

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

\[
\mathbb{P}\left( \sup_{1 \leq i \leq N} \left| \sum_{i=1}^{N} d_i \right| > c\lambda K_{\infty,0}^{e}(S_2f, e^{-1-\lambda^2}) \right) \leq e^{1-\lambda^2}.
\]
(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),
\[
K_{\infty, p}^p(A^{*} f, t^{1/p}) \geq (A^{*} f \cdot w_t)_p(t) \geq \|A^{*} f\|_p.
\]
Now, consider \( 1 \leq p < q < \infty \), \( \psi(t) = \tilde{\psi}(t) = t^p \) (see Example 4.3), and \( \Omega = [0, 1] \), where we use the dyadic filtration \( (F_k^p)_{k=0}^{\infty} \) from Section 4.
Let \( A : E_n \to L_0^p(\Omega, \mathcal{F}, \mathbb{P}) \), where \( E_n \) is the set of all mean-zero martingale differences with respect to \( (F_k^p)_{k=0}^{\infty} \) and \( A((d_k^p)_k)_k := (\sum_{k=1}^{\infty} d_k^p)_k \). If \( f_n := (d_k^p)_k \in E_n \) is the sequence of martingale differences generated by
\[
f(t) := \sqrt{1 + \log(t)} - \sqrt{1 + \log(1/t)} \|1 \|_{L_1(0, 1]}, \text{ then Theorems } 4.12 \text{ and } 4.6 \text{ imply, for } h = \sup_{n} (A^{*} f)_n,
\]
\[
\tilde{h}(t) := c \sqrt{1 + \log(1/t)} \text{ for } 0 < t < 1.
\]
Now one can show \( K_{\infty, p}[\sup_{n} A^{*} f_n, t^{1/p}] \leq c \sqrt{1 + \log(t)} \|f\|_p \) for \( 1 \leq t < \infty \) so that
\[
\lim_{t \to \infty} K_{\infty, p}[\sup_{n} A^{*} f_n, t^{1/p}] / \tilde{\psi}(1 + \log(t)) = 0
\]
but
\[
\inf_{t \geq 1} K_{\infty, p}[A^{*} f_n, t^{1/p}] \geq \|A^{*} f_1\|_p > 0.
\]

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 \((C = 0)\) imply the equivalence of the K-functional \( K_{\infty, p}^{\psi}(B, f, t) \) and \( K(f, t, H_{\infty, p}^{\psi}, H_{p}^{\psi}) \), the K-functional with respect to the Hardy spaces generated by \( B \). This observation can be found with more restrictive assumptions in [28] and is prepared in Lemma 7.2 for our purposes.

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

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

**Lemma 7.1.** Let \( T : \mathcal{A}^X((F_k^p)_{k=1}^{\infty}) \to E \to L_0^p(\Omega, \mathcal{F}, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume \( T \) to be quasilinear with constant \( \gamma_1 \geq 1 \) and local. Then for every stopping time \( \tau, f \in E \), and \( k \in \{0, \ldots, n\} \) one has a.s.
\[
T(f^\tau) \leq \gamma_1^k T(f^k) \text{ on } \{\tau = k\} \quad \text{and} \quad T(f^\tau) \leq 2\gamma_1^k T(f^k) \chi_{\{\tau < n\}}.
\]

**Proof.** On \( \{\tau = k\} \) we get a.s.
\[
T(f^\tau) \leq \gamma_1^k T(f^\tau - f^{\tau \wedge k}) + T(f^{\tau \wedge k}) = \gamma_1^k T(f^{\tau \wedge k}) \leq \gamma_1^k [T(f^{\tau \wedge k} - f^k) + T(f^k)] = \gamma_1^k T(f^k).
\]
For the second inequality we remark that \( T(f^\tau) = 0 \) a.s. on \( \{\tau = n\} \). For \( 0 \leq k < n \) from the first step we obtain \( T(f^\tau) \leq \gamma_1 T(f^\tau + T(f^\tau)) \leq \gamma_1 [T(f^\tau + \gamma_1^k T(f^k))] \leq 2\gamma_1^k T(f^k) \) a.s. on \( \{\tau = k\} \).

The next lemma is proved for \( 0 < p \leq 1 \) and for \( \sigma \)-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-functional, now available for all \( 0 < p < \infty \) and for quasilinear operators.

**Lemma 7.2.** Let \( T : \mathcal{A}^X((F_k^p)_{k=1}^{\infty}) \to E \to L_0^p(\Omega, \mathcal{F}, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume \( T \) to be quasilinear with constant \( \gamma_1 \geq 1 \), local, and predictable. Let \( \lambda \geq 0, f \in E \), and let us define \( \tau := \inf\{k \mid T(f^k) > \lambda\} \wedge n \). Then for all \( 0 < p < \infty \) one has
\[
T(f^\tau) \leq \gamma_1^\tau \lambda \text{ a.s. and } \|T(f^\tau)\|_p \leq 2\gamma_1^\tau \left( \int_{\{T(f^\tau) > \lambda\}} (T(f^\tau))^p \text{d}\mathbb{P} \right)^{1/p}.
\]

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

**Proposition 7.3.** Let \( A, B, C : \mathcal{A}^X((F_k^p)_{k=1}^{\infty}) \to E \to L_0^p(\Omega, \mathcal{F}, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume that \( A \) is quasilinear with constant \( \gamma_1 \geq 1 \), local, measurable, that \( B \) is quasilinear with constant \( \gamma_2 \geq 1 \), local, predictable, and monotone, and that \( C f = c(f) \chi_{\Omega} \) where \( c(f) \in \mathbb{R} \) with \( c(0) = 0 \) and \( c(f^+) \leq \|c(f)\| \) for all stopping times \( \tau \). Then \( \{E, A, B + C\} \) has (EP) with constant \( c = \max(\gamma_1^2, \gamma_2^2) + 1 \).

**Proof.** Let \( \lambda > 0 \) and \( f \in E \). If \( c(f) > \lambda \), then we choose \( g = 0 \) such that
\[
\chi_{\{B + C \leq \lambda\}} A^{*} g = A^{*} g = 0 \leq A^{*} f \text{ a.s. and } (B + C)g = 0 \text{ a.s.}\]
If \( c(f) \leq \lambda \), then we choose \( g := f^* \in E \) where \( r := \inf \{ k \mid B(f^{k+1}) > \lambda \} \) \& \( n \). Lemma 7.2 gives \( Bg = B(f^*) \leq \frac{\lambda}{\gamma_A} \) so that \( (B + C)g \leq Bg + Cf \leq \left( \frac{\lambda}{\gamma_A} + 1 \right) \lambda \) a.s. on \( \Omega \). Since \( \{ Bf \leq \lambda \} = \{ \tau = n \} \) we have

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

Finally, the left-hand side of the assertion of Lemma 7.1 implies \( A^*(f^*) \leq \frac{\lambda}{\gamma_A} A^*(f) \).

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

**Corollary 7.4.** Let \( A, B : A^\infty((F_k)_{k=0}^\infty) \supseteq E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume that \( A \) is quasinilvar with constant \( \gamma_A \geq 1 \), local, and measurable, and that \( B \) is quasinilvar with constant \( \gamma_B \geq 1 \), local, predictable, and monotone. Let \( \psi \in D \) and suppose that

\[ c(f) := \sup \{ \| (A^k \psi^*)^T \|_{k=0} \|BMO_\psi \| - \|B(f^*)\|_{\infty} \} \geq 0, \]

with the supremum taken over all stopping times \( \tau \), is finite for all \( f \in E \). Then there is some \( d > 0 \), depending on \( \psi \in D \) and \( \gamma_A, \gamma_B \geq 1 \) only, such that for \( 1 \leq p < \infty \), \( \tau \geq 1 \), \( w_1 := w_1^\psi \), and \( f \in E \),

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

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

**Definition 7.5.** Let \( T : A^\infty((F_k)_{k=0}^\infty) \supseteq E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \), \( t \geq 0 \), \( f \in E \), and \( 1 \leq p, q \leq \infty \).

\[
K(f, t, H^T, H^T, H^T) := \inf \{ \|Tg\|_q + t\|Th\|_p \mid g \in E, f = g + h \},
\]

\[
K(f, t, BM\psi^\psi, H^T, H^T) := \inf \{ \|Tg\|_q + t\|Th\|_p \mid g \in E, f = g + h \}.
\]

**Proposition 7.6.** Let \( T : A^\infty((F_k)_{k=0}^\infty) \supseteq E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume \( T \) to be quasinilvar with constant \( \gamma_T \geq 1 \), local, predictable, and monotone. Then for all \( f \in E \) and \( 1 \leq t, p < \infty \) one has

\[
\frac{1}{\gamma_T^2} K(f, t, H^\infty, H^\infty, H^\infty, H^\infty) \leq K_{\psi, \psi}[Tf, t] \leq \gamma_T K(f, t, H^\infty, H^\infty, H^\infty, H^\infty).
\]

**Proof.** On the one hand, \( Tf \leq \gamma_T[Tg + Th] \) a.s. gives \( K_{\psi, \psi}[Tf, t] \leq \gamma_T K(f, t, H^\infty, H^\infty, H^\infty) \). On the other hand, for \( \lambda := \gamma_T(1/\lambda) \) Lemma 7.2 says

\[
\|T(f^*)\|_\infty \leq \frac{1}{\gamma_T^2} \gamma_T(Tf^*)^{1/\lambda} \leq \frac{1}{\gamma_T^2} (Tf^*)^{1/\lambda}.
\]

and

\[
d\|T(f^*)\|_p \leq 2\gamma_T^2 (Tf^*)^{1/\lambda} \]

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 \( A, B : A^\infty((F_k)_{k=0}^\infty) \supseteq E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \), where \( E \) is closed under starting and stopping. Assume that \( A \) is quasinilvar with constant \( \gamma_A \geq 1 \), local, and measurable, and that \( B \) is quasinilvar with constant \( \gamma_B \geq 1 \), local, predictable, and monotone. Let \( \psi \in D \) and suppose

\[
\| (A^k \psi)^T \|_{k=0} \leq \| Bf \|_{\infty} \quad \text{for all } f \in E.
\]

Then there is some \( d > 0 \), depending on \( \psi, \gamma_A, \gamma_B \) only, such that for \( 1 \leq p < \infty, t \geq 1, f \in E \),

\[
K(f, t; B, BM\psi^\psi, H^T, H^T) \leq d\psi^{-1}(p)K_{\psi, \psi}[Bf, t].
\]

**Proof.** Let \( f \in E, 1 \leq p < \infty, t > 1 \). According to Proposition 7.6 we find \( g, h \in E \) satisfying \( f = g + h \) and \( \|Bg\|_{\infty} + t\|Bh\|_p \leq 4\gamma_B K_{\psi, \psi}[Bf, t] \).

Applying Theorem 1.7 to \( h \in E \) yields

\[
\| (A^k g)^T \|_{k=0} \leq \| BMO_\psi + t\|Ah\|_p \leq \| Bg \|_{\infty} + td\psi^{-1}(p)\|Bh\|_p \leq (d + 1)\psi^{-1}(p)4\gamma_B K_{\psi, \psi}[Bf, t].
\]

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 \( X \otimes_{\infty} \mathbb{R} \). This is done with the help of

**Definition 7.8.** Let \( B : A^\infty((F_k)_{k=0}^\infty) \supseteq E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \), and \( 0 < p \leq 1 \). Then \( B_p : E \rightarrow L_0^1(\Omega, F, \mathbb{P}) \) is defined for \( f = (d_k)_{k=0}^\infty \) by

\[
B_p(f)(\omega) := \inf \left\{ \left( \sum_{l=1}^L |B(d^{(l-1)-d(l+1)})(\omega) |^p \right)^{1/p} \mid 0 = n_0 < n_1 < \ldots < n_L = n \right\}
\]

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 \( f : 2^M \rightarrow [0, \infty) \) be a function satisfying \( f(\emptyset) = 0 \) and

\[
f(C_1 \cup C_2) \leq \gamma[f(C_1) + f(C_2)] \quad \text{for all disjoint } C_1, C_2 \subseteq 2^M,
\]
where \( \gamma \geq 1 \). Defining \( 0 < p \leq 1 \) via \( 2^{1/p-1} = \gamma \) one has

\[
(f\left( \bigcup_{i=1}^{L} C_i \right))^p \leq \sum_{i=1}^{L} (f(C_i))^p \quad \text{for all pairwise disjoint } C_1, \ldots, C_L \subset 2^M.
\]

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

1. First we show that \( (f(C_i))^p \leq 2^{-k_i} s \) for \( i = 1, \ldots, L \), \( k_i \in \{0,1,\ldots\} \), and \( \sum_{i=1}^{L} 2^{-k_i} \leq 1 \) imply \( (f\left( \bigcup_{i=1}^{L} C_i \right))^p \leq s \) whenever the \( C_i \) 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 \( \sum_{i=1}^{L} 2^{-k_i} = 1 \). 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 \( h \geq 0 \). If \( I := \{ i \mid k_i = h+1 \} \) we obtain a set of even cardinality since \( \sum_{i=1}^{L} 2^{-k_i} = 1 \). If \( i, j \in I \) then one gets

\[
(f(C_i \cup C_j))^p \leq \gamma^p (f(C_i) + f(C_j))^p \leq 2^{-k_i} s
\]

so that we can write \( \bigcup_{i=1}^{L} C_i \) as disjoint union of \( D_n (n = 1, \ldots, N) \) with \( (f(D_n))^p \leq 2^{-m_n} s \) where \( \max(m_1, \ldots, m_N) \leq h \) 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 on \( h \).

2. To prove the statement of our lemma we can assume that \( f(C) > 0 \) for all \( C \neq \emptyset \) (otherwise we choose \( f_{C,0} := f(C) + \varepsilon \) if \( C \neq \emptyset \) and \( f_{\emptyset} := 0 \) with \( \varepsilon > 0 \)). Let \( C_1 = \bigcup_{i=1}^{L} C_i \) be a disjoint union of non-empty sets. If \( f(C_1))^p \leq \sum_{i=1}^{L} (f(C_i))^p \) and choose \( k_1 \in \{1,2,\ldots\} \) with \( 2^{-k_1} s \leq (f(C_i))^p \leq 2^{-k_i} s \), then we get \( \sum_{i=1}^{L} 2^{-k_i} \leq 1 \) and \( (f(C_i))^p \leq s \) from step (1), which proves our assertion.

**Lemma 7.10.** Let \( B : AX((F_k)^p_{k=0}) \to G \to L_0^0(\Omega, F, |p|) \) be a quasinorm operator with constant \( \gamma_B \geq 1 \). Then for \( 0 < p \leq 1 \) with \( 2^{1/p-1} = \gamma_B \) one has:

1. \( B_B f \leq B f \leq 2^{1/p} B_B f \) and \( B_B^{(k-1)f^k} = B^{(k-1)f^k} \) for all \( f \in E \) and \( 1 \leq k \leq n \).
2. \( [B_B(f^p)]^p \leq [B_B(f^p)]^p + [B_B(f-f^p)]^p \) for all \( f \in E \) and \( 0 \leq k \leq n \).
3. If \( B \) is local (measurable, predictable, monotone) then so is \( B_B \).

**Proof.** Most of the things are evident (for example, \( B_B f \leq B f \) follows simply from \( d_0 = 0 \) for \( (d_k^p_{k=0}) \in E \)). The point is \( B f \leq 2^{1/p} B_B f \), which we get from Lemma 7.9.

**Proposition 7.11.** Let \( f = (d_k^p_{k=0}) \in AX((F_k)^p_{k=0}) \) with \( d_0 = 0 \). Assume that

\[
A, B : \{\pm f^k \mid \sigma, \tau \text{ stopping times} \} \to E \to L_0^0(\Omega, F, |p|)
\]

are quasinlinear with constants \( \gamma_A, \gamma_B \geq 1 \), local, and measurable, and assume that \( B \) is monotone. Suppose that \( B^{(k-1)f^k} \leq u_B \) a.s. for \( k = 1, \ldots, n \), where \( (u_k^p_{k=0}) \in AX((F_k)^p_{k=0}) \) is a fixed predictable sequence and \( u_0 = 0 \). If \( F \subset AX((F_k)^p_{k=0}) \) and \( S, T : F \to L_0^0(\Omega, F, |p|) \) are defined by

\[
F := \{(g_k, w_k)^p_{k=0} \mid (g_k, w_k)^p_{k=0} = \pm (\chi_{\sigma < k \leq \tau})(d_k, w_k)^p_{k=0}, \quad \sigma, \tau \text{ stopping times} \}
\]

\[
S((g_k, w_k)^p_{k=0}) := A((g_k)^p_{k=0}),
\]

\[
T((g_k, w_k)^p_{k=0}) := \{ \sup_{0 \leq k \leq n} |w_k| \} \vee B((g_k)^p_{k=0}),
\]

then \((F, S, T)\) has (EP) with some constant \( c > 0 \) depending on \( \gamma_A \) and \( \gamma_B \) only.

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

\[
B^{(k-1)f^k} \leq \gamma_B B^{(k-1)f^k} + B^{(k-1)f^k} = \gamma_B B^{(k-1)f^k}
\]

so that \( B^{(k-1)f^k} \leq \gamma_B B^{(k-1)f^k} \) a.s. on \( \Omega \).

(2) We define \( U : F \to L_0^0(\Omega, F, |p|) \) by

\[
U((g_k, w_k)^p_{k=0}) := \sup_{0 \leq k \leq n} \{|B_B((g_k \chi_{\sigma < k \leq \tau})^p_{k=0})|^p + \gamma_B^2 |w_k + 1|^p \}/p,
\]

where \( u_{n+1} \equiv 0 \). If \( 0 < p \leq 1 \) is chosen so that \( 2^{1/p-1} = \gamma_B \), and \( B_B \) is explained in Definition 7.8. It is clear that \( S \) and \( U \) are quasinormable and local, and

\[
2^{-1/p} T((g_k, w_k)^p_{k=0}) \leq U((g_k, w_k)^p_{k=0}) \leq \gamma_B^2 2^{1/p} T((g_k, w_k)^p_{k=0}).
\]

\( S \) is measurable and \( U \) is monotone. The point is that \( U \) is predictable since for \( 1 \leq k \leq n \),

\[
B_B^{(k-1)f^k} \leq B_B^{(k-1)f^k} + B_B^{(k-1)f^k} \leq \gamma_B^2 B_B^{(k-1)f^k}
\]

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 \( \psi \in D, f = (d_k^p_{k=0}) \in AX((F_k)^p_{k=0}) \) with \( d_0 = 0 \), and

\[
A, B : \{\pm f^k \mid \sigma, \tau \text{ stopping times} \} \to E \to L_0^0(\Omega, F, |p|)
\]

be quasinlinear operators with constants \( \gamma_A, \gamma_B \geq 1 \), local, and measurable, and let \( B \) be monotone. Suppose that \( B^{(k-1)f^k} \leq u_B \) a.s. for \( k = 1, 2, \ldots, \).
where \((v_k)_{k=0}^n \in \mathcal{A}(\mathcal{F}_k)_{k=0}^n\) is a fixed predictable sequence. Assume that
\[
\|A(f)\|_{\text{BMOM}_\nu} \leq \sup_{\sigma < k \leq T} \|v_k \vee B(f)\|_\infty
\]
for all stopping times \(\sigma\) and \(T\). Then there exists \(d > 0\), depending on \(\psi \in \mathcal{D}, \gamma_A\), and \(\gamma_B\) only, such that for \(1 \leq p < \infty, t \geq 1, \) and \(w_t := w_t^\psi,\)
\[
K_{\psi, p}[A^t, B^t] \leq d\psi^{-1}(p)K_{\psi, p}[\sup v_k \vee B(f), t, t/p].
\]
(32)

In particular, for \(t = 1\) one has \(\|A^t\|_p \leq \psi^{-1}(p)\sup v_k \vee B(f).\)

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 \(\mathcal{D} = \mathbb{D}_n = \{e_1, \ldots, e_n \mid e_1, \ldots, e_n = \pm 1\}\) is the Cantor group equipped with the normalized Haar measure \(\mu_N\). As filtration \((\mathcal{F}_k)_{k=0}^n\) we use \(\mathcal{F}_0 := \{\emptyset, \mathbb{D}_n\}\), and for \(1 \leq k \leq n,\)
\[
\mathcal{F}_k := \{\{e_1, \ldots, e_n \mid (e_{k+1}, \ldots, e_n = \pm 1) \mid e_1 = \pm 1, \ldots, e_k = \pm 1\}\}
\]}

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

**Definition 8.1.** An operator \(A : M_{\mathcal{X}}^{\mathcal{Y}} \to L_1^{\mathcal{Y}}(\mathbb{D}_n)\) is called self-similar provided that for all \(0 < k < n, \) all atoms \(D \in \mathcal{F}_k\), and all \(f = (d_k)_{k=0}^n \in M_{\mathcal{X}}^{\mathcal{Y}}\) with \(\text{supp}(d_k) \subseteq D\) for all \(l = 0, \ldots, n\) and \(d_0 = \ldots = d_k = 0\) one has
\[
A(f)_{f_k, \ldots, f_k} = (Af)_{f_k, \ldots, f_k} = (\theta_1, \ldots, \theta_k) \in \{-1, 1\},
\]
with
\[
f(\theta_1, \ldots, \theta_k) := (d_{\theta_1, \ldots, f_k})_{k=0}^n,\]
and
\[
h_{f_k, \ldots, f_k}(e_1, \ldots, e_n) := h(e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)
\]
for some \(h \in L_1^{\mathcal{Y}}(\mathbb{D}_n).\)

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

Proposition 8.2. Let \(A, B : M_{\mathcal{X}}^{\mathcal{Y}} \to L_1^{\mathcal{Y}}(\mathbb{D}_n)\) be self-similar, sublinear, and local, where \(A\) is assumed to be measurable and \(B\) is assumed to be monotone.

(1) If \(\varphi \in C_0,\) then \(|Af|_{M_0^\varphi}(\mathbb{D}_n) \leq \|Bf\|_\infty\) for all \(f \in M_{\mathcal{X}}^{\mathcal{Y}}\) implies
\[
\|Bf\|_{\text{BMOM}_\varphi} \leq 2\Delta(\varphi)\|Bf\|_\infty\] for \(f \in M_{\mathcal{X}}^{\mathcal{Y}}.\)

(2) If \(0 < s < 1/4,\) then \(\mu_n(Af) > Bf\|_\infty\) for all \(f \in M_{\mathcal{X}}^{\mathcal{Y}}\) implies
\[
\|Bf\|_{\text{BMOM}_\varphi} \leq 6\max\left(1, \frac{1}{\log(1/(4s))}\right)\|Bf\|_\infty\] for \(f \in M_{\mathcal{X}}^{\mathcal{Y}}.\)

**Proof.** (1) Let \(0 < k < n\) and assume \(g = (d_k)^{\mathcal{Y}}_{k=0} \in M_{\mathcal{X}}^{\mathcal{Y}}\) and \(h = (d_h)^{\mathcal{Y}}_{h=0} \in M_{\mathcal{X}}^{\mathcal{Y}}\) where \(C, D \in \mathcal{F}_k\) are disjoint and \(d_0 = \ldots = d_k = 0.\) If the operator \(T : M_{\mathcal{X}}^{\mathcal{Y}} \to L_1^{\mathcal{Y}}(\mathbb{D}_n)\) 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 \(f = (d_k)^{\mathcal{Y}}_{k=0} \in M_{\mathcal{X}}^{\mathcal{Y}}\) with \(d_0 = \ldots = d_k = 0\) we get by induction
\[
T((f - f^{k-1})\chi_{\mathcal{D}}) = \sum_{\theta_1, \ldots, \theta_k = \pm 1} (Tf)(\theta_1, \ldots, \theta_k).
\]

(2) Let \(f = (d_k)_{k=0}^n \in M_{\mathcal{X}}^{\mathcal{Y}} - 1 < k \leq n,\) and \(D \in \mathcal{F}_k\) be an atom. Let \(\tilde{D} \supseteq D\) be the dyadic predecessor, \(\tilde{D} \in \mathcal{F}_{k-1}.\) It follows for \(\lambda \leq 0\) and \(g = \sum_{\theta_1, \ldots, \theta_k = \pm 1} (f(\theta) - f^{k-1})\chi_{\mathcal{D}}(\theta_1, \ldots, \theta_k)\)

that
\[
(\mu_{\lambda})(|f - f^{k-1}|) \leq \max(\mu_{\lambda})(f - f^{k-1}) \leq 2\max(\mu_{\lambda})(f - f^{k-1}) \chi_{\tilde{D}} + \lambda) \leq 2\mu_{\lambda}(Ag > \lambda)\]

(3) To prove the first assertion of our proposition we derive, under the assumptions of step (2),
\[
|f - f^{k-1}|_{M_0^\varphi(D, \mu)} \leq \Delta(\varphi)\|Ag\|_{M_0^\varphi(D)} \leq \|B(f - f^{k-1})\|_{\text{L}_m}(\tilde{D}) = 2\Delta(\varphi)\|Bf\|_{\text{L}_m}(\tilde{D}).
\]

The case \(k = 0, 1\) leads trivially to the same estimate. Finally, letting \(D = \cup D_i\) be a disjoint union of atoms \(D_i \in \mathcal{F}_k (k \geq 1)\) one gets
\[
|f - f^{k-1}|_{M_0^\varphi(D, \mu)} \leq \sup|f - f^{k-1}|_{M_0^\varphi(D_i, \mu)} \leq 2\Delta(\varphi)\|Bf\|_{\infty}.
\]

(4) To prove the second assertion we consider the estimate
\[
2\|Bf\|_\infty \geq \|B(f - f^{k-1})\|_{\text{L}_m}(\tilde{D}) = \|Bg\|_\infty,\] so that step (2) and our assumption yield
\[
(\mu_{\lambda})(|f - f^{k-1}|) > 2\|Bf\|_\infty \leq 2\max(Ag > \|Bg\|_\infty) \leq 2s.
\]
Similarly to step (3) we get \((\mu_n)_D(\sum |A f^i - A f^{i-k}| > 2 \|B f\|_\infty) \leq 2s\) 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 both assertions of the proposition above with Proposition 7.3 \((C = 0)\) and Theorem 1.7, then we get

**Corollary 8.3.** Let \(A, B : \mathcal{M}_n^X \to L_1(\mathbb{D}_n)\) be self-similar, sublinear, and local, where \(A\) is assumed to be measurable and \(B\) is assumed to be predictable and monotonic. Then for all \(0 < s < 1/4\) there is some \(d_s > 0\), not depending on \(n, A, B, s\), such that

\[
\|A^* f\|_p > d_s p \|B f\|_p
\]

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

\[
\mu_n(A g) > \|B g\|_\infty > s.
\]

**Remark 8.4.** (1) D. L. Burkholder has shown in [9] (Lemma 3.1) that for the UMD-transforms \(A_0, B_0 : \mathcal{M}_n^X \to L_1(\mathbb{D}_n)\) with \(A_0((d_k)_{k=0}^n) := \|\sum_{k=1}^n d_k e_k\|_X\) and \(B_0((d_k)_{k=0}^n) := \|\sum_{k=1}^n d_k e_k\|_X\), where \(d_k \in \{-1, 1\}\) are fixed, the inequality \(A_0^* f > 6 c \|B_0 f\|_1\) a.s. for some \(f \in \mathcal{M}_n^X\) and some \(c > 0\) gives some \(g \in \mathcal{M}_n^X\) with

\[
\mu_n(A_0^* g) > c \|B_0 g\|_\infty \geq 1/2.
\]

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

\[
B_0 f := \sup_{0 \leq i < n} \left[ \|d_i\|_X + \|d_{i+1}\|_X \right] \quad \text{where} \quad f = (d_i)_{i=0}^n.
\]

\(B_0\) is predictable since for \(1 \leq k \leq n\) one has \(B_0 f^k = \sup_{0 \leq i < k} \left[ \|d_i\|_X + \|d_{i+1}\|_X \right]\) and that \(\|d_k\|\) is \(\mathcal{F}_{k-1}\)-measurable. Moreover, the estimate \((p-1)/(3p)) \|B_0 f\|_p \leq \|B_0 f\|_p \leq \|B_0 f\|_p\) 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 can be found concerning the following martingale-type quantities.

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

\[
\left\| \sum_{k=1}^n T d_k \right\|_{L_\alpha} \leq c \left( \sum_{k=1}^n \|d_k\|_{L_\beta}^p \right)^{1/p}_\beta
\]

and

\[
\mu_n \left( \left\| \sum_{k=1}^n T d_k \right\|_{L_\beta} > d \left( \sum_{k=1}^n \|d_k\|_{L_\beta}^p \right)^{1/p}_\beta \right) \leq s.
\]

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

\[
\sup_n \text{Mtp}_p(n T | L_{p, \infty}, L_p) \leq c_0 \sup_n \text{Mtp}_p(n T | L_2, L_\infty).
\]

The following extends this inequality in two directions.

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

\[
\text{Mtp}_p(n T | L_\alpha, L_\alpha) \leq c \text{Mtp}_p(n T | L_{0, s}, L_\infty) \leq \frac{c}{s} \text{Mtp}_p(T | L_1, L_\infty).
\]

**Proof.** To apply Proposition 8.2(2) we use \(A, B : \mathcal{M}_n^X \to L_1(\mathbb{D}_n)\) with

\[
A((d_k)_{k=0}^n)(\omega) := \left\| T \left( \sum_{k=1}^n d_k(\omega) \right) \right\|_{L_\beta}
\]

and

\[
B((d_k)_{k=0}^n)(\omega) := \alpha \left( \sum_{k=1}^n \|d_k(\omega)\|_{L_\beta}^p \right)^{1/p}_\beta,
\]

where \(\alpha := \text{Mtp}_p(n T | L_{0, s}, L_\infty)\). Then we finish with Theorem 1.7.

**Acknowledgements.** I would like to thank S. Kwapien for calling my attention to the papers of P. Hitzen 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**


Cohomology groups, multipliers and factors in ergodic theory

by

M. LE MAÇNCZYK (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 $\tau : (Y, C, \nu) \to (Y, C, \nu)$ of a Lebesgue space $(Y, C, \nu)$ call any of its invariant $\sigma$-algebras a factor. Denote by

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

the centralizer of $\tau$. Endowed with the weak topology in which

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

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

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

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

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

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

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

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

$$T_\phi(x, S) = (T_\xi, \varphi(x)S),$$

where

$$\varphi : C(\tau) \to C(\mathcal{H})$$

is a $\mathcal{H}$-equivariant *-isomorphism. Thus $T_\phi$ is a cocycle.

[276]