Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Functional Analysis 263 (2012) 129–166

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

A new look at the John–Nirenberg and John–Strömberg theorems for BMO ☆

Michael Cwikel [a], Yoram Sagher [b], Pavel Shvartsman [a,*]

[a] *Department of Mathematics, Technion – Israel Institute of Technology, Haifa 32000, Israel*
[b] *Department of Mathematics, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA*

Received 24 October 2011; accepted 4 April 2012

Available online 14 April 2012

Communicated by G. Schechtman

Abstract

We develop some techniques for studying various versions of the function space *BMO*. Special cases of one of our results give alternative proofs of the celebrated John–Nirenberg inequality and of related inequalities due to John and to Wik. Our approach enables us to pose a simply formulated "geometric" question, for which an affirmative answer would lead to a version of the John–Nirenberg inequality with dimension free constants.

Keywords: BMO; John–Nirenberg inequality; Mean oscillation; John–Strömberg functional; "Dimension free" inequality

1. Introduction. Our main question

We begin by inviting the reader to consider and hopefully even answer the following question. We will subsequently refer to it as "Question A".

☆ The research of the first and third named authors was supported by funding from the Martin and Sima Jelin Chair in Mathematics, by the Technion V.P.R. Fund and by the Fund for Promotion of Research at the Technion.
* Corresponding author.
E-mail addresses: mcwikel@math.technion.ac.il (M. Cwikel), yoram.sagher@gmail.com (Y. Sagher), pshv@math.technion.ac.il (P. Shvartsman).

0022-1236/$ – see front matter
http://dx.doi.org/10.1016/j.jfa.2012.04.003

Do there exist two absolute constants $\tau \in (0, 1/2)$ and $s > 0$ which have the following property?

For every positive integer d and for every cube Q in $\mathbb{R}^d$, whenever E_+ and E_- are two disjoint measurable subsets of Q whose d-dimensional Lebesgue measures satisfy

$$\min\{\lambda(E_+), \lambda(E_-)\} > \tau\lambda(Q \setminus E_+ \setminus E_-),$$

then there exists some cube W contained in Q for which

$$\min\{\lambda(W \cap E_+), \lambda(W \cap E_-)\} \geqslant s\lambda(W).$$

We are led to consider this question because of our interest in the space *BMO* of functions of bounded mean oscillation introduced by John and Nirenberg [13]. We recall that these are the functions $f : \mathbb{R}^d \to \mathbb{R}$ which have the property that

$$\sup_Q \frac{1}{\lambda(Q)} \int_Q |f - f_Q| \, d\lambda < \infty$$

where the supremum is taken over all cubes Q in $\mathbb{R}^d$ and where f_Q is the average of f on Q.

We will show that an affirmative answer to Question A would have very interesting consequences for the study of a remarkable property of functions of bounded mean oscillation. It would imply (see Theorem 9.1) that the following "dimension free" version

$$\lambda\bigl(\bigl\{x \in Q: \bigl|f(x) - m_f\bigr| \geqslant \alpha\bigr\}\bigr) \leqslant \max\left\{\frac{1}{2\tau}, 2\sqrt{\frac{1}{2\tau}}\right\} \cdot \lambda(Q) \cdot \exp\left(-\frac{\alpha s \log \frac{1}{2\tau}}{8\|f\|_{BMO}}\right) \tag{1.1}$$

of the John–Nirenberg inequality [13] holds for every $\alpha \geqslant 0$. It would also imply some slightly stronger inequalities. (Here m_f is any median of the measurable function f on the cube Q in $\mathbb{R}^d$.)

Having formulated our question, let us now state to what extent we have been able, so far, to answer it or to simplify it.

For each particular value of $d \in \mathbb{N}$ we *can* find numbers, $\tau \in (0, 1/2)$ and $s > 0$ which *do* have the property sought in Question A. Furthermore we can show that their having this property, implies that the inequality (1.1) is satisfied.

We do not yet have an answer to Question A, because at least one of our constants τ and s depends on d. We can take, for example, $\tau = \sqrt{2} - 1$, but, for that choice of τ, we have only been able to obtain a value of s which depends on d, namely $s = 2^{-d}(3 - 2\sqrt{2})$.

Regardless of whether τ and s really have to depend on the dimension, it seems of interest that, in the expression of the form $C\lambda(Q) \cdot \exp(-\frac{c\alpha}{\|f\|_{BMO}})$ on the right-hand side of our version (1.1) of the John–Nirenberg inequality, we have revealed a quite explicit connection between the constants C and c and a geometric property expressed by the constants τ and s.

It also seems of interest that the "geometric" condition sought in Question A is, more or less "equivalent" to an analytic condition which compares certain kinds of *BMO* "norms" of functions f on $\mathbb{R}^d$ with related kinds of *BMO* "norms" of their rearrangements f^* on $(0, \infty)$. (The implications, in two opposite directions, which express this "sort of equivalence" are precisely formulated and established in Theorems 8.2 and 8.4.)

Our results can be expressed in more abstract terms, and they apply to other versions of the space *BMO* including the one considered by Wik [30], where cubes are replaced by "false cubes".

A preliminary version of this paper with additional, possibly tedious, explanations and other details, intended mainly for readers less familiar with some of the topics treated here, can be found in [6]. It also contains some slightly different alternative formulations of Question A which may ultimately be of interest.

2. Notation, terminology and some more introduction

Throughout this paper d will denote a positive integer and λ will denote d-dimensional Lebesgue measure on $\mathbb{R}^d$. The value of d will always be clear from the context. When $d = 1$ we will also often use the notation $|E|$ instead of $\lambda(E)$ for each measurable subset E of $\mathbb{R}$. By a *cube* in $\mathbb{R}^d$ we will always mean a d-dimensional closed cube with sides parallel to the axes.

Definition. To save tedious repetitions of terminology, we will say that a set E is *admissible* if it is a measurable subset (i.e., a Lebesgue measurable subset) of $\mathbb{R}^d$ and its d-dimensional Lebesgue measure satisfies $0 < \lambda(E) < \infty$.

For each admissible set E and each measurable real valued function f whose domain of definition contains E, we define the *mean oscillation* of f on E by

$$\mathbf{O}(f, E) := \inf_{c\in\mathbb{R}} \frac{1}{\lambda(E)} \int_E |f - c|\, d\lambda. \tag{2.1}$$

It is convenient to fix some notation for two other frequently used variants of the functional $\mathbf{O}(f, E)$. So we set

$$\mathbf{A}(f, E) := \frac{1}{\lambda(E)} \int_E |f - f_E|\, d\lambda$$

for every function f which is integrable on E, and where $f_E := \frac{1}{\lambda(E)} \int_E f\, d\lambda$. We also set

$$\mathbf{D}(f, E) := \frac{1}{\lambda(E)^2} \iint_{E\times E} \big|f(x) - f(y)\big|\, d\lambda(x)\, d\lambda(y)$$

("**A**" and "**D**" are our abbreviations for "average" and "double integral" respectively). We recall that the set of medians of f on E consists of all numbers $c \in \mathbb{R}$ which satisfy

$$\lambda\big(\big\{x \in E\colon\ f(x) < c\big\}\big) \leqslant \frac{1}{2}\lambda(E) \quad \text{and} \quad \lambda\big(\big\{x \in E\colon\ f(x) > c\big\}\big) \leqslant \frac{1}{2}\lambda(E).$$

This set is always non-empty, and the infimum in (2.1) is attained, i.e.,

$$\mathbf{O}(f, E) = \frac{1}{\lambda(E)} \int_E |f - c|\, d\lambda$$

whenever c is a median of f. We also note that

$$\mathbf{O}(f, E) \leqslant \mathbf{A}(f, E) \leqslant \mathbf{D}(f, E) \leqslant 2\mathbf{O}(f, E) \tag{2.2}$$

for all functions f which are integrable on E. The above standard facts have easy proofs (which are recalled in Appendix 11.1 of [6, pp. 55–57]).

Definition 2.1. Let D be some measurable subset of $\mathbb{R}^d$ with positive measure and let $\mathcal{E}$ be some collection of admissible subsets E of D. We define the space $BMO(D,\mathcal{E})$ to consist of all (equivalence classes of) measurable functions $f : D \to \mathbb{R}$ for which the seminorm

$$\|f\|_{BMO(D,\mathcal{E})} := \sup_{E\in\mathcal{E}} \mathbf{O}(f,E) \tag{2.3}$$

is finite.

One may also define $BMO(D,\mathcal{E})$ equivalently via either one of the seminorms

$$\|f\|^{(\mathbf{A})}_{BMO(D,\mathcal{E})} := \sup_{E\in\mathcal{E}} \mathbf{A}(f,E)$$

or

$$\|f\|^{(\mathbf{D})}_{BMO(D,\mathcal{E})} = \sup_{E\in\mathcal{E}} \mathbf{D}(f,E) \tag{2.4}$$

which (cf. (2.2)) are each equivalent to the seminorm (2.3) to within constants of equivalence 1 and 2.

Of course if f coincides a.e. with a constant function then $\|f\|_{BMO(D,\mathcal{E})} = 0$. The reverse implication may also be true for suitable choices of D and $\mathcal{E}$. In all cases the seminorm $\|\cdot\|_{BMO(D,\mathcal{E})}$ defines a norm on suitable equivalence classes of functions in $BMO(D,\mathcal{E})$ which may, for suitable choices of D and $\mathcal{E}$, be simply equivalence classes of functions modulo constants.

The most frequently considered way of choosing D and $\mathcal{E}$ is

$$\begin{cases} D \text{ is either } \mathbb{R}^d \text{ or some fixed cube in } \mathbb{R}^d \text{ and } \mathcal{E} \text{ is chosen} \\ \text{to be } \mathcal{Q}(D), \text{ the collection of all cubes contained in } D. \end{cases} \tag{2.5}$$

As the reader no doubt recalls, functions of the space $BMO(D,\mathcal{E})$ were first introduced and studied by John and Nirenberg [13] for the case where D is a cube in $\mathbb{R}^d$ and $\mathcal{E} = \mathcal{Q}(D)$. The original motivation for studying these functions apparently came from John's study [11] of problems in the theory of elasticity, related in particular to the concept of elastic strain. One of the first applications of [13] was in a paper [18] by Moser extending Harnack's theorem about harmonic functions to functions which are solutions of elliptic second order PDEs. But the space of these functions and its analogues have since turned out to also have many other deep properties and numerous other, sometimes quite surprising applications in analysis. One particularly notable example of such an application is the connection with H^p spaces revealed in the paper [7] of Fefferman and Stein.

The choice of D and $\mathcal{E}$ specified in (2.5) is only one among several possible interesting choices, and we will list four more examples of such choices now, taking the opportunity to also fix our notation for them, notation which will be used throughout the paper. In each of these examples we will take the set D to either be $\mathbb{R}^d$ or some measurable subset of $\mathbb{R}^d$ with non-empty interior.

$$\begin{cases} \mathcal{E} \text{ is chosen to be } \mathcal{D}(D), \text{ the collection of all dyadic cubes} \\ \text{contained in } D, \end{cases} \tag{2.6}$$

$$\begin{cases} \mathcal{E} \text{ is chosen to be } \mathcal{B}(D), \text{ the collection of all Euclidean balls} \\ \text{contained in } D, \end{cases} \tag{2.7}$$

$$\begin{cases} \mathcal{E} \text{ is chosen to be } \mathcal{K}(D), \text{ the collection of all bounded closed} \\ \text{convex subsets of } D \text{ which have non-empty interiors,} \end{cases} \tag{2.8}$$

$$\begin{cases} \mathcal{E} \text{ is chosen to be } \mathcal{W}(D), \text{ the collection of all \textit{special rectangles}} \\ \text{contained in } D. \end{cases} \tag{2.9}$$

By special rectangles we mean all those subsets of $\mathbb{R}^d$ which are the Cartesian products $I_1 \times I_2 \times \cdots \times I_d$ of d bounded closed intervals of positive length, where, for each $j = 1, 2, \ldots, d$, the length $|I_j|$ of I_j equals either $\min_{k=1,2,\ldots,d} |I_k|$ or $2\min_{k=1,2,\ldots,d} |I_k|$. Such sets, and their associated space $BMO(\mathbb{R}^d, \mathcal{W}(\mathbb{R}^d))$ were introduced and studied by Wik in [30]. He used the terminology "false cubes" for special rectangles. Below we will describe his results in more detail.

Of course the seminorm $\|f\|_{BMO(D,\mathcal{K}(D))}$ is larger than any of the other seminorms $\|f\|_{BMO(D,\mathcal{E})}$ arising from the other choices of $\mathcal{E}$ listed just above, and for this reason it will be of less interest for us here for the particular aims of this paper. However we remark that a result of Nazarov, Sodin and Vol'berg ([19, p. 13] and [20]) shows that every polynomial $P : \mathbb{R}^d \to \mathbb{R}$ of degree n satisfies

$$\big\|\log|P|\big\|_{BMO(D,\mathcal{K}(\mathbb{R}^d))} \leqslant \frac{4+\log 4}{2} n. \tag{2.10}$$

It is remarkable that there is no dependence on the dimension d in this inequality. We are naturally led to ask whether the left side of (2.10) can also be bounded from below by cn for some absolute positive constant c. If this can be shown to be the case, then other results in [19] would imply that a dimension free version of John–Nirenberg inequality holds, for $D = \mathbb{R}^d$ and $\mathcal{E} = \mathcal{K}(\mathbb{R}^d)$, at least for all functions of the special form $\log|P|$. An analogous question with analogous consequences can be asked for the apparently more difficult and perhaps more interesting case where $\mathcal{E}$ is chosen to be $\mathcal{Q}(\mathbb{R}^d)$.

There are also other more "exotic" versions of the space BMO. But these seem to be quite beyond the scope of what we will study in this paper. For example, the measure λ may be replaced by a more general measure, and, furthermore, the underlying set $\mathbb{R}^d$ may be replaced by other suitable sets. There is even a version of BMO in the setting of martingales.

We have already alluded above to the following result of John and Nirenberg, which is (a very slight rewording of) Lemma 1 of [13], and is in fact the main result of that paper:

Theorem 2.2. *Let D be a cube in $\mathbb{R}^d$ and let $f : D \to \mathbb{R}$ be a function belonging to the space $BMO(D, \mathcal{Q}(D))$. Then*

$$\lambda\big(\big\{x \in D\colon \big|f(x) - f_D\big| > \alpha\big\}\big) \leqslant B\lambda(D)\exp\bigg(-\frac{b\alpha}{\|f\|^{(\mathbf{A})}_{BMO(D,\mathcal{Q}(D))}}\bigg) \quad \textit{for every } \alpha > 0, \tag{2.11}$$

where B and b are constants which depend only on the dimension d.

The inequality (2.11), together with various generalizations and variants of it, will be at once our main motivation and our main interest in this paper. In fact (2.11) is the key to obtaining various other properties of *BMO* and has been widely studied further since its original discovery. The proof of (2.11) in [13] uses a famous lemma of Calderón and Zygmund. Among other proofs of (2.11), one of the simpler ones is due to Bennett, DeVore and Sharpley [3] (see the remark at the end of Section 3 on p. 607 of [3]) using a covering lemma which appeared previously in [4].

Remark 2.3. It is a simple exercise to show that the constant B in (2.11) must necessarily satisfy $B \geqslant 1$. Furthermore, versions of (2.11) have been proved in which $B = 2$. (See e.g., [27] or the results of [30] which we shall also discuss below.)

Remark 2.4. It should be noted that a version of the John–Nirenberg inequality was obtained by Korenovskii [15] who showed, for $d = 1$, that the optimal value of the constant b in (2.11) is $2/e$. See also [16, p. 77].

It is also interesting to note that, again for the case where $d = 1$, various sharp constants for different forms of the John–Nirenberg inequality have been obtained by Slavin, Vasyunin and Vol'berg. Their papers [25,26,29] are devoted to the Bellman function approach to this inequality. Naturally, we are very interested in the question of whether it might ultimately be possible to apply the Bellman function technique to obtain new estimates for these constants when $d > 1$ and when $d \to +\infty$. We are grateful to the referee and to Gady Kozma for drawing our attention to these papers.

In this paper we begin the development of some techniques for studying *BMO* type spaces which are apparently somewhat different from those used so far. We will use them to give an alternative proof of a somewhat abstract result which includes, as special cases, both (2.11) and an analogous result of Wik for special rectangles, which we will describe in a moment. Our proof will not be obviously shorter or simpler than the analogous proofs in [13] and [3] and [30] and elsewhere. Nor does it give better constants than the ones obtained by previously published proofs. But it seems distinctly possible that, with further development and refinement, some elements of our approach here may be able to give new information about the behavior of the constants B and b, for large values of the dimension d and perhaps even ultimately lead to determining whether these constants can both be taken to be independent of the dimension d. In Wik's analogue of the inequality (2.11) in the context (2.9) of special rectangles (see (2.13)), the relevant constants are independent of d. Our proof in the same context also gives constants independent of d, but Wik's constants are better than ours. (See Remark 9.2 for details.)

Let us recall some of Wik's results more explicitly, since we will later wish to compare them with ours. For each measurable function $f : \mathbb{R}^d \to \mathbb{R}$ he defines $\|f\|'_{BMO}$ to be the seminorm (2.4) where $\mathcal{E} = \mathcal{W}(\mathbb{R}^d)$ is the collection of all special rectangles (or "false cubes") in $\mathbb{R}^d$, i.e., in our notation

$$\|f\|'_{BMO} = \|f\|^{(\mathbf{D})}_{BMO(\mathbb{R}^d, \mathcal{W}(\mathbb{R}^d))}. \tag{2.12}$$

He defines $\|f\|_{BMO}$ analogously, except that here $\mathcal{E}$ is $\mathcal{Q}(\mathbb{R}^d)$, i.e.,

$$\|f\|_{BMO} = \|f\|^{(\mathbf{D})}_{BMO(\mathbb{R}^d, \mathcal{Q}(\mathbb{R}^d))}.$$

The particular interest of the seminorm $\|f\|'_{BMO}$ lies in the following inequality which is proved by Wik:

$$\lambda\bigl(\bigl\{x \in Q\colon \bigl|f(x) - m_Q\bigr| \geqslant \alpha\bigr\}\bigr) \leqslant 2\lambda(Q)\exp\left(-\frac{\alpha \ln 2}{16\|f\|'_{BMO}}\right). \tag{2.13}$$

(We have written it here using notation slightly different from that of [30], to make it more convenient for comparison with (2.11).) This holds for every special rectangle Q, and for every $\alpha \geqslant 0$ and for every number m_Q which is a median of f on Q.

The final result of Wik's paper (the corollary on p. 199 of [30]) is a variant of (2.11) expressed in terms of medians. Using the considerations described in Lemma 11.5 of [6] we see that Wik's result implies a version of (2.11) in which the constants b and B are given by

$$b = \frac{\ln 2}{32(2 + 6\sqrt{\frac{d}{\pi}})} \quad \text{and} \quad B = 2e^b.$$

We conclude this section by listing some of the features of our approach, some of which have already been discussed or alluded to:

- Probably the most important feature of this paper is that it provides the framework for posing "Question A", whose positive resolution would, as we show, give a dimension free John–Nirenberg inequality.
- In one of the decisive steps of our proof of the John–Nirenberg inequality (see Theorem 8.2), we find that in some sense we have reduced our argument to the case where we only have to consider functions which take the three values 0, 1 and 2.
  As kindly pointed out by the referee, an apparently similar phenomenon, for $d = 1$, has been noted in the paper [29, p. 14], and this suggests the possibility that a Bellman function point of view may perhaps also ultimately provide an alternative explanation of why, also for $d > 1$, everything is reducible to the estimates of functions having only 3 values.
- Instead of working with the seminorms

  $$\|f\|_{BMO(D,\mathcal{Q}(D))} \quad \text{or} \quad \|f\|^{(\mathbf{D})}_{BMO(D,\mathcal{Q}(D))} \quad \text{or} \quad \|f\|^{(\mathbf{A})}_{BMO(D,\mathcal{Q}(D))}$$

  we mainly use another functional, which we denote by $\|f\|_{BMO_{0,s}}$ or $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{Q}(D))}$. This functional was introduced by John [12] and then further studied by Strömberg [28]. The condition sought in Question A implies an inequality of the form

  $$\|f^*\|_{BMO_{0,\sigma}(\mathbb{R})} \leqslant \|f\|_{BMO_{0,s}(\mathbb{R}^d)}$$

  for suitable values of the parameter σ and a suitable class of functions f. Conversely (see Theorem 8.4) if such an inequality holds for some other appropriate value of σ, then it implies the condition sought in Question A.
- Some additional features are mentioned on pp. 7–8 of [6].

3. Properties of non-increasing rearrangements

In this section we shall recall some properties of the non-increasing rearrangements of measurable functions which are defined on an arbitrary measure space (Ω, Σ, μ). Most, indeed probably all of these are well known. A detailed discussion of them can be found, for example, in [9]. Among other relevant references we mention [5] and [21].

For each measurable $f : \Omega \to \mathbb{R}$, one first defines the distribution function $f_* : (0,\infty) \to [0,\infty]$ of f by

$$f_*(\alpha) = \mu\bigl(\bigl\{\omega \in \Omega\colon\ \bigl|f(\omega)\bigr| > \alpha\bigr\}\bigr).$$

One can then define the non-increasing rearrangement $f^* : (0,\infty) \to [0,\infty)$ of f, provided $f_*(\alpha)$ is finite for some positive α. It is given by the formula

$$f^*(t) = \inf\bigl\{\alpha > 0\colon\ f_*(\alpha) \leqslant t\bigr\}$$

for each $t > 0$. It is, roughly speaking, the right continuous "inverse" of the distribution function.

Remark. In all our applications here, we will only need to consider the non-increasing rearrangements of functions for which the set on which they are non-zero has finite measure. Thus the required condition about the finiteness of the distribution function will always be fulfilled.

Here are the properties of the non-increasing rearrangement that we will need in this paper. They all hold for any measurable function $f : \Omega \to \mathbb{R}$ for which $f_*(\alpha) < \infty$ for some $\alpha > 0$. Recall that we denote the one-dimensional Lebesgue measure of subsets G of $(0,\infty)$ by $|G|$.

(i) f^* is non-negative, non-increasing and right continuous on $(0,\infty)$.
(ii) f and f^* have the same distribution functions, i.e., they satisfy

$$\bigl|\bigl\{t > 0\colon\ f^*(t) > \alpha\bigr\}\bigr| = \mu\bigl(\bigl\{\omega \in \Omega\colon\ \bigl|f(\omega)\bigr| > \alpha\bigr\}\bigr) \quad \text{for all } \alpha \in [0,\infty). \tag{3.1}$$

(iii) If $f_*(\beta) < \infty$ for some $\beta \geqslant 0$ then,

$$\bigl|\bigl\{t > 0\colon\ f^*(t) \geqslant \alpha\bigr\}\bigr| = \mu\bigl(\bigl\{\omega \in \Omega\colon\ \bigl|f(\omega)\bigr| \geqslant \alpha\bigr\}\bigr) \quad \text{for all } \alpha \in (\beta,\infty) \tag{3.2}$$

and

$$\bigl|\bigl\{t > 0\colon\ f^*(t) = \alpha\bigr\}\bigr| = \mu\bigl(\bigl\{\omega \in \Omega\colon\ \bigl|f(\omega)\bigr| = \alpha\bigr\}\bigr) \quad \text{for all } \alpha \in (\beta,\infty). \tag{3.3}$$

(iv) The set $\{t > 0\colon\ f^*(t) > \alpha\}$ is the open interval $(0, f_*(\alpha))$ for each $\alpha \in [0,\infty)$.
(v) If $\mu(\Omega) < \infty$, then a variant of (3.3) holds for all $\alpha \in [0,\infty)$, namely

$$\mu\bigl(\bigl\{\omega \in \Omega\colon\ \bigl|f(\omega)\bigr| = \alpha\bigr\}\bigr) = \bigl|\bigl\{t \in \bigl(0, \mu(\Omega)\bigr)\colon\ f^*(t) = \alpha\bigr\}\bigr|. \tag{3.4}$$

We refer, e.g. to [9] and to [6, pp. 9–10], for proofs of these properties.

We close this section with the following lemma which will be needed in Section 4. Its proof is also a straightforward exercise. (Cf. the proof of Lemma 3.2 in [6, p. 10].)

Lemma 3.1. *Let Q be an admissible subset of $\mathbb{R}^d$. Let $g : Q \to [0, \infty)$ be a measurable function. Then,*

$$\left|\left\{t \in \left(0, \lambda(Q)\right):\ \left|g^*(t) - c\right| \leqslant \alpha\right\}\right| = \lambda\left(\left\{x \in Q:\ \left|g(x) - c\right| \leqslant \alpha\right\}\right)$$

for all $c \in \mathbb{R}$ and all $\alpha \geqslant 0$.

4. The functional of John and Strömberg for characterizing *BMO*

Given an admissible subset E of $\mathbb{R}^d$, a real valued function f which is defined and measurable on E, and a number $s \in (0, 1)$, it is convenient to introduce the notation $\mathbf{J}(f, E, s)$ for a special functional which was introduced and studied in [12] and then considered in greater generality in [28]. Thus we set

$$\mathbf{J}(f, E, s) = \inf_{c\in\mathbb{R}}\left(\inf\left\{\alpha \geqslant 0:\ \lambda\left(\left\{x \in E:\ \left|f(x) - c\right| > \alpha\right\}\right) < s\lambda(E)\right\}\right). \tag{4.1}$$

(Here, as always in this paper, λ denotes d-dimensional Lebesgue measure on $\mathbb{R}^d$.) In [12] and [28] the set E is always taken to be a cube, and the functional $\mathbf{J}(f, E, s)$ is shown to be a kind of counterpart, a very useful counterpart, of the functionals $\mathbf{O}(f, E)$, $\mathbf{A}(f, E)$ and $\mathbf{D}(f, E)$.

There is another, perhaps more convenient formula for $\mathbf{J}(f, E, s)$, namely

$$\mathbf{J}(f, E, s) = \inf_{c\in\mathbb{R}}\left((f - c)\chi_E\right)^{*(L)}\left(s\lambda(E)\right).$$

(Here $u^{*(L)}$ denotes the *left* continuous rearrangement of a measurable function u.) This formula is mentioned e.g., in [17,22,24]. In this section we shall obtain yet another formula for $\mathbf{J}(f, E, s)$ in terms of rearrangements. (See Proposition 4.3.)

In our case E will often be a cube or special rectangle, or, more generally, a member of some collection $\mathcal{E}$ of admissible subsets which is used, as in (2.3), together with some measurable set $D \subset \mathbb{R}^d$, to define a seminorm for some version of the space *BMO*. Indeed for such $\mathcal{E}$ and D, following the model of [12] and [28], and analogously to the seminorm defined by (2.3), we consider the functional

$$\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} := \sup_{E\in\mathcal{E}} \mathbf{J}(f, E, s). \tag{4.2}$$

Remark. In the case where $D = \mathbb{R}^d$ and $\mathcal{E} = \mathcal{Q}(\mathbb{R}^d)$ it is known [10,24] that this quantity is equivalent to a certain K-functional. More explicitly,

$$\|f\|^{(\mathbf{J},e^{-t})}_{BMO(D,\mathcal{E})} \sim K\left(t, f; L^\infty\left(\mathbb{R}^d\right), BMO\left(\mathbb{R}^d, \mathcal{Q}\left(\mathbb{R}^d\right)\right)\right).$$

Despite the choice of notation in (4.2), $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}$ is not a norm nor even a seminorm. At least it is homogeneous, i.e., as follows almost immediately from the definition,

$$\mathbf{J}(rf, E, s) = |r|\mathbf{J}(f, E, s) \quad \text{and so}$$
$$\|rf\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} = |r|\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \quad \text{for each } r \in \mathbb{R}. \tag{4.3}$$

Let us note another simple property of these functionals: If T is an invertible affine transformation of $\mathbb{R}^d$, i.e., if $Tx = rx + x_0$ for some non-zero $r \in \mathbb{R}$ and $x_0 \in \mathbb{R}^d$, and if $g(x) = f(rx + x_0)$, then a simple routine calculation (see Appendix 11.5 on pp. 61–62 of [6]) shows that

$$\mathbf{J}(g, E, s) = \mathbf{J}(f, rE + x_0, s) \tag{4.4}$$

for each admissible set E contained in the domain of g. Consequently

$$\|g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} = \|f\|^{(\mathbf{J},s)}_{BMO(T(D),T(\mathcal{E}))} \tag{4.5}$$

where $T(\mathcal{E})$ is of course the collection of sets $\{T(E)\colon\ E \in \mathcal{E}\}$. In various natural examples, where $D = \mathbb{R}^d$ and $\mathcal{E}$ is any one of the collections $\mathcal{Q}(\mathbb{R}^d)$, $\mathcal{D}(\mathbb{R}^d)$, $\mathcal{B}(\mathbb{R}^d)$, $\mathcal{K}(\mathbb{R}^d)$ or $\mathcal{W}(\mathbb{R}^d)$ we of course have $T(D) = D$ and $T(\mathcal{E}) = \mathcal{E}$.

Suppose that $D = \mathbb{R}^d$ and (as in (2.5)) $\mathcal{E}$ is the collection $\mathcal{Q}(\mathbb{R}^d)$ of all cubes in $\mathbb{R}^d$. In this case it will sometimes be convenient to adopt the notation of [28] and write

$$\|f\|_{BMO_{0,s}} = \|f\|^{(\mathbf{J},s)}_{BMO(\mathbb{R}^d,\mathcal{Q}(\mathbb{R}^d))} \tag{4.6}$$

and also

$$\|f\|_{BMO} = \|f\|_{BMO(\mathbb{R}^d,\mathcal{Q}(\mathbb{R}^d))}.$$

It is known that

$$s\|f\|_{BMO_{0,s}} \leqslant \|f\|_{BMO} \leqslant C_d\|f\|_{BMO_{0,s}} \tag{4.7}$$

for every measurable $f : \mathbb{R}^d \to \mathbb{R}$, whenever $0 < s \leqslant \frac{1}{2}$, where C_d is a constant depending only on the dimension d. This result was originally obtained by John [12] for $0 < s < \frac{1}{2}$, and then extended by Strömberg [28] to include the case $s = \frac{1}{2}$. Thus the functional $\mathbf{J}(f, Q, s)$ enables one to characterize *BMO* functions in an alternative way.

The result (4.7) is false for $s > 1/2$, although the definition (4.1) is valid for all $s \in (0, 1)$. This is because $\mathbf{J}(f, E, s) = 0$ and $\|f\|_{BMO_{0,s}} = 0$ for certain non-constant functions f whenever $s > 1/2$. (Cf. the remark on p. 522 of [28].)

Remark 4.1. The essential content of (4.7) is the second inequality. Let us recall the elementary proof of (a more general version of) the first inequality of (4.7). By Chebyshev's inequality we have

$$\lambda\big(\big\{x \in E\colon\ \big|f(x) - c\big| > \alpha\big\}\big) \leqslant \frac{1}{\alpha}\int_E |f - c|\, d\lambda = \frac{\lambda(E)}{\alpha}\mathbf{O}(f, E)$$

for each admissible E, each $\alpha > 0$, each f which is measurable on E, and each median c of f on E. Thus, every α satisfying $\alpha > \frac{1}{s}\mathbf{O}(f, E)$ also satisfies

$$\lambda\big(\big\{x \in E\colon\ \big|f(x) - c\big| > \alpha\big\}\big) < s\lambda(E)$$

for some $c \in \mathbb{R}$. Accordingly, $\mathbf{J}(f, E, s) \leqslant \frac{1}{s}\mathbf{O}(f, E)$ which immediately implies that

$$\|f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)} \leqslant \frac{1}{s}\|f\|_{BMO(D,\mathcal{E})}. \tag{4.8}$$

The first inequality in (4.7) is a special case of (4.8).

The method which we develop in this paper will obviously imply an alternative proof of (4.7), but (so far) only for quite small values of s.

We will sometimes need to use the following very simple result.

Lemma 4.2. *Suppose that the function $\varphi : \mathbb{R} \to \mathbb{R}$ satisfies*

$$\left|\varphi(s) - \varphi(t)\right| \leqslant |s - t|$$

for all $s, t \in \mathbb{R}$. Then

$$\mathbf{J}(\varphi \circ f, E, s) \leqslant \mathbf{J}(f, E, s)$$

for every admissible set E, every $s \in (0, 1)$, and every real valued function f which is defined and measurable on E.

Proof. This follows immediately from the obvious inclusion

$$\left\{x \in E\colon \left|\varphi\bigl(f(x)\bigr) - \varphi(c)\right| > \alpha\right\} \subset \left\{x \in E\colon \left|f(x) - c\right| > \alpha\right\}$$

and the definition of $\mathbf{J}(f, E, s)$. □

We remark that, for each f, E and s as above, and for each $c \in \mathbb{R}$ and each $\alpha \geqslant 0$, the condition

$$\lambda\bigl(\left\{x \in E\colon \left|f(x) - c\right| > \alpha\right\}\bigr) < s\lambda(E)$$

is equivalent to

$$\lambda\bigl(\left\{x \in E\colon \left|f(x) - c\right| \leqslant \alpha\right\}\bigr) > (1 - s)\lambda(E).$$

So we also have

$$\mathbf{J}(f, E, s) = \inf_{c \in \mathbb{R}}\bigl(\inf\left\{\alpha \geqslant 0\colon \lambda\bigl(\left\{x \in E\colon \left|f(x) - c\right| \leqslant \alpha\right\}\bigr) > (1 - s)\lambda(E)\right\}\bigr). \tag{4.9}$$

The following proposition gives us another way to calculate and "visualize" $\mathbf{J}(f, E, s)$, at least for functions which are either univariate and monotone, or non-negative. This other way, for some purposes, seems to be an easier alternative than working with the original definition. It enables us to work with just one variable (here denoted by u), instead of having to deal with the two variables α and c in the original definition.

Proposition 4.3.

(i) *For each $q > 0$ and each non-increasing right continuous function $h : (0, q) \to \mathbb{R}$, the formula*

$$\mathbf{J}\big(h, (0, q), s\big) = \frac{1}{2} \inf\big\{h(u) - h\big(u + (1 - s)q\big) \colon 0 < u < sq\big\} \tag{4.10}$$

holds for each $s \in (0, 1)$.

(ii) *Furthermore, the formula*

$$\mathbf{J}(f, Q, s) = \frac{1}{2} \inf\big\{(f\chi_Q)^*(u) - (f\chi_Q)^*\big(u + (1 - s)\lambda(Q)\big) \colon 0 < u < s\lambda(Q)\big\} \tag{4.11}$$

holds for each admissible subset Q of $\mathbb{R}^d$, each $s \in (0, 1)$ and each ***non-negative*** *real valued function f which is defined and measurable on Q.*

Remark. In our main applications of this proposition the set Q will be a cube or a special rectangle. But we stress that, despite the choice of letter, the set Q in (4.11) can be an *arbitrary* admissible subset.

Remark. Restated informally, part (ii) of this proposition tells us that $2\mathbf{J}(f, Q, s)$ is the "minimum" amount that $(f\chi_Q)^*$ can decrease on any closed subinterval of $(0, \lambda(Q))$ of length exactly $(1 - s)\lambda(Q)$.

Remark. It is easy to see from the original definition or from the formula (4.11), that, for each fixed Q and f the function $s \mapsto \mathbf{J}(f, Q, s)$ is non-increasing. As is explained in Appendix 11.6 of [6, p. 62] (but is not needed for any other purposes in this paper), $s \mapsto \mathbf{J}(f, Q, s)$ is also left continuous, but in general not right continuous.

Proof of Proposition 4.3. We will first deal with part (i). (To understand the rather simple ideas behind our proof of (4.10), the reader may care to first look at the rather shorter and simpler proof given in Remark 4.9 of [6, pp. 16–17] for the special case where h is strictly decreasing and uniformly continuous on $(0, q)$, and to draw some relevant pictures of the graph of h.)

Let β equal the right side of (4.10). We will now prove one "half" of (4.10), namely that $\mathbf{J}(h, (0, q), s) \leqslant \beta$. Obviously $\beta \geqslant 0$ and there exists a non-increasing sequence $\{\beta_n\}_{n\in\mathbb{N}}$ which tends to β and a sequence $\{u_n\}_{n\in\mathbb{N}}$ of numbers satisfying $0 < u_n < sq$ such that

$$\beta_n = \frac{1}{2}\big(h(u_n) - h\big(u_n + (1 - s)q\big)\big).$$

Since $u_n + (1-s)q < q$ and h is right continuous, there exists v_n such that $u_n + (1-s)q < v_n < q$ and

$$0 \leqslant h\big(u_n + (1 - s)\lambda(Q)\big) - h(v_n) \leqslant \frac{1}{n}.$$

If we set $c_n = \frac{1}{2}(h(u_n) + h(v_n))$ and $\alpha_n = \frac{1}{2}(h(u_n) - h(v_n))$ then

$$\begin{aligned}[u_n, v_n] &\subset \{t \in (0,q)\colon\ h(v_n) \leqslant h(t) \leqslant h(u_n)\} \\ &= \{t \in (0,q)\colon\ c_n - \alpha_n \leqslant h(t) \leqslant c_n + \alpha_n\} \\ &= \{t \in (0,q)\colon\ |h(t) - c_n| \leqslant \alpha_n\}.\end{aligned}$$

It follows that

$$\left|\{t \in (0,q)\colon\ |h(t) - c_n| \leqslant \alpha_n\}\right| \geqslant v_n - u_n > (1-s)q.$$

Consequently (by (4.9)) we have $\mathbf{J}(h,(0,q),s) \leqslant \alpha_n$ for each n. Since

$$\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \beta_n = \beta$$

this shows that $\mathbf{J}(h,(0,q),s) \leqslant \beta$.

Next we shall prove the reverse of the preceding inequality, namely that $\beta \leqslant \mathbf{J}(h,(0,q),s)$. Here again we will use sequences denoted by $\{\alpha_n\}_{n\in\mathbb{N}}$, $\{c_n\}_{n\in\mathbb{N}}$, $\{u_n\}_{n\in\mathbb{N}}$ and $\{v_n\}_{n\in\mathbb{N}}$. But they will be defined differently from their definitions in the preceding part of the proof. By (4.9), there exists a non-increasing sequence $\{\alpha_n\}_{n\in\mathbb{N}}$ of non-negative numbers which tends to $\mathbf{J}(h,(0,q),s)$ and a sequence $\{c_n\}_{n\in\mathbb{N}}$ of real numbers such that

$$\left|\{t \in (0,q)\colon\ c_n - \alpha_n \leqslant h(t) \leqslant c_n + \alpha_n\}\right| > (1-s)q. \tag{4.12}$$

Let us define

$$u_n := \inf\{t \in (0,q)\colon\ h(t) \leqslant c_n + \alpha_n\}$$

and

$$v_n := \sup\{t \in (0,q)\colon\ h(t) \geqslant c_n - \alpha_n\}.$$

Then, by definition, for each $m \in \mathbb{N}$, we have that

$$[u_n + 1/m, v_n - 1/m] \subset \{t \in (0,q)\colon\ c_n - \alpha_n \leqslant h(t) \leqslant c_n + \alpha_n\} \subset [u_n, v_n] \cap (0,q).$$

Since we can choose m arbitrarily large, this implies that the intervals

$$(u_n, v_n) \quad \text{and} \quad [u_n, v_n] \cap (0,q)$$

must have the same length as the interval

$$\{t \in (0,q)\colon\ c_n - \alpha_n \leqslant h(t) \leqslant c_n + \alpha_n\}.$$

In view of (4.12), this gives us that $v_n - u_n > (1-s)q$. Furthermore, $0 \leqslant u_n < v_n \leqslant q$. Therefore, for some sufficiently small $\varepsilon_n > 0$, we have $v_n - u_n > 2\varepsilon_n + (1-s)q$ and

$$0 \leqslant u_n < u_n + \varepsilon_n < u_n + \varepsilon_n + (1-s)q < u_n + \varepsilon_n + v_n - u_n - 2\varepsilon_n = v_n - \varepsilon_n < v_n \leqslant q.$$

Since the two points $u_n + \varepsilon_n$ and $u_n + \varepsilon_n + (1-s)q$ are both in $(0, q)$, the number β defined above satisfies

$$\begin{aligned} 2\beta &\leqslant h(u_n + \varepsilon_n) - h\big(u_n + \varepsilon_n + (1-s)q\big) \\ &\leqslant h(u_n + \varepsilon_n) - h(v_n - \varepsilon_n). \end{aligned}$$

By the definitions of u_n and v_n this last expression is dominated by

$$c_n + \alpha_n - (c_n - \alpha_n) = 2\alpha_n.$$

Thus $\beta \leqslant \alpha_n$ for all n. This gives us the remaining required inequality $\beta \leqslant \mathbf{J}(h, (0, q), s)$ and completes the proof of (4.10) and part (i) of the proposition.

Now we turn to part (ii) and the proof of the formula (4.11). We will see that in fact (4.11) can be deduced from (4.10), essentially by a careful application of the fact that the functions f and $(f\chi_Q)^*$, when restricted to Q and to $(0, \lambda(Q))$ respectively, have the same distribution function.

The function $(f\chi_Q)^*$ is non-increasing and right continuous on $(0, \infty)$ and therefore also on the subinterval $(0, \lambda(Q))$. So, we can set $q = \lambda(Q)$ and $h = (f\chi_Q)^*$ and apply (4.10) to obtain that

$$\begin{aligned} &\mathbf{J}\big((f\chi_Q)^*, \big(0, \lambda(Q)\big), s\big) \\ &\quad = \frac{1}{2} \inf\big\{(f\chi_Q)^*(u) - (f\chi_Q)^*\big(u + (1-s)\lambda(Q)\big)\colon\ 0 < u < s\lambda(Q)\big\}. \end{aligned} \tag{4.13}$$

We remark that we have used the notation $(f\chi_Q)^*$ rather than f^* in (4.11) because, in future applications of this proposition, f might possibly be defined on all of $\mathbb{R}^d$ or on some other set which is strictly larger that Q. (Indeed the statement of the proposition explicitly allows for this possibility.) To simplify the notation in the rest of our proof we will let

$$g = f\,|_Q,$$

i.e., $g: Q \to [0, \infty)$ will denote the function defined *only* on Q which is the restriction of f to Q. Thus we can unambiguously write g^* instead of $(f\chi_Q)^*$, and of course $\mathbf{J}(f, Q, s) = \mathbf{J}(g, Q, s)$. In view of (4.13), in order to complete the proof of (4.11) and part (ii) of Proposition 4.3, it will suffice to show that

$$\mathbf{J}(g, Q, s) = \mathbf{J}\big(g^*, \big(0, \lambda(Q)\big), s\big). \tag{4.14}$$

In view of (4.9), we can immediately obtain (4.14) if we know that

$$\big|\big\{t \in \big(0, \lambda(Q)\big)\colon\ \big|g^*(t) - c\big| \leqslant \alpha\big\}\big| = \lambda\big(\big\{x \in Q\colon\ \big|g(x) - c\big| \leqslant \alpha\big\}\big)$$

for all $c \in \mathbb{R}$ and all $\alpha \geqslant 0$. This is exactly the result which was proved in Lemma 3.1 and therefore the proof of part (ii) of Proposition 4.3 is complete. □

5. Non-increasing functions of one variable in BMO. Some simple calculations

Suppose that $d = 1$, that D is a bounded open interval, and that $f : D \to \mathbb{R}$ is non-increasing and right continuous. Our main aim in this section is to prove that a slight variant of the John–Strömberg inequality (namely (5.7)) holds for these very special choices of D and f.

The proof of (5.7) in this special case is of course much simpler than any known proofs of the John–Nirenberg or John–Strömberg inequalities for the general case. But the results of other sections will enable us to deduce the general case from this special case, albeit with not particularly good constants, and with restrictions on the range of the parameter s appearing in the John–Strömberg functional.

We obtain (5.7) as a consequence of the following two lemmata. The first of these bounds the functional $\|f\|^{(\mathbf{J},s)}_{BMO(I,\mathcal{Q}(I))}$ by another functional which has been found to be useful in various contexts and is more or less connected to the functional $\sup_{t>0} f^{**}(t) - f^{*}(t)$ which was introduced in [3]. Other results about these and similar functionals can be found, for example, in [1,2,17,23] and in a large number of subsequent papers.

For simplicity, we only consider (and in fact only need to consider) the interval $I = (0, 1)$ at this stage.

The following lemma can be deduced from a much more general result of Lerner [17, Theorem 3.1, p. 52]. It has a simple and short proof which we include for the reader's convenience, since the proof of Lerner's result in [17] is rather elaborate.

Lemma 5.1. *Suppose that $s \in (0, 1/2)$ and $\rho = \frac{s}{1-s}$. Suppose that $f : (0, 1) \to \mathbb{R}$ is a non-increasing right continuous function. Then*

$$\sup_{t\in(0,1/2]} \big(f(\rho t) - f(t)\big) \leqslant 2\|f\|^{(\mathbf{J},s)}_{BMO((0,1),\mathcal{Q}((0,1)))}.$$

Proof. The properties of f permit us to use the formula (4.10) of Proposition 4.3. Let (a, b) be an arbitrary open subinterval of $(0, 1)$. Via an obvious change of variables (translation, e.g. apply (4.4) with $r = 1$ and $x_0 = a$) the formula (4.10) tells us that

$$\mathbf{J}\big(f, (a, b), s\big) = \frac{1}{2} \inf\big\{f(a + u) - f\big(a + u + (1 - s)(b - a)\big)\colon\ 0 < u < s(b - a)\big\}. \tag{5.1}$$

Let $[c, d]$ be an arbitrary closed subinterval of (a, b) of length $(1 - s)(b - a)$. Then d must satisfy $d > a + (1 - s)(b - a)$ and c must satisfy $c < b - (1 - s)(b - a)$. From these estimates it follows that

$$f(d) \leqslant f\big(a + (1 - s)(b - a)\big) \quad \text{and} \quad f\big(b - (1 - s)(b - a)\big) \leqslant f(c).$$

These estimates imply that

$$f\big(b - (1 - s)(b - a)\big) - f\big(a + (1 - s)(b - a)\big) \leqslant f(c) - f(d).$$

Taking the infimum over all subintervals $[c, d]$ of (a, b) which have length $(1 - s)(b - a)$ and applying (5.1), we see that

$$f\big(b-(1-s)(b-a)\big)-f\big(a+(1-s)(b-a)\big)\leqslant 2\mathbf{J}\big(f,(a,b),s\big)\leqslant 2\|f\|_{BMO(I,\mathcal{Q}(I))}^{(\mathbf{J},s)} \tag{5.2}$$

whenever $0\leqslant a<b\leqslant 1$.

In particular, for an arbitrary $t\in(0,1/2]$, let us choose $a=0$ and $b=\frac{t}{1-s}$. Since $s\in(0,1/2)$ we have $b\in(t,2t)\subset(t,1)$. For these choices of a and b, the left-hand side of (5.2) equals $f(\rho t)-f(t)$. So the proof of the lemma is complete. □

Our second lemma in this section enables us to bound the size of our function f by an expression depending on the functional $\sup_{t\in(0,1/2]}(f(\rho t)-f(t))$ and consequently to obtain an inequality which is quite close to the one that we need.

Lemma 5.2. *The inequality*

$$f(u)-f(v)\leqslant\left(1+\frac{\log\frac{v}{u}}{\log(1/\rho)}\right)\sup_{t\in(0,1/2]}\big(f(\rho t)-f(t)\big) \tag{5.3}$$

holds for every non-increasing function $f:(0,1)\to\mathbb{R}$, *every* $\rho\in(0,1)$, *and every* u *and* v *satisfying* $0<u<v\leqslant 1/2$.

As an immediate consequence we obtain

Corollary 5.3. *If* f *and* ρ *are as in the preceding lemma and if*

$$\sup_{t\in(0,1/2]}\big(f(\rho t)-f(t)\big)\leqslant c,$$

then

$$\big|\big\{t\in(0,1)\colon\ f(t)-f(1/2)\geqslant\alpha\big\}\big|\leqslant\frac{1}{2\rho}\exp\left(-\frac{\alpha\log(1/\rho)}{c}\right)\quad \textit{for each }\alpha\geqslant 0. \tag{5.4}$$

Proof of the lemma and its corollary. Let N be the unique positive integer for which $\rho^N v\leqslant u<\rho^{N-1}v$. Then $(1/\rho)^{N-1}<\frac{v}{u}$ and so $N<1+\frac{\log\frac{v}{u}}{\log(1/\rho)}$. Hence

$$\begin{aligned}f(u)-f(v)&\leqslant f\big(\rho^N v\big)-f\big(\rho^0 v\big)=\sum_{n=1}^{N}\big(f\big(\rho^n v\big)-f\big(\rho^{n-1}v\big)\big)\\&\leqslant N\sup_{t\in(0,1/2]}\big(f(\rho t)-f(t)\big).\end{aligned}$$

This, combined with our estimate for N, establishes (5.3).

Now let us prove (5.4) under the stated hypothesis. Setting $v=1/2$ in (5.3) gives us that

$$f(u)-f(1/2)\leqslant c\left(1-\frac{\log 2u}{\log(1/\rho)}\right)\quad\text{for all } u\in(0,1/2). \tag{5.5}$$

For each $\alpha\geqslant 0$, the set $\{t\in(0,1)\colon\ f(t)-f(1/2)\geqslant\alpha\}$ is of course an interval contained in (0, 1/2]. It follows from (5.5) that the length of this interval cannot exceed $\frac{1}{2\rho}\exp(-\frac{\alpha\log(1/\rho)}{c})$. □

The preceding two lemmata and corollary have the following immediate consequence.

Let $f : I \to \mathbb{R}$ be a non-increasing right continuous function on the interval $I = (0, 1)$. Then, for each $s \in (0, 1/2]$,

$$\left|\left\{t \in I:\ f(t) - f\left(\frac{1}{2}\right) \geqslant \alpha\right\}\right| \leqslant \frac{1-s}{2s} \cdot \exp\left(-\frac{\alpha \log(\frac{1}{s} - 1)}{2\|f\|^{(\mathbf{J},s)}_{BMO(I,\mathcal{Q}(I))}}\right) \quad \text{for all } \alpha \geqslant 0. \tag{5.6}$$

Note that here we can also permit s to take the limiting value $s = 1/2$ and we can permit $\|f\|^{(\mathbf{J},s)}_{BMO(I,\mathcal{Q}(I))}$ to be infinite (provided we agree to interpret both $1/\infty$ and $0/\infty$ as 0). In such cases the right-hand side of (5.6) is greater than or equal to $1/2$ which means that (5.6) is also true, trivially so, in these "limiting" cases.

It will now be a very simple matter to deduce a more general version of (5.6) for the case where I is an arbitrary open interval (a, b). Suppose that $f : (a, b) \to \mathbb{R}$ is right continuous and non-increasing. Define $g : (0, 1) \to \mathbb{R}$ by $g(t) = f(a + (b - a)t)$. Then (cf. (4.4) and (4.5)) we have $\|f\|^{(\mathbf{J},s)}_{BMO(I,\mathcal{Q}(I))} = \|g\|^{(\mathbf{J},s)}_{BMO((0,1),\mathcal{Q}(0,1))}$. Furthermore, the set

$$E_1 = \left\{t \in (a, b):\ f(t) - f\left(\frac{a+b}{2}\right) \geqslant \alpha\right\}$$

coincides with the set $(b - a)E_2 + a$ where

$$E_2 = \{x \in (0, 1):\ g(x) - g(1/2) \geqslant \alpha\}.$$

Therefore, applying (5.6) to the function g and multiplying both sides of the resulting inequality by $b - a = |I|$ gives us that

$$\left|\{t \in I:\ f(t) - f(c_I) \geqslant \alpha\}\right| \leqslant \frac{1-s}{2s} \cdot |I| \cdot \exp\left(-\frac{\alpha \log(\frac{1}{s} - 1)}{2\|f\|^{(\mathbf{J},s)}_{BMO(I,\mathcal{Q}(I))}}\right) \tag{5.7}$$

for every $\alpha \geqslant 0$ and every $s \in (0, 1/2]$ and for every open interval I, where c_I denotes the midpoint of I.

This is the inequality that we need to apply in the proof of our main result, Theorem 9.1 of Section 9.

One immediate consequence of (5.7) together with the inequality (4.8) recalled in Remark 4.1, is that

$$\left|\{t \in I:\ f(t) - f(c_I) \geqslant \alpha\}\right| \leqslant \frac{1-s}{2s} \cdot |I| \cdot \exp\left(-\frac{\alpha s \log(\frac{1}{s} - 1)}{2\|f\|_{BMO(I,\mathcal{Q}(I))}}\right). \tag{5.8}$$

6. A reduction of the proof of the John–Strömberg theorem to a special case

Having, in the previous section, prepared the auxiliary results that we need about special functions of one variable, we now turn to consider functions of several variables.

In our (very slightly) different notation, Lemma 3.1 on p. 517 of [28] states that

$$\lambda\big(\big\{x \in Q\colon \big|f(x) - m_f(Q)\big| > \alpha\big\}\big) \leqslant C\lambda(Q)\exp\biggl(-\frac{c\alpha}{\|f\|_{BMO_{0,s}}}\biggr) \tag{6.1}$$

for all $\alpha \geqslant 0$ and $s \in (0, 1/2]$.

Here C and c are positive constants depending only on d and $m_f(Q)$ is a median of f on the arbitrary cube Q in $\mathbb{R}^d$.

(Here we are again using the notation specified in (4.6). Note that there is a small misprint in [28], namely the factor $\lambda(Q)$ (or $|Q|$) has been omitted there.)

Our main goal in this section is to show that, in order to prove the inequality (6.1) for the specified values of α and s, and some other inequalities like it, it suffices to obtain such inequalities, but with different values of the constants c and C, in the special case where f is a non-negative function taking only integer values. This fact will be precisely formulated as Theorem 6.3. (The question of whether such inequalities actually do hold in that special case will be deferred to Section 9. We will be able to answer it there, with the help of results from other sections.)

Here we can just as easily work in the rather more general context of the space $BMO(D, \mathcal{E})$ of Definition 2.1. Indeed, doing so will be convenient, since we will later want to apply the result of this section in such a general context, which will include, for example, the particular case of special rectangles (2.9) as well as the case of usual cubes (2.5). Thus, throughout this section D will denote some arbitrary but fixed measurable subset of $\mathbb{R}^d$ and $\mathcal{E}$ will denote some arbitrary but fixed collection of admissible subsets of D.

Our first (easy) step is to reduce everything to the case of non-negative functions.

Lemma 6.1. *Let E be a fixed admissible set in $\mathcal{E}$, let s be a fixed number in $(0, 1/2]$, and let c and C be positive constants. Suppose that every non-negative measurable function $f : D \to [0, \infty)$ satisfies the inequality*

$$\lambda\big(\big\{x \in E\colon f(x) > \alpha\big\}\big) \leqslant C\lambda(E)\exp\biggl(-\frac{c\alpha}{\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}}\biggr) \quad \text{for all } \alpha \geqslant 0. \tag{6.2}$$

Then every measurable function $f : D \to \mathbb{R}$ satisfies

$$\lambda\big(\big\{x \in E\colon \big|f(x) - m\big| > \alpha\big\}\big) \leqslant 2C\lambda(E)\exp\biggl(-\frac{c\alpha}{\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}}\biggr) \quad \text{for all } \alpha \geqslant 0 \tag{6.3}$$

whenever m is a median of f on E.

In fact this same conclusion also holds under weaker hypotheses, namely if it is only known that (6.2) *holds for those non-negative functions $f : D \to [0, \infty)$ having the additional property that*

$$\lambda\big(\big\{x \in E\colon f(x) > 0\big\}\big) \leqslant \frac{1}{2}\lambda(E). \tag{6.4}$$

Remark. Our "natural" applications of Lemma 6.1, will be in the case where the collection $\mathcal{E}$ includes the set D itself, and we choose $E = D$.

Proof of Lemma 6.1. We have to prove (6.3) for an arbitrary measurable function $f : D \to \mathbb{R}$ with median m on E. Let $g = f - m$. Obviously $\mathbf{J}(g, E, s) = \mathbf{J}(f, E, s)$ and $\|g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} = \|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}$ and so it will suffice to prove that

$$\lambda\bigl(\bigl\{x \in E:\ \bigl|g(x)\bigr| > \alpha\bigr\}\bigr) \leqslant 2C\lambda(E)\exp\Biggl(-\frac{c\alpha}{\|g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}}\Biggr) \quad \text{for all } \alpha \geqslant 0. \tag{6.5}$$

The left-hand side of (6.5) equals

$$\begin{aligned}&\lambda\bigl(\bigl\{x \in E:\ g(x) > \alpha\bigr\}\bigr) + \lambda\bigl(\bigl\{x \in E:\ g(x) < -\alpha\bigr\}\bigr)\\ &\quad = \lambda\bigl(\bigl\{x \in E:\ g_+(x) > \alpha\bigr\}\bigr) + \lambda\bigl(\bigl\{x \in E:\ g_-(x) > \alpha\bigr\}\bigr)\end{aligned}$$

where, as usual, $g_+ = \max\{g, 0\}$ and $g_- = g_+ - g = \max\{-g, 0\}$. We can apply Lemma 4.2 with $\varphi(t) = \max\{t, 0\}$ to obtain that

$$\|g_+\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant \|g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \quad \text{and} \quad \|g_-\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant \|-g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}.$$

Obviously $\|-g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} = \|g\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}$. Thus, if we apply (6.2) to each of the non-negative functions g_+ and g_- and sum the results, we obtain (6.5) and therefore that f indeed satisfies (6.3).

It remains to justify the claim in the last sentence of the statement of the lemma. Since m is a median of f, it follows that 0 is a median of g and so the two functions g_+ and g_- to which we have applied (6.2) satisfy

$$\lambda\bigl(\bigl\{x \in E:\ g_+(x) > 0\bigr\}\bigr) \leqslant \frac{1}{2}\lambda(E) \quad \text{and} \quad \lambda\bigl(\bigl\{x \in E:\ g_-(x) > 0\bigr\}\bigr) \leqslant \frac{1}{2}\lambda(E). \qquad \square$$

Our next step is to reduce the proof of (6.2) to the case of appropriate integer valued functions.

Lemma 6.2. *Let E, $\mathcal{E}$, s, c and C be as in the statement of Lemma* 6.1. *Suppose that* (6.2) *holds for every non-negative measurable function $f : D \to [0, \infty)$ which takes only integer values and satisfies $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant 1/2$ and* (6.4). *Then* (6.2) *also holds for every measurable $f : D \to [0, \infty)$, which satisfies* (6.4), *but with the constants c and C replaced respectively by $c_1 = c/4$ and $C_1 = \max\{C, e^c\}$.*

Remark. It will be clear from the following proof that we can also obtain the following additional result: Suppose that in the above lemma we impose the stronger condition that (6.2) holds also for every non-negative measurable function $f : D \to [0, \infty)$ which takes only integer values and satisfies $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant 1/2$ but does not necessarily satisfy (6.4). Then we obtain the stronger conclusion that (6.2) holds for every measurable $f : D \to [0, \infty)$, but with the constants c and C replaced by $c_1 = c/4$ and $C_1 = \max\{C, e^c\}$. I.e., in this case we can obtain (6.2) also for functions $f : D \to [0, \infty)$ which do not satisfy (6.4).

Proof of Lemma 6.2. We shall use the function $\varphi : [0, \infty) \to [0, \infty)$ which is defined by

$$\varphi(t) = \begin{cases} 0, & 0 \leqslant t \leqslant 1/2,\\ n, & n - 1/2 < t \leqslant n + 1/2 \text{ for each } n \in \mathbb{N}.\end{cases}$$

We will need the following three obvious or easily verified properties of φ:

$$\varphi\big([0,\infty)\big)=\mathbb{N}\cup\{0\},$$
$$\varphi(t)-\varphi(s)\in\{0,1\}\quad\text{whenever}\quad 0\leqslant s\leqslant t\leqslant s+1/2 \tag{6.6}$$

and

$$\varphi(t)\geqslant t-1/2\quad\text{for all } t\geqslant 0. \tag{6.7}$$

Suppose that $f:D\to[0,\infty)$ is an arbitrary measurable function which satisfies (6.4) and also

$$0<\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}<\frac{1}{3}. \tag{6.8}$$

The composed function $\varphi\circ f$ also satisfies (6.4) since

$$\begin{aligned}\lambda\big(\big\{x\in E\colon\ \varphi\circ f(x)>0\big\}\big)&=\lambda\big(\big\{x\in E\colon\ \varphi\circ f(x)\geqslant 1\big\}\big)\\&=\lambda\big(\big\{x\in E\colon\ f(x)>1/2\big\}\big)\leqslant\frac{1}{2}\lambda(E).\end{aligned}$$

We will next show that, furthermore, $\varphi\circ f$ satisfies

$$\|\varphi\circ f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}\leqslant 1/2. \tag{6.9}$$

Let W be an arbitrary set in $\mathcal{E}$. Then (6.8) implies that $\mathbf{J}(f,W,s)<1/3$. Therefore (cf. (4.9)) there exists some $\alpha\in[0,1/3)$ and some $\gamma\in\mathbb{R}$ such that

$$\lambda\big(\big\{x\in W\colon\ \big|f(x)-\gamma\big|\leqslant\alpha\big\}\big)>(1-s)\lambda(W). \tag{6.10}$$

Let us choose

$$\gamma_1=\frac{1}{2}\big(\varphi(\gamma-\alpha)+\varphi(\gamma+\alpha)\big)\quad\text{and}\quad\alpha_1=\frac{1}{2}\big(\varphi(\gamma+\alpha)-\varphi(\gamma-\alpha)\big).$$

Since $0\leqslant\alpha<1/3$ we obtain from (6.6) that α_1 is either 0 or $1/2$. Since φ is non-decreasing, we also obtain that

$$\begin{aligned}\big\{x\in W\colon\ \big|f(x)-\gamma\big|\leqslant\alpha\big\}&=\big\{x\in W\colon\ \gamma-\alpha\leqslant f(x)\leqslant\gamma+\alpha\big\}\\&\subset\big\{x\in W\colon\ \varphi(\gamma-\alpha)\leqslant\varphi\circ f(x)\leqslant\varphi(\gamma+\alpha)\big\}\\&=\big\{x\in W\colon\ \gamma_1-\alpha_1\leqslant\varphi\circ f(x)\leqslant\gamma_1+\alpha_1\big\}\\&=\big\{x\in W\colon\ \big|\varphi\circ f(x)-\gamma_1\big|\leqslant\alpha_1\big\}.\end{aligned}$$

Thus we deduce, using (6.10) and (4.9) once more, that $\mathbf{J}(\varphi\circ f,W,s)\leqslant\alpha_1\leqslant 1/2$ and this establishes (6.9).

Since the function $\varphi \circ f$ is also non-negative and integer valued, we have, according to the hypotheses of the lemma, that

$$\lambda\big(\big\{x \in E\colon\ \varphi \circ f(x) > \alpha\big\}\big) \leqslant C\lambda(E)\exp\bigg(-\frac{c\alpha}{\|\varphi \circ f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)}}\bigg) \quad \text{for all } \alpha \geqslant 0. \tag{6.11}$$

The inequality (6.7) implies that

$$\big\{x \in E\colon\ f(x) > \alpha\big\} \subset \big\{x \in E\colon\ \varphi \circ f(x) > \alpha - 1/2\big\},$$

and consequently, for all $\alpha > 1/2$, it follows, using (6.11) and then (6.9) and then (6.8), that

$$\begin{aligned}\lambda\big(\big\{x \in E\colon\ f(x) > \alpha\big\}\big) &\leqslant C\lambda(E)\exp\bigg(-\frac{c(\alpha - 1/2)}{\|\varphi \circ f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)}}\bigg)\\ &\leqslant C\lambda(E)\exp\big(-2c(\alpha - 1/2)\big).\end{aligned}$$

If we now restrict α to the range $\alpha \geqslant 1$, we also have $\alpha - 1/2 \geqslant \alpha/2$ and so

$$\lambda\big(\big\{x \in E\colon\ f(x) > \alpha\big\}\big) \leqslant C\lambda(E)\exp(-c\alpha) \leqslant C_1\lambda(E)\exp(-c\alpha),$$

recalling, as stated in the lemma, that $C_1 = \max\{C, e^c\}$.

Now let us consider the range of values $0 \leqslant \alpha < 1$. Of course

$$\lambda\big(\big\{x \in E\colon\ f(x) > \alpha\big\}\big) \leqslant \lambda(E)$$

for these (and all other) values of α. Furthermore, for each $\alpha \in [0, 1)$ we have $C_1 e^{-c\alpha} \geqslant 1$ and therefore

$$\lambda\big(\big\{x \in E\colon\ f(x) > \alpha\big\}\big) \leqslant C_1\lambda(E)\exp(-c\alpha).$$

We have thus now shown that, subject to the hypotheses of the lemma, the inequality

$$\lambda\big(\big\{x \in E\colon\ f(x) > \beta\big\}\big) \leqslant C_1\lambda(E)\exp(-c\beta) \tag{6.12}$$

holds for all $\beta \geqslant 0$ and for all those measurable functions $f : D \to [0, \infty)$ which satisfy

$$0 < \|f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)} < 1/3 \tag{6.13}$$

and (6.4). Now we can easily obtain the required inequality (6.2) without having to impose (6.13). Given an arbitrary measurable function $f : D \to [0, \infty)$ satisfying (6.4) and any $\alpha > 0$, we let $\widetilde{f} = f/4\|f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)}$ and choose $\beta = \alpha/4\|f\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)}$. Then by homogeneity (cf. (4.3)), we have $\|\widetilde{f}\|_{BMO(D,\mathcal{E})}^{(\mathbf{J},s)} = 1/4$ and of course $\widetilde{f}$ also satisfies (6.4). So we can apply (6.12) to $\widetilde{f}$ and obtain (6.2), completing the proof of the lemma. □

We can summarize the results of this section by the following theorem, whose proof follows immediately from the previous two lemmata.

Theorem 6.3. *Let E be a fixed admissible set in $\mathcal{E}$, let s be a fixed number in* $(0, 1/2]$, *and let c and C be positive constants. Let Φ be the set of all non-negative measurable functions $f : D \to [0,\infty)$ which take only integer values, and which satisfy $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant 1/2$ and $\lambda(\{x \in E\colon f(x) > 0\}) \leqslant \frac{1}{2}\lambda(E)$. Suppose that every $f \in \Phi$ satisfies*

$$\lambda\bigl(\bigl\{x \in E\colon f(x) > \alpha\bigr\}\bigr) \leqslant C\lambda(E)\exp\biggl(-\frac{c\alpha}{\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}}\biggr) \quad \textit{for all } \alpha \geqslant 0. \tag{6.14}$$

Then every measurable function $f : D \to \mathbb{R}$ satisfies

$$\lambda\bigl(\bigl\{x \in E\colon \bigl|f(x) - m\bigr| > \alpha\bigr\}\bigr) \leqslant 2\max\bigl\{C, e^c\bigr\}\lambda(E)\exp\biggl(-\frac{c\alpha}{4\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})}}\biggr) \tag{6.15}$$

for all $\alpha \geqslant 0$, whenever m is a median of f on E.

7. The "geometrical" component of our proof

7.1. A "balancing act" between two subsets of a cube. The "bi-density" constant

We begin by stating a simple result which is a sort of "prototype" of the main result that we seek in this section. It will also be a tool for proving that main result.

Lemma 7.1. *Let Q be a cube in $\mathbb{R}^d$ and let E be a measurable subset of Q such that $0 < \lambda(E) < \lambda(Q)$. Then there exists a cube W contained in Q such that*

$$\lambda(W \setminus E) = \lambda(W \cap E) = \frac{1}{2}\lambda(W). \tag{7.1}$$

A slightly more general result is presented in Lemma 7.5 below. Setting $s = 1/2$ in that lemma will give the result just stated here.

If the cube Q in Lemma 7.1 is dyadic, and we want to only consider subcubes W which are also dyadic, then we cannot hope in general to obtain one of them which satisfies (7.1). Instead, as the next lemma states, we can obtain a dyadic subcube W satisfying a rather weaker property.

Lemma 7.2. *Let Q be a dyadic cube in $\mathbb{R}^d$ and let E be a measurable subset of Q such that $0 < \lambda(E) < \lambda(Q)$. Then there exists a dyadic cube W contained in Q such that*

$$\min\bigl\{\lambda(W \setminus E), \lambda(W \cap E)\bigr\} \geqslant 2^{-d}\bigl(1 - 2^{-d}\bigr)\lambda(W). \tag{7.2}$$

For a simple proof of this lemma we refer the reader to [6, p. 25].

Remark 7.3. We have not bothered to check whether it is possible to obtain a stronger conclusion in Lemma 7.2 where the constant $2^{-d}(1 - 2^{-d})$ in (7.2) is replaced by some larger constant. However any such improvement would not be very significant, since the simple example where E is a dyadic subcube of Q shows that the constant in (7.2) cannot exceed 2^{-d}.

Referring back to the terminology introduced in (2.5), (2.6), (2.7), (2.8) and (2.9) we can see that Lemma 7.1 gives us information about the collection $\mathcal{Q}(D)$ of subcubes of D, and Lemma 7.2 gives us analogous information about the collection $\mathcal{D}(D)$ of dyadic subcubes of D. The following notion will put the results of these two lemmata in a more general context.

Definition 7.4. Let $\mathcal{E}$ be a collection of admissible subsets of $\mathbb{R}^d$. We say that a number δ is a *bi-density constant* for $\mathcal{E}$ if, for each $Q \in \mathcal{E}$ and for each measurable set E for which $0 < \lambda(Q \cap E) < \lambda(Q)$, there exists some set $W \in \mathcal{E}$ with $W \subset Q$ such that

$$\min\{\lambda(W \cap E), \lambda(W \setminus E)\} \geqslant \delta\lambda(W).$$

Thus Lemma 7.1 tells us that $\delta = 1/2$ is a bi-density constant for $\mathcal{Q}(D)$. The next lemma will show (when we substitute $s = 1/2$) that this is also the case for the collections $\mathcal{K}(D)$ and $\mathcal{W}(D)$ (defined above in (2.8) and (2.9)). The value 1/2 is in fact optimal since, clearly, any bi-density constant for any collection $\mathcal{E}$ always has to satisfy $\delta \leqslant 1/2$. Lemma 7.2 and Remark 7.3 tell us that every bi-density constant δ for $\mathcal{D}(D)$ must satisfy $2^{-d}(1 - 2^{-d}) \leqslant \delta \leqslant 2^{-d}$.

Lemma 7.5. *(See [6, p. 26].) Let K be a convex subset of $\mathbb{R}^d$ with non-empty interior and let E be a measurable subset of K such that $0 < \lambda(E) < \lambda(K)$. Then, given an arbitrary number $s \in (0, 1)$, there exists a cube $W \subset K$ for which*

$$\frac{\lambda(W \cap E)}{\lambda(W)} = s. \tag{7.3}$$

7.2. Some preparations for a "balancing act" between three subsets of a cube

In the previous subsection we considered cubes Q which are the unions of *two* disjoint sets E and $Q \setminus E$ which both have positive measure, and we have obtained a subcube W of Q whose intersections with E and with $Q \setminus E$ are both "significant" proportions of W. Our main goal in the sequel will be to "upgrade" these kinds of results to a situation where the cube Q (or a more general set) is the union of *three* disjoint sets, which we may denote by E_+ and E_- and G. We will show (in Theorem 7.7) that, under certain conditions, there is a kind of "tri-density constant". Let us try to express this a little more explicitly: If G is a "relatively small" part of Q then we will show that there is a subcube W of Q whose intersections with E_+ and E_- are "significant" proportions of W. We will formulate this result in a more general context where Q and W are not necessarily cubes, but are members of some suitable collection $\mathcal{E}$ of admissible sets. In Section 8 we will see the implications of this property of three subsets for the study of various versions of *BMO*.

In order to express our main result for a more general choice of collections $\mathcal{E}$ of admissible sets, we need to define some more notions. Our point of departure for doing this comes from considering two important examples:

(i) Every cube Q in $\mathbb{R}^d$ is of course the union of 2^d non-overlapping subcubes, each having volume $2^{-d}\lambda(Q)$. Then each of these subcubes can of course in turn be subdivided into 2^d non-overlapping subcubes of volume $2^{-2d}\lambda(Q)$, ... and this process can be continued indefinitely.

(ii) Every special rectangle Q in $\mathbb{R}^d$ is of course the union of 2 non-overlapping special rectangles, each having volume $2^{-1}\lambda(Q)$. Then each of these special rectangles can of course in turn subdivided into 2 non-overlapping special rectangles of volume $2^{-2}\lambda(Q)$, ... and this process too can be continued indefinitely.

Here is a notion which incorporates these two examples, and ultimately other examples.

Definition 7.6. Let E be an admissible set in $\mathbb{R}^d$. Let $M \geqslant 2$ be a positive integer. We will say that the doubly indexed sequence $\{E_{j,k}\}_{j\geqslant 0,\, 1\leqslant k\leqslant M^j}$ of admissible sets is a *multilevel decomposition of E with multiplicity M* if it satisfies the following conditions:

(i) $E_{0,1} = E$, and, more generally,

$$E = \bigcup_{k=1}^{M^j} E_{j,k}$$

for each fixed $j \geqslant 0$.

(ii) For each fixed $j \geqslant 1$ the sets $E_{j,k}$ satisfy $\lambda(E_{j,k} \cap E_{j,k'}) = 0$ whenever $k \neq k'$.

(iii) $\lambda(E_{j,k}) = M^{-j}\lambda(E)$ for each $j \geqslant 0$ and $k \in \{1, 2, \dots, M^j\}$.

(iv) For each fixed $j \geqslant 1$, and $k \in \{1, 2, \dots, M^j\}$, the set $E_{j,k}$ is the union of M sets from among the M^{j+1} sets $E_{j+1,m}$. More explicitly,

$$E_{j,k} = \bigcup_{m=M(k-1)+1}^{Mk} E_{j+1,m}.$$

(v) The diameters of the sets $E_{j,k}$ tend to zero uniformly as j tends to infinity, i.e.,

$$\lim_{j\to\infty}\Big(\max_{1\leqslant k\leqslant M^j} \operatorname{diam} E_{j,k}\Big) = 0. \tag{7.4}$$

The preceding definition leads us immediately to this next one.

Definition. Let $\mathcal{E}$ be a collection of admissible subsets of $\mathbb{R}^d$ and let $M \geqslant 2$ be an integer. We will say that $\mathcal{E}$ is *M-multidecomposable* if every set $E \in \mathcal{E}$ has a multilevel decomposition $\{E_{j,k}\}_{j\geqslant 0,\, 1\leqslant k\leqslant M^j}$ of multiplicity M where all of the sets $E_{j,k}$ are also in $\mathcal{E}$.

So, of course, for any open subset D of $\mathbb{R}^d$, the collections $\mathcal{Q}(D)$ and $\mathcal{D}(D)$ are both 2^d-multidecomposable, and the collection $\mathcal{W}(D)$ is 2-multidecomposable. It is probably easy to show that the collection $\mathcal{K}(D)$ is also 2-decomposable.

7.3. Our main "geometrical" result. The promised "balancing act" between three subsets of a (generalized) cube

Theorem 7.7. *Let $\mathcal{E}$ be an M-multidecomposable collection of admissible subsets of $\mathbb{R}^d$ for some $M \geqslant 2$. Let τ be a positive number. Let δ be a bi-density constant for $\mathcal{E}$.*

Suppose that Q is a set in $\mathcal{E}$ and there exist three pairwise disjoint measurable sets E_+, E_- and G which satisfy

$$Q = E_+ \cup E_- \cup G$$

and

$$\min\{\lambda(E_+), \lambda(E_-)\} > \tau\lambda(G). \tag{7.5}$$

Then there exists a subset W of Q such that $W \in \mathcal{E}$ and

$$\min\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\} \geqslant s\lambda(W) \tag{7.6}$$

where

$$s = \begin{cases} \min\{\frac{\tau-\tau^2}{M(1+\tau)}, \delta\}, & 0 < \tau \leqslant \sqrt{2}-1, \\ \min\{\frac{3-2\sqrt{2}}{M}, \delta\}, & \sqrt{2}-1 \leqslant \tau. \end{cases}$$

It will be convenient to explicitly state some immediate consequences of Theorem 7.7 for some special choices of Q and $\mathcal{E}$.

Corollary 7.8. *Suppose that Q is, respectively* (i) *a cube, or* (ii) *a dyadic cube or* (iii) *a special rectangle in $\mathbb{R}^d$ and that, respectively,* (i) $\mathcal{E} = \mathcal{Q}(Q)$, *or* (ii) $\mathcal{E} = \mathcal{D}(Q)$, *or* (iii) $\mathcal{E} = \mathcal{W}(Q)$. *Suppose that Q is the disjoint union of the three sets E_-, E_+ and G and that*

$$\min\{\lambda(E_+), \lambda(E_-)\} > (\sqrt{2}-1)\lambda(G).$$

Then there exists a set $W \subset Q$ which is, respectively, (i) *a cube, or* (ii) *a dyadic cube or* (iii) *a special rectangle, and which satisfies*

$$\min\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\} \geqslant s\lambda(W)$$

where, respectively, (i) $s = 2^{-d}(3-2\sqrt{2})$, *or* (ii) $s = 2^{-d}(3-2\sqrt{2})$ (*again*), *or* (iii) $s = (3-2\sqrt{2})/2$.

Proof. We simply apply Theorem 7.7, substituting the known values for δ and M in the formula $s = \min\{\frac{3-2\sqrt{2}}{M}, \delta\}$ in each of the three cases. □

Theorem 7.7 and Corollary 7.8 motivate us to introduce another notion. This notion will enable a convenient formulation of the main question raised by this paper, and also a convenient proof of the consequences that an affirmative answer to that question would have.

Definition 7.9. Let $\mathcal{E}$ be a collection of admissible subsets of $\mathbb{R}^d$. Let τ and s be positive numbers with the following property:

Let Q be a set in $\mathcal{E}$ and let E_+, E_- and G be arbitrary pairwise disjoint admissible sets whose union is Q. Suppose that

$$\min\{\lambda(E_+), \lambda(E_-)\} > \tau\lambda(G). \tag{7.7}$$

Then there exists a set $W \subset Q$ which is also in $\mathcal{E}$ and for which

$$\min\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\} \geqslant s\lambda(W). \tag{7.8}$$

Then we will say that (τ, s) *is a John–Strömberg pair for* $\mathcal{E}$.

Remark. The preceding definition is formulated for all possible positive values of τ and s. However, for our particular applications we are interested only in cases where $\tau < 1/2$. This is why, in the formulation of Question A, we apply this latter restriction to τ.

Example 7.10. We can reformulate Corollary 7.8 as follows: We take $\tau = \sqrt{2} - 1$. Then, for any open set D of $\mathbb{R}^d$, we have that:

(i) $(\sqrt{2} - 1, 2^{-d}(3 - 2\sqrt{2}))$ is a John–Strömberg pair for $\mathcal{Q}(D)$ and also for $\mathcal{D}(D)$.
(ii) $(\sqrt{2} - 1, \frac{3-2\sqrt{2}}{2})$ is a John–Strömberg pair for $\mathcal{W}(D)$.

Remark 7.11. In the special case where $\mathcal{E} = \mathcal{Q}(D)$ for some open subset D of $\mathbb{R}^d$ then, obviously, (τ, s) is a John–Strömberg pair if and only if the condition appearing in the second paragraph of Definition 7.9 holds for just one particular cube Q in $\mathbb{R}^d$, for example for the unit cube $Q = [0, 1]^d$. The cube Q does not have to be contained in D. (In fact the particular choice of D is irrelevant in this case. (τ, s) is a John–Strömberg pair for $\mathcal{Q}(D)$ if and only if it is a John–Strömberg pair for $\mathcal{Q}(\mathbb{R}^d)$.)

Remark 7.12. We can now express Question A concisely in the language of Definition 7.9. Question A simply asks whether there exist two absolute constants $s > 0$ and $\tau \in (0, 1/2)$ such that (τ, s) is a John–Strömberg pair for $\mathcal{Q}(\mathbb{R}^d)$ for *every* $d \in \mathbb{N}$.

Remark 7.13. In Definition 7.9 the set G may be chosen to be empty. Therefore any number s, which happens to form a John–Strömberg pair (τ, s) for $\mathcal{E}$ with some positive number τ, will also automatically be a bi-density constant for $\mathcal{E}$. The particular value of τ is immaterial here.

Remark. Although the only restriction that we explicitly impose on s is that it has to be positive, we see from (7.8) that some cube W, which has positive measure, must satisfy

$$\lambda(W) \geqslant \lambda(E_+ \cap W) + \lambda(E_- \cap W) \geqslant 2\min\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\} \geqslant 2s\lambda(W).$$

Thus the constant s in the above definition can never be greater than $1/2$.

Proof of Theorem 7.7. In view of the regularity of λ there exist compact subsets H_+ and H_- of E_+ and E_- respectively such that

$$\min\{\lambda(H_+), \lambda(H_-)\} > \tau\lambda(G).$$

If there exists $W \in \mathcal{E}$ such that $W \subset Q$ and

$$\min\bigl\{\lambda(H_+ \cap W), \lambda(H_- \cap W)\bigr\} \geqslant s\lambda(W)$$

then obviously W also satisfies (7.6). This means that we may assume without loss of generality that E_+ and E_- are themselves compact sets. Since they are also disjoint, it follows that

$$\rho := \operatorname{dist}(E_+, E_-) > 0.$$

We will say that the set W is a *good set* if $W \in \mathcal{E}$ and $W \subset Q$ and

$$\min\bigl\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\bigr\} > \tau\lambda(G \cap W).$$

(We mention that a slight variant of this definition plays a role in Section 10.2 of [6, pp. 47–52], in particular in the proof of the rather technical Theorem 10.6 there (pp. 49–52), which has some similarities with some of the arguments here.)

Let $\{Q_{j,k}\}_{j\geqslant 0,\, 1\leqslant k\leqslant M^j}$ be a multilevel decomposition of Q of multiplicity M where all of the sets $Q_{j,k}$ are in $\mathcal{E}$. A sequence $\{Q^{(n)}\}_{0\leqslant n<\ell}$ of sets in $\mathcal{E}$, where ℓ can be finite or infinite, will be called a *chain* if $Q^{(0)} = Q$ and, if for each n such that $1 \leqslant n < \ell$ we have $Q^{(n)} = Q_{n,k_n}$ for some integer $k_n \in \{1, 2, \dots, M^n\}$. When $\ell = \infty$, we have (by (7.4)) that $\lim_{n\to\infty} \operatorname{diam} Q^{(n)} = 0$. So there must exist some n_1 such that $\operatorname{diam} Q^{(n)} < \rho$ for all $n \geqslant n_1$. Thus, for $n \geqslant n_1$ the set $Q^{(n)}$ cannot intersect with both of E_+ and E_- and so cannot be a good set.

Let construct a particular chain $\{Q^{(n)}\}_{0\leqslant n<\ell}$ in the following way. We of course have to start with $Q^{(0)} = Q$. If among all the sets $Q_{1,k}$ for $k \in \{1, 2, \dots, M\}$ there is no set which is good, then we set $\ell = 1$ and our construction is complete. Otherwise we choose $Q^{(1)}$ to be a good set from the above list. Next we check whether, among those of the sets $Q_{2,k}$ for $k \in \{1, 2, \dots, M^2\}$ which are contained in $Q^{(1)}$, there is one which is a good set. If so we choose $Q^{(2)}$ to be such a set. If not, we set $\ell = 2$ and our construction is complete. The continuation of this process is now clear. At the nth stage we seek a good set, which we will call $Q^{(n)}$, from among those of the sets $Q_{n,k}$ which are contained in $Q^{(n-1)}$. If no such set exists, then we set $\ell = n$ and the construction is complete. In view of the arguments presented in the previous paragraph, we must have $\ell < \infty$, i.e., the construction necessarily has to terminate after finitely many steps.

Thus we have obtained a good set $Q^{(\ell-1)}$ which is one of the sets $Q_{\ell-1,k}$ for some integer k and it will enable us to complete the proof of the theorem. Let us denote the M sets of the form $Q_{\ell,m}$ which are contained in $Q^{(\ell-1)}$ by $W_1, W_2, \dots, W_M$. By our construction, none of these sets are good sets. For the rest of the proof we will assume that

$$\lambda\bigl(Q^{(\ell-1)} \cap E_-\bigr) \leqslant \lambda\bigl(Q^{(\ell-1)} \cap E_+\bigr).$$

If the reverse inequality holds then we will simply use an exact analogue of the proof that we are about to give, where we will simply interchange the roles of E_+ and E_-.

We have to consider three cases:

Case (i). Suppose that $\lambda(Q^{(\ell-1)} \cap G) = 0$. Then, since δ is a bi-density constant for $\mathcal{E}$, and since $\lambda(Q^{(\ell-1)} \cap E_+)$ and $\lambda(Q^{(\ell-1)} \setminus E_+) = \lambda(Q^{(\ell-1)} \cap E_-)$ are both positive, there exists $W \in \mathcal{E}$ such that $W \subset Q^{(\ell-1)}$ and

$$\min\{\lambda(W\cap E_+),\lambda(W\cap E_-)\}\geqslant\delta\lambda(W)\geqslant s\lambda(W), \tag{7.9}$$

completing the proof of the theorem.

Case (ii). Suppose that $\lambda(W_m\cap E_-)=\lambda(W_m)$ for some $m\in\{1,2,\dots,M\}$.

Then

$$\begin{aligned}\lambda\big(Q^{(\ell-1)}\cap E_+\big)\geqslant\lambda\big(Q^{(\ell-1)}\cap E_-\big)\geqslant\lambda(W_m\cap E_-)=\lambda(W_m)=M^{-1}\lambda\big(Q^{(\ell-1)}\big)\\ \geqslant s\lambda\big(Q^{(\ell-1)}\big).\end{aligned}$$

So we see that in this case the set $W=Q^{(\ell-1)}$ has the properties required to complete the proof of the theorem.

Case (iii). This is the remaining case where the first and second cases are excluded. I.e., we have $\lambda(Q^{(\ell-1)}\cap G)>0$ and

$$\lambda(W_m\cap E_-)<\lambda(W_m)\quad\text{for all } m\in\{1,2,\dots,M\}. \tag{7.10}$$

The inequality

$$\lambda\big(Q^{(\ell-1)}\cap E_-\big)>\tau\lambda\big(Q^{(\ell-1)}\cap G\big)>0$$

(which holds because $Q^{(\ell-1)}$ is good) can be rewritten (in view of condition (ii) of Definition 7.6) as

$$\sum_{m=1}^{M}\lambda(W_m\cap E_-)>\sum_{m=1}^{M}\tau\lambda(W_m\cap G)>0. \tag{7.11}$$

Let N be the (possibly empty) set of all integers m in $\{1,2,\dots,M\}$ which satisfy

$$\lambda(W_m\cap G)=0.$$

If $\lambda(W_m\cap E_-)>0$ for some $m\in N$, then, since (7.10) also holds, we can again, analogously to what was done in case (i), invoke the bi-density condition to obtain some $W\in\mathcal{E}$ with $W\subset W_m$ which satisfies (7.9) and so completes the proof. This means that we can now assume that $\lambda(W_m\cap E_-)=0$ for all $m\in N$. Therefore (7.11) can be rewritten as

$$\sum_{m\in\{1,2,\dots,M\}\setminus N}\lambda(W_m\cap E_-)>\sum_{m\in\{1,2,\dots,M\}\setminus N}\tau\lambda(W_m\cap G)>0.$$

It follows that there exists at least one $m\in\{1,2,\dots,M\}\setminus N$ which satisfies

$$\lambda(W_m\cap E_-)>\tau\lambda(W_m\cap G)>0.$$

Recall that, by our construction, W_m is not good. Therefore we must have

$$\lambda(W_m\cap E_+)\leqslant\tau\lambda(W_m\cap G).$$

We use this and the preceding inequality to obtain that

$$\begin{aligned}\lambda(W_m \cap E_-) > \tau\lambda(W_m \cap G) &= \tau\big[\lambda(W_m) - \lambda(W_m \cap E_-) - \lambda(W_m \cap E_+)\big]\\ &\geqslant \tau\big[\lambda(W_m) - \lambda(W_m \cap E_-) - \tau\lambda(W_m \cap G)\big]\\ &\geqslant \tau\big[\lambda(W_m) - \lambda(W_m \cap E_-) - \tau\lambda(W_m)\big]\\ &= \big(\tau - \tau^2\big)\lambda(W_m) - \tau\lambda(W_m \cap E_-).\end{aligned}$$

This implies that

$$\lambda(W_m \cap E_-) > \frac{\tau - \tau^2}{1+\tau}\lambda(W_m) = M^{-1}\frac{\tau - \tau^2}{1+\tau}\lambda\big(Q^{(\ell-1)}\big).$$

Let us consider the case where τ is in the range $0 < \tau \leqslant \sqrt{2} - 1$. Since

$$\lambda\big(Q^{(\ell-1)} \cap E_+\big) \geqslant \lambda\big(Q^{(\ell-1)} \cap E_-\big) \geqslant \lambda(W_m \cap E_-)$$

and $M^{-1}\frac{\tau-\tau^2}{1+\tau} = s$ for this range of values of τ, we see that in this case $W = Q^{(\ell-1)}$ satisfies (7.6). In the remaining case, where $\tau > \sqrt{2} - 1$, we simply observe that the given condition (7.5) also holds when τ is replaced by the smaller number $\sqrt{2} - 1$ and so we can apply the same argument for this smaller value of τ to obtain the required conclusion. This completes the proof of the theorem. □

8. Applying our "geometrical" result. The non-increasing rearrangement of a BMO function

Suppose that a function f of d-variables is in $BMO(D, \mathcal{Q}(D))$ for some cube D in $\mathbb{R}^d$. It is known [3,8] that this implies that the function of one variable f^*, i.e., the non-increasing rearrangement of f, is in $BMO(I, \mathcal{Q}(I))$ and that

$$\|f^*\|_{BMO(I,\mathcal{Q}(I))} \leqslant C\|f\|_{BMO(D,\mathcal{Q}(D))} \tag{8.1}$$

for some constant C depending only on d. (Of course here we are in fact considering only the restriction of f^* to the interval $(0, \lambda(D))$.) Apparently the optimal value of C for which (8.1) holds for all such f is not yet known. Nor is it known yet whether or not C can be chosen to in fact be independent of d. It is known [15] that $C = 1$ when $d = 1$. An analogous result holds for dyadic intervals [14].

In this section we wish to obtain inequalities analogous to (8.1), in terms of the functional of John and Strömberg, namely inequalities of the form

$$\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} \leqslant C\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \tag{8.2}$$

where, as above, I is the interval $I = (0, \lambda(D))$ and s and σ are suitably chosen numbers in $(0, 1/2]$. We will be particularly interested in the cases where D is a cube or a special rectangle and $\mathcal{E}$ is $\mathcal{Q}(D)$ or $\mathcal{W}(D)$. But our results for these will be consequences of an analogous result for more general choices of D and $\mathcal{E}$.

The constant C in our versions of (8.2) will be $C = 1$ and this is apparently the best possible constant for any and every choice of the parameters σ and s in $(0, 1/2]$. We will be able to take our parameter σ to be any number satisfying

$$\frac{1}{2} \geqslant \sigma > \frac{2\sqrt{2}-2}{2\sqrt{2}-1} \approx 0.453082.$$

(As mentioned earlier, values of $\sigma > 1/2$ are not particularly interesting, since the functional $\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))}$ can equal 0 also when f is not a constant function.) Our parameter s will depend on our choice of $\mathcal{E}$. In fact it will be exactly the parameter given by the formula $s = \min\{\frac{3-2\sqrt{2}}{M}, \delta\}$ in Theorem 7.7. Thus, exactly as in Corollary 7.8, we will have $s = 2^{-d}(3-2\sqrt{2})$ if $\mathcal{E}$ is either $\mathcal{Q}(D)$ or $\mathcal{D}(D)$ and D is, respectively a cube, or a dyadic cube, and, furthermore, we will have $s = (3-2\sqrt{2})/2$ if $\mathcal{E} = \mathcal{W}(D)$ and D is a special rectangle.

We should now stress that we will not obtain our versions of the inequality (8.2), which we have spent the last few paragraphs describing, in as much generality as the reader may have been led to expect. We will only obtain them for the very special class of those measurable functions f on D which take only non-negative integer values, and which satisfy $\|f\|^{(\mathbf{J},s)}_{BMO(D,\mathcal{E})} \leqslant 1/2$. But, in view of the results of Section 6, notably Theorem 6.3, this will be sufficient for our subsequent purposes.

It is perhaps surprising, and perhaps even amusing, that, at a certain stage, it will turn out to be sufficient to consider an even more restricted subclass of the very special class of functions just referred to, namely those functions which only assume at most three consecutive integer values, which may just as well be 0, 1 and 2.

All the results which we have just described are immediate consequences of the following theorem (Theorem 8.1). In turn Theorem 8.1 will follow from a more general theorem (Theorem 8.2) which will be formulated after this one.

Remark. In the formulation of both of the following two theorems we consider a collection $\mathcal{E}$ of admissible sets and a particular set $Q \in \mathcal{E}$. Instead of $\mathcal{E}$, we use the subcollection $\mathcal{E}(Q)$ of all sets in $\mathcal{E}$ which are contained in Q. We are essentially forced to do this because our function f is defined only on Q. Obviously, if f happens to be defined on all sets in $\mathcal{E}$ the conclusions of both theorems will remain true if we replace $\mathcal{E}(Q)$ by the larger collection $\mathcal{E}$.

Theorem 8.1. *Let $\mathcal{E}$ be an M-multidecomposable collection for some $M \geqslant 2$. Let δ be a bi-density constant for $\mathcal{E}$. Let Q be a set in $\mathcal{E}$ and let $\mathcal{E}(Q)$ be the collection of all sets in $\mathcal{E}$ which are contained in Q. Suppose that the function $f : Q \to \mathbb{N} \cup \{0\}$ is measurable and satisfies $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E}(Q))} \leqslant 1/2$ for $s = \min\{\delta, \frac{3-2\sqrt{2}}{M}\}$. Let σ be a number in the range*

$$\frac{2\sqrt{2}-2}{2\sqrt{2}-1} < \sigma \leqslant \frac{1}{2}. \tag{8.3}$$

Then the function $f^* : (0, \lambda(Q)) \to [0, \infty)$, i.e., the non-increasing rearrangement of f restricted to the interval $I := (0, \lambda(Q))$ satisfies

$$\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} \leqslant \|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E}(Q))}.$$

We want to obtain this theorem as a consequence of the following somewhat more abstract and general theorem, which is formulated in terms of John–Strömberg pairs (τ, s) for the collection $\mathcal{E}$. By introducing this extra level of abstraction we also make it possible to formulate the consequences of an affirmative answer to Question A in a (hopefully) clearer and more organized way.

Theorem 8.2. *Let $\mathcal{E}$ be a collection of admissible subsets of $\mathbb{R}^d$. Let Q be a set in $\mathcal{E}$ and let $\mathcal{E}(Q)$ be the collection of all sets in $\mathcal{E}$ which are contained in Q. Let $\tau \in (0,1/2)$ and $s \in (0,1/2)$ be such that (τ, s) is a John–Strömberg pair for $\mathcal{E}$. Suppose that the function $f : Q \to \mathbb{N} \cup \{0\}$ is measurable and satisfies $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E}(Q))} \leqslant 1/2$. Let σ be a number in the range*

$$\frac{2\tau}{1+2\tau} < \sigma \leqslant \frac{1}{2}. \tag{8.4}$$

Then the function $f^* : (0, \lambda(Q)) \to [0,\infty)$, i.e., the non-increasing rearrangement of f restricted to the interval $I := (0, \lambda(Q))$, satisfies

$$\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} \leqslant \|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E}(Q))}. \tag{8.5}$$

Remark 8.3. We find it interesting that the “geometric” condition of the existence of a John–Strömberg pair turns out to be in some sense “almost equivalent” to the “analytic” condition expressed by the inequality (8.5). This is revealed by combining Theorem 8.2 with an auxiliary result (Theorem 8.4) which we will defer to the end of this section.

Proofs of Theorems 8.1 and 8.2. Let us first show that Theorem 8.1 is a consequence of Theorem 8.2. Since the conclusions of both theorems are the same, this simply amounts to showing that the conditions imposed on $\mathcal{E}$ and s and σ in Theorem 8.1 suffice to guarantee that $\mathcal{E}$ and s and σ satisfy the hypotheses of Theorem 8.2 for some suitable choice of τ. In fact we will choose $\tau = \sqrt{2} - 1$ so that $\frac{2\tau}{1+2\tau} = \frac{2\sqrt{2}-2}{2\sqrt{2}-1}$. So when, in Theorem 8.1, we require σ to satisfy (8.3), this ensures that σ will be in the range specified in (8.4). It remains only to check that (τ, s) is a John–Strömberg pair for $\mathcal{E}$ when $\tau = \sqrt{2} - 1$ and $s = \min\{\delta, \frac{3-2\sqrt{2}}{M}\}$. But this is exactly what is stated by Theorem 7.7 for these choices of τ and s.

Thus we can now turn to the proof of Theorem 8.2. Since we only have to deal with sets of $\mathcal{E}$ which are contained in Q we may suppose from here onwards that $\mathcal{E} = \mathcal{E}(Q)$.

The fact that all values taken by f are in $\mathbb{N} \cup \{0\}$ readily implies that the same is true for all values of f^*. To explain this more precisely, since $\lambda(Q) < \infty$, we can invoke (3.4) to obtain that

$$\lambda\bigl(\bigl\{x \in Q\colon\ f(x) = m\bigr\}\bigr) = \bigl|\bigl\{t \in \bigl(0, \lambda(Q)\bigr)\colon\ f^*(t) = m\bigr\}\bigr| \quad \text{for all } m \in \mathbb{N} \cup \{0\}. \tag{8.6}$$

Then the fact that

$$\sum_{m=0}^{\infty} \lambda\bigl(\bigl\{x \in Q\colon\ f(x) = m\bigr\}\bigr) = \lambda(Q)$$

implies that the subset of $(0, \lambda(Q))$ where f^* takes non-integer values has measure zero.

Let us first dispose of three easier special cases where we can readily see that (8.5) holds.

The first of these cases is when f takes only one value (on sets of positive measure). Then of course $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} = \|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} = 0$ no matter how we choose σ and s. So we obtain (8.5).

The second case is when f takes only two values (on sets of positive measure). In this case $f = a\chi_A + b\chi_{Q\setminus A}$ for some $A \subset Q$ such that $0 < \lambda(A) < \lambda(Q)$ and where $b < a$. Since (cf. Remark 7.13) s is also a bi-density constant for $\mathcal{E}$, there exists some set $W_0 \in \mathcal{E}$ contained in Q such that

$$\min\{\lambda(W_0 \cap A), \lambda(W_0 \setminus A)\} \geqslant s\lambda(W_0).$$

This means that the restriction of $(f\chi_{W_0})^*$ to the interval $(0, \lambda(W_0))$ is given by the formula $(f\chi_{W_0})^* = a\chi_{(0,r)} + b\chi_{[r,\lambda(W_0))}$ for some number $r \in [s\lambda(W_0), (1-s)\lambda(W_0)]$. For all choices of the number u which satisfy $u \in (0, s\lambda(W_0))$ we have $u \in (0, r)$ and

$$u + (1-s)\lambda(W_0) \in [r, \lambda(W_0)).$$

Consequently, $(f\chi_{W_0})^*(u) = a$ and $(f\chi_{W_0})^*(u + (1-s)\lambda(W_0)) = b$. This implies, by Proposition 4.3, that $\mathbf{J}(f, W_0, s) = \frac{a-b}{2}$ for this particular set W_0. Proposition 4.3 also tells us that $\mathbf{J}(f, W, s) \leqslant \frac{a-b}{2}$ for all other sets W in $\mathcal{E}$, so we conclude that $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} = \frac{a-b}{2}$. Since $1/2$ is a bi-density constant for $\mathcal{Q}(I)$ and $\sigma \in (0, 1/2]$ it follows that σ is also a bi-density constant for $\mathcal{Q}(I)$. So we can show, by applying reasoning to f^* and σ, which is exactly analogous to the reasoning just applied to f and s, that $\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} = \frac{a-b}{2}$. Thus we see that (8.5) holds in this case also.

The third and last of these easier cases is when $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} < 1/2$. We will deal with this case by showing that here f has to be a constant. The fact that f and f^* are integer valued, together with the formula (4.11), tells us that the two functionals $\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))}$ and $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}$ are each infima of appropriate subsets of the set $\{(n-1)/2\colon n \in \mathbb{N}\}$ of non-negative half-integers. Thus these functionals themselves can only take non-negative half-integer values. More explicitly, if $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} < 1/2$, then $\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} = 0$. If f is constant a.e., then so is f^* and, as in the first case, we see, trivially that (8.5) holds. We shall now show that this is the only possibility. Suppose, on the contrary, that f is not a constant a.e. Then let b be the smallest non-negative integer (there must exist at least two such integers) for which the set $B = \{x \in Q\colon f(x) = b\}$ has positive measure. Then the set

$$A = \{x \in Q\colon f(x) > b\} = \{x \in Q\colon f(x) \geqslant b+1\}$$

must also have positive measure and we must have $\lambda(B) + \lambda(A) = \lambda(Q)$.

Let us now define the functions $\varphi : \mathbb{R} \to \mathbb{R}$ and $\psi : \mathbb{R} \to \mathbb{R}$ by $\psi(t) = \max\{t, b\}$ and then $\varphi(t) = \min\{b+1, \psi(t)\}$. Obviously ψ and therefore also φ are both 1-Lipschitz functions and therefore, by Lemma 4.2, we have that

$$\|\varphi \circ f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} \leqslant \|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} = 0. \tag{8.7}$$

We also have $\varphi \circ f = b\chi_B + (b+1)\chi_A$ almost everywhere. Therefore, the same calculation that we did in case (ii) for a function taking *only two different values* a.e., gives here that

$$\|\varphi \circ f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})} = \frac{|(b+1) - b|}{2} = \frac{1}{2}.$$

This contradicts (8.7) and shows that f indeed must be a constant, completing our treatment of this case.

Having disposed of these three cases, we can, from this point onwards, assume that f takes three or more different values on sets of positive measure, and we can also assume that $\|f\|_{BMO(Q,\mathcal{E})}^{(\mathbf{J},s)} = 1/2$. Let us suppose that (8.5) does not hold, i.e., that

$$\|f^*\|_{BMO(I,\mathcal{Q}(I))}^{(\mathbf{J},\sigma)} > 1/2. \tag{8.8}$$

We will complete the proof of Theorem 8.2 by showing that (8.8) leads to a contradiction. It follows from (8.8) that there exists some subinterval $I' = (a,b)$ of I such that $\mathbf{J}(f^*, I', \sigma) > 1/2$. This means that the number $\alpha = 1/2$ is not in the "competition" for the infimum in the formula analogous to (4.9) for $\mathbf{J}(f^*, I', \sigma)$. So, for every $c \in \mathbb{R}$, we have

$$\left|\left\{t \in (a,b)\colon\ \left|f^*(t) - c\right| \leqslant 1/2\right\}\right| \leqslant (1-\sigma)(b-a).$$

If k is any integer and if we choose $c = k + 1/2$, then, since f and f^* only take integer values, we have that

$$\begin{aligned}\left\{t \in (a,b)\colon\ \left|f^*(t) - c\right| \leqslant 1/2\right\} &= \left\{t \in (a,b)\colon\ c - 1/2 \leqslant f^*(t) \leqslant c + 1/2\right\} \\ &= \left\{t \in (a,b)\colon\ f^*(t) \in \{k, k+1\}\right\}.\end{aligned}$$

So, in fact,

$$\left|\left\{t \in (a,b)\colon\ f^*(t) \in \{k, k+1\}\right\}\right| \leqslant (1-\sigma)(b-a) \quad \text{for each } k \in \mathbb{Z}. \tag{8.9}$$

(It may of course happen that this set is empty for some or even most values of k.)

Since

$$(a,b) = \bigcup_{k \in \mathbb{N}} \left\{t \in (a,b)\colon\ f^*(t) < k\right\}$$

and since the set $\{t \in (a,b)\colon\ f^*(t) < 0\}$ is empty, there exists a unique integer $k \in \mathbb{N}$ such that

$$\begin{aligned}\left|\left\{t \in (a,b)\colon\ f^*(t) < k-1\right\}\right| \leqslant \frac{\sigma}{2}(b-a) &< \left|\left\{t \in (a,b)\colon\ f^*(t) < k\right\}\right| \\ &= \left|\left\{t \in (a,b)\colon\ f^*(t) \leqslant k-1\right\}\right|.\end{aligned} \tag{8.10}$$

The interval (a,b) is the union of the three disjoint sets $\{t \in (a,b)\colon\ f^*(t) < k-1\}$, $\{t \in (a,b)\colon\ f^*(t) \in \{k-1, k\}\}$ and $\{t \in (a,b)\colon\ f^*(t) \geqslant k+1\}$. By (8.9) and (8.10), the measure of the union of the first two of these does not exceed $(1 - \frac{\sigma}{2})(b-a)$. Therefore we conclude that

$$\left|\left\{t \in (a,b)\colon\ f^*(t) \geqslant k+1\right\}\right| \geqslant \frac{\sigma}{2}(b-a). \tag{8.11}$$

This implies that there exists some integer $m \geqslant k+1$ for which the set

$$\left\{t \in (a,b)\colon\ f^*(t) = m\right\}$$

has positive measure. Let us also show that the set $\{t \in (a,b)\colon\ f^*(t) = k-1\}$ has positive measure. Its measure satisfies

$$\begin{aligned}\bigl|\bigl\{t \in (a,b)\colon\ f^*(t) = k-1\bigr\}\bigr| &= \bigl|\bigl\{t \in (a,b)\colon\ f^*(t) \leqslant k-1\bigr\}\bigr| - \bigl|\bigl\{t \in (a,b)\colon\ f^*(t) \leqslant k-2\bigr\}\bigr| \\ &= \bigl|\bigl\{t \in (a,b)\colon\ f^*(t) < k\bigr\}\bigr| - \bigl|\bigl\{t \in (a,b)\colon\ f^*(t) < k-1\bigr\}\bigr|\end{aligned}$$

which, by (8.10), is indeed positive.

We now know that on the interval (a,b) the function f^* assumes at least one value strictly larger than k and at least one value strictly less than k. Since f is non-increasing, this means that

$$\bigl\{t \in (a,b)\colon\ f^*(t) = k\bigr\} = \bigl\{t \in \bigl(0, \lambda(Q)\bigr)\colon\ f^*(t) = k\bigr\}. \tag{8.12}$$

Now let us define the three sets

$$E_- = \bigl\{x \in Q\colon\ f(x) \leqslant k-1\bigr\}, \qquad G = \bigl\{x \in Q\colon\ f(x) = k\bigr\} \quad \text{and}$$
$$E_+ = \bigl\{x \in Q\colon\ f(x) \geqslant k+1\bigr\}.$$

By properties of the non-increasing rearrangement, or, more specifically, in view of (8.6), we obtain that the λ measures of these three sets are equal, respectively, to

$$\bigl|\bigl\{t \in \bigl(0, \lambda(Q)\bigr)\colon\ f^*(t) \leqslant k-1\bigr\}\bigr| \quad \text{and} \quad \bigl|\bigl\{t \in \bigl(0, \lambda(Q)\bigr)\colon\ f^*(t) = k\bigr\}\bigr|$$

and $|\{t \in (0,\lambda(Q))\colon\ f^*(t) \geqslant k+1\}|$. Consequently, using (8.12) and then (8.9), we see that $\lambda(G) \leqslant (1-\sigma)(b-a)$. Then (8.10) and (8.11) give us that $\lambda(E_-) \geqslant \frac{\sigma}{2}(b-a)$ and $\lambda(E_+) \geqslant \frac{\sigma}{2}(b-a)$. Therefore we have

$$\min\bigl\{\lambda(E_-), \lambda(E_+)\bigr\} \geqslant \frac{\sigma}{2(1-\sigma)}\lambda(G). \tag{8.13}$$

It is a routine matter to check that the condition (8.4) is equivalent to

$$\tau < \frac{\sigma}{2(1-\sigma)} \leqslant \frac{1}{2}. \tag{8.14}$$

Let us pause for a moment to point out that we have finally reached the *only* step of the proof which needs a non-trivial "geometric" input, namely the fact that (τ, s) is a John–Strömberg pair for $\mathcal{E}$.

To know that we have this fact in the particular case that appears in the formulation of Theorem 8.1 we need to apply our "geometric" Theorem 7.7. To know that we have this fact for other particular collections $\mathcal{E}$, or with better values of the constant s, we would need an affirmative answer to Question A, or to some other question. In the present theorem we have simply "axiomatized the 'geometric' problem away" by invoking a convenient definition.

Now let us resume our formal proof: The estimates (8.14) and (8.13) give us the inequality (7.7) which appears in Definition 7.9. Since we have required that (τ, s) is a John–Strömberg pair for $\mathcal{E}$, this guarantees that there exists a set $W \in \mathcal{E}$ for which $W \subset Q$ and

$$\min\bigl\{\lambda(E_+ \cap W), \lambda(E_- \cap W)\bigr\} \geqslant s\lambda(W). \tag{8.15}$$

Let us now, analogously to what we did in the easy case (iii) above, define the functions $\varphi:\mathbb{R}\to\mathbb{R}$ and $\psi:\mathbb{R}\to\mathbb{R}$ by $\psi(t)=\max\{t,k-1\}$ and then $\varphi(t)=\min\{k+1,\psi(t)\}$. Here again it is obvious that ψ and therefore also φ are both 1-Lipschitz functions. Therefore, again by Lemma 4.2 and by our hypotheses on f, we have that

$$\|\varphi\circ f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}\leqslant\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}\leqslant 1/2. \tag{8.16}$$

Note also that the function $\varphi\circ f$ takes precisely three values, namely $k-1$, k and $k+1$.

Our final step will be to show that the set W which satisfies (8.15) must also satisfy

$$\mathbf{J}(\varphi\circ f,W,s)\geqslant 1. \tag{8.17}$$

This will contradict (8.16) and so show that the assumption (8.8) must be false, and thus will suffice to complete the proof of Theorem 8.2.

To simplify the notation, let us set $g=\varphi\circ f$. As already remarked above, this non-negative function takes only the three values $k-1$, k, and $k+1$. More precisely, when we restrict g to the cube W, it takes these three values, respectively, on the sets $E_-\cap W$, $G\cap W$ and $E_+\cap W$, whose union is W.

Thus the restriction of $(g\chi_W)^*$ to the interval $(0,\lambda(W))$ is given by

$$(g\chi_W)^*(t)=\begin{cases}k+1, & 0<t<a,\\ k, & a\leqslant t<b,\\ k-1, & b\leqslant t<\lambda(W)\end{cases}$$

where $a=\lambda(E_+\cap W)$, $b-a=\lambda(G\cap W)$ and $\lambda(W)-b=\lambda(E_-\cap W)$.

Let $I=[u,u+(1-s)\lambda(W)]$ be an arbitrary closed interval of length $(1-s)\lambda(W)$ which is contained in $(0,\lambda(W))$. Since the left interval $(0,a)$ has length not less than $s\lambda(W)$ we conclude that the left endpoint of I must line in $(0,a)$. Similarly, since the right interval $[b,\lambda(W))$ also has length not less than $s\lambda(W)$, the right endpoint of must lie in $[b,\lambda(W))$. We deduce that

$$(g\chi_W)^*(u)-(g\chi_W)^*\big(u+(1-s)\lambda(Q)\big)=(k+1)-(k-1)=2.$$

Therefore, by Proposition 4.3, we have that $\mathbf{J}(g,W,s)=1$ which establishes (8.17). As already explained above, this suffices to complete the proof of Theorem 8.1. □

We conclude this section by stating the auxiliary result alluded to above in Remark 8.3, which is a sort of converse to Theorem 8.2. We refer to [6, p. 35] for its proof and for further remarks about it.

Theorem 8.4. *(See [6, Theorem 8.5, p. 35].) Let $\mathcal{E}$ be a collection of admissible subsets of $\mathbb{R}^d$. Let s and σ be two given numbers in $(0,1/2)$ and let τ be any number in $(0,1)$ satisfying $0<\sigma<\tau/(1+2\tau)$. Suppose that, for each $Q\in\mathcal{E}$, the inequality*

$$\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))}\leqslant\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E}(Q))}$$

holds for the interval $I=(0,\lambda(Q))$ and for every measurable function $f:Q\to\mathbb{R}$ which assumes only the three values 0, 1 and 2. Here $\mathcal{E}(Q)$ denotes the collection of all those sets in $\mathcal{E}$ which are contained in Q.

Then (τ,s) is a John–Strömberg pair for $\mathcal{E}$.

9. Putting all the pieces together

Now at last we can combine our results from previous sections to obtain our versions of the John–Nirenberg and John–Strömberg inequalities. The following theorem does this. It also explicitly and immediately shows (keeping in mind Remark 7.12) the consequence of an affirmative answer to Question A, thus proving what we claimed at the very beginning of this paper.

The hypotheses imposed on $\mathcal{E}$, Q, τ and s in this theorem are exactly those which were imposed in Theorem 8.2. But here the parameter σ does not need to be explicitly mentioned.

Theorem 9.1. *Let $\mathcal{E}$ be a collection of admissible subset of $\mathbb{R}^d$. Let Q be a set in $\mathcal{E}$ which contains all other sets of $\mathcal{E}$. Let $\tau \in (0,1/2)$ and $s \in (0,1/2)$ be such that (τ,s) is a John–Strömberg pair for $\mathcal{E}$. Then, for every constant r in the range $1 \leqslant r \leqslant \frac{1}{2\tau}$, the inequalities*

$$\lambda\bigl(\bigl\{x \in Q\colon \bigl|f(x)-m\bigr| \geqslant \alpha\bigr\}\bigr) \leqslant \max\{r, 2\sqrt{r}\} \cdot \lambda(Q) \cdot \exp\biggl(-\frac{\alpha \log r}{8\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}}\biggr) \tag{9.1}$$

and

$$\lambda\bigl(\bigl\{x \in Q\colon \bigl|f(x)-m\bigr| \geqslant \alpha\bigr\}\bigr) \leqslant \max\{r, 2\sqrt{r}\} \cdot \lambda(Q) \cdot \exp\biggl(-\frac{\alpha s \log r}{8\|f\|_{BMO(Q,\mathcal{E})}}\biggr) \tag{9.2}$$

hold for every $\alpha \geqslant 0$, every measurable $f : Q \to \mathbb{R}$, and every median m of f on Q.

Remark. The inequality (1.1) which appears at the very beginning of this paper is of course simply (9.2) with $r = \frac{1}{2\tau}$. This is in some sense the most "pertinent" value of r to choose since it gives the best control of the left-hand side of (9.2) when we consider large values of α.

Proof of Theorem 9.1. Let us prepare some ingredients which will later enable us to apply Theorem 6.3. In our application the sets D and E which appear in the statement of Theorem 6.3 will both be taken to equal the set Q specified in the formulation here of Theorem 9.1. Let $f : Q \to \mathbb{N} \cup \{0\}$ be an arbitrary function in the class Φ which is defined in Theorem 6.3. Let σ be a number satisfying the condition (8.4) which is imposed in Theorem 8.2. Then, since f and σ and s satisfy the hypotheses of Theorem 8.2, we can deduce from Theorem 8.2 that

$$\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))} \leqslant \|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}, \tag{9.3}$$

where I is the interval $(0,\lambda(Q))$. The fact that f^* is right continuous and non-increasing on I ensures that, for every $\alpha \geqslant 0$, it satisfies the inequality corresponding to (5.7) which here takes the form

$$\bigl|\bigl\{t \in I\colon f^*(t) - f^*\bigl(\lambda(Q)/2\bigr) \geqslant \alpha\bigr\}\bigr| \leqslant \frac{1-\sigma}{2\sigma} \cdot |I| \cdot \exp\biggl(-\frac{\alpha \log(\frac{1}{\sigma}-1)}{2\|f^*\|^{(\mathbf{J},\sigma)}_{BMO(I,\mathcal{Q}(I))}}\biggr). \tag{9.4}$$

As an element of Φ, the function f must satisfy

$$\lambda\bigl(\bigl\{x \in Q\colon f(x) > 0\bigr\}\bigr) \leqslant \frac{1}{2}\lambda(Q). \tag{9.5}$$

This implies that $f^*(t) = 0$ for all $t > \lambda(Q)/2$ and therefore also that $f^*(\lambda(Q)/2) = 0$. (Use properties (i) and (ii) of Section 3.)

Now, for every $\alpha > 0$ we can apply (3.2) to the left-hand side of (9.4) and use (9.3) to bound the right-hand side of (9.4) from above. This gives us that

$$\lambda\big(\big\{x \in Q\colon\ f(x) \geqslant \alpha\big\}\big) \leqslant \frac{1-\sigma}{2\sigma} \cdot \lambda(Q) \cdot \exp\left(-\frac{\alpha \log(\frac{1}{\sigma}-1)}{2\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}}\right). \tag{9.6}$$

We deduce, using (9.6) when $\alpha > 0$, and using (9.5), together with the fact that $\frac{1-\sigma}{2\sigma} \geqslant \frac{1}{2}$, when $\alpha = 0$, that

$$\lambda\big(\big\{x \in Q\colon\ f(x) > \alpha\big\}\big) \leqslant \frac{1-\sigma}{2\sigma} \cdot \lambda(Q) \cdot \exp\left(-\frac{\alpha \log(\frac{1}{\sigma}-1)}{2\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}}\right)$$

for all $\alpha \geqslant 0$. This last inequality corresponds to the inequality (6.14) of Theorem 6.3, and the fact that it holds for every $f \in \Phi$ is exactly what we need to justify applying Theorem 6.3. Since here the constants C and c of (6.14) are respectively $(1-\sigma)/(2\sigma)$ and $\log(\frac{1}{\sigma}-1)/2$, the formula (6.15) furnished by Theorem 6.3 takes the form

$$\lambda\big(\big\{x \in Q\colon\ \big|f(x)-m\big| \geqslant \alpha\big\}\big) \leqslant 2\max\left\{\frac{1-\sigma}{2\sigma}, \sqrt{\frac{1-\sigma}{\sigma}}\right\} \cdot \lambda(Q) \cdot \exp\left(-\frac{\alpha \log(\frac{1}{\sigma}-1)}{8\|f\|^{(\mathbf{J},s)}_{BMO(Q,\mathcal{E})}}\right).$$

Let us substitute $r = \frac{1}{\sigma} - 1 = \frac{1-\sigma}{\sigma}$ in this inequality. Since σ satisfies (8.4) and since, as already observed during the proof of Theorem 8.2, (8.4) is equivalent to (8.14), it follows that $\tau < 1/2r \leqslant 1/2$ and so $1 \leqslant r < 1/2\tau$. Thus the above inequality shows that (9.1) holds for all $r \in [1, 1/2\tau)$, and therefore also, by continuity, at the endpoint $r = 1/2\tau$. Finally (9.2) follows immediately, in view of (4.8). □

Remark 9.2. Let us consider Theorem 9.1 in the particular case where Q is a special rectangle and $\mathcal{E} = \mathcal{W}(Q)$. Then we know (cf. Corollary 7.8 or Example 7.10) that we can take $\tau = \sqrt{2} - 1$ and $s = (3 - 2\sqrt{2})/2 \approx 0.0857864$. So the parameter r can range between 1 and $1/(2\sqrt{2} - 2) \approx 1.20711$. Consequently, the right-hand side of (9.2) can be, for example,

$$2/\sqrt{2\sqrt{2}-2} \cdot \lambda(Q) \cdot \exp\left(-\frac{\alpha(3-2\sqrt{2})\log(1/(2\sqrt{2}-2))}{16\|f\|'_{BMO}}\right)$$

where $\|f\|'_{BMO}$ is the seminorm defined in (2.12). This expression is approximately equal to $2.197 \cdot \lambda(Q) \cdot \exp(-\frac{0.002\alpha}{\|f\|'_{BMO}})$. Wik obtains a smaller expression, approximately equal to $2\lambda(Q)\exp(-\frac{0.043\alpha}{\|f\|'_{BMO}})$.

Acknowledgments

As already indicated in our text we were offered some interesting remarks by the referee and by Gady Kozma, and we thank them both for these.

References

[1] R.J. Bagby, D.S. Kurtz, Covering lemmas and the sharp function, Proc. Amer. Math. Soc. 93 (1985) 291–296.
[2] R.J. Bagby, D.S. Kurtz, A rearranged good-λ inequality, Trans. Amer. Math. Soc. 293 (1986) 71–81.
[3] C. Bennett, R.A. DeVore, R. Sharpley, Weak-L^∞ and BMO, Ann. of Math. 113 (1981) 601–611.
[4] C. Bennett, R. Sharpley, Weak-type inequalities for H^p and BMO, Proc. Sympos. Pure Math. 35 (I) (1979) 201–229.
[5] J. Bergh, J. Löfström, Interpolation Spaces. An Introduction, Springer, Berlin, 1976.
[6] M. Cwikel, Y. Sagher, P. Shvartsman, A new look at the John–Nirenberg and John–Strömberg theorems for *BMO*. Lecture Notes, arXiv:1011.0766v1 [math.FA].
[7] C. Fefferman, E.M. Stein, H^p spaces of several variables, Acta Math. 129 (1972) 137–193.
[8] A.M. Garsia, E. Rodemich, Monotonicity of certain functionals under rearrangements, Ann. Inst. Fourier (Grenoble) 24 (1974) 67–116.
[9] R.A. Hunt, On $L(p,q)$ spaces, Enseign. Math. 12 (1966) 249–276.
[10] B. Jawerth, A. Torchinsky, Local sharp maximal functions, J. Approx. Theory 43 (1985) 231–270.
[11] F. John, Rotation and strain, Comm. Pure Appl. Math. 14 (1961) 391–413.
[12] F. John, Quasi isometric mappings, in: Semineri 1962–1963 di Analisi Algebra, Geometria e Topologia, vol. II, Rome, 1965.
[13] F. John, L. Nirenberg, On functions of bounded mean oscillation, Comm. Pure Appl. Math. 14 (1961) 415–426.
[14] I. Klemes, A mean oscillation inequality, Proc. Amer. Math. Soc. 93 (1985) 497–500.
[15] A.A. Korenovskii, The connection between mean oscillations and exact exponents of summability of functions, Mat. Sb. 181 (12) (1990) 1721–1727 (in Russian); translation in Math. USSR-Sb. 71 (2) (1992) 561–567.
[16] A.A. Korenovskii, Mean Oscillations and Equimeasurable Rearrangements of Functions, Lect. Notes Unione Mat. Ital., vol. 4, Springer/UMI, Berlin/Bologna, 2007.
[17] A.K. Lerner, Maximal functions with respect to differential bases measuring mean oscillation, Anal. Math. 24 (1998) 41–58.
[18] J. Moser, On Harnack's theorem for elliptic differential equations, Comm. Pure Appl. Math. 14 (1961) 577–591.
[19] F. Nazarov, M. Sodin, A. Vol'berg, The geometric Kannan–Lovász–Simonovits lemma, dimension-free estimates for volumes of sublevel sets of polynomials, and distribution of zeroes of random analytic functions, arXiv:math/0108212v2 [math.CA].
[20] F. Nazarov, M. Sodin, A. Vol'berg, The geometric Kannan–Lovász–Simonovits lemma, dimension-free estimates for the distribution of the values of polynomials, and the distribution of the zeros of random analytic functions, Algebra i Analiz 14 (2) (2002) 214–234 (in Russian); translation in St. Petersburg Math. J. 14 (2) (2003) 351–366.
[21] Y. Sagher, Lecture notes on measure theory and real analysis, in preparation.
[22] Y. Sagher, P. Shvartsman, On the John–Strömberg–Torchinsky characterization of *BMO*, J. Fourier Anal. Appl. 4 (1998) 521–548.
[23] Y. Sagher, P. Shvartsman, Rearrangement-function inequalities and interpolation theory, J. Approx. Theory 119 (2002) 214–251.
[24] P. Shvartsman, The K-functional of the pair $(L_\infty(w), BMO)$, in: Function Spaces, Interpolation Spaces, and Related Topics, Haifa, 1995, in: Israel Math. Conf. Proc., vol. 13, Bar-Ilan Univ., Ramat Gan, 1999, pp. 183–203.
[25] L. Slavin, V. Vasyunin, Sharp results in the integral-form John–Nirenberg inequality, Trans. Amer. Math. Soc. 363 (2011) 4135–4169.
[26] L. Slavin, V. Vasyunin, Sharp L^p estimates on *BMO*, arXiv:1110.1771v1 [math.CA].
[27] E.M. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. With the Assistance of T.S. Murphy, Monographs in Harmonic Analysis, III, Princeton Math. Ser., vol. 43, Princeton University Press, Princeton, NJ, 1993, xiv+695 pp.
[28] J.-O. Strömberg, Bounded mean oscillation with Orlicz norms and duality of Hardy spaces, Indiana Math. J. 28 (1979) 511–544.
[29] V. Vasyunin, A. Vol'berg, Sharp constants in the classical weak form of the John–Nirenberg inequality, preprint.
[30] I. Wik, On John and Nirenberg's theorem, Ark. Mat. 28 (1990) 193–200.