Vol. 102, No. 3 DUKE MATHEMATICAL JOURNAL © 2000

BMO FOR NONDOUBLING MEASURES

J. MATEU, P. MATTILA, A. NICOLAU, AND J. OROBITG

1. Introduction. The Calderón-Zygmund theory of singular integrals has been traditionally considered with respect to a measure satisfying a doubling condition. Recently, Tolsa [T] and, independently, Nazarov, Treil, and Volberg [NTV] have shown that this standard doubling condition was not really necessary. Likewise, in the homogeneous spaces setting, functions of bounded mean oscillation, BMO, and its predual H^1, the atomic Hardy space, play an important role in the theory of singular integrals.

This note is an attempt to find good substitutes for the spaces BMO and H^1 when the underlying measure is nondoubling. Our hope was that we would have been able to prove some results of Tolsa, Nazarov, Treil, and Volberg, via BMO-H^1 interpolation, but in this respect we were unsuccessful.

Let μ be a nonnegative Radon measure on $\mathbb{R}^n$. A function $f \in L^1_{\mathrm{loc}}(\mu)$ is said to belong to BMO(μ) if the inequality

$$\int_Q |f(x) - f_Q| \, d\mu(x) \le C\mu(Q) \tag{1}$$

holds for all cubes Q with sides parallel to the coordinate axes; $f_Q = (\mu(Q))^{-1} \int_Q f \, d\mu$ denotes the mean value of f over the cube Q. The smallest bound C for which (1) is satisfied is then taken to be the "norm" of f in this space, and it is denoted by $\|f\|_*$.

One says that BMO(μ) has the John-Nirenberg property when there exist positive constants c_1 and c_2 so that whenever $f \in \mathrm{BMO}(\mu)$, then for every $\lambda > 0$ and every cube Q with sides parallel to the coordinate axes, one has

$$\mu\big(\{x \in Q : |f(x) - f_Q| > \lambda\}\big) \le c_1 e^{-c_2\lambda/\|f\|_*} \mu(Q).$$

It is well known that if a measure μ is doubling (i.e., there exists a constant $C = C(\mu)$ such that $\mu(2Q) \le C\mu(Q)$ for all cubes Q), then it satisfies the John-Nirenberg inequality. We give examples of nonnegative Radon measures on $\mathbb{R}^n$ which do not

Received 8 January 1999. Revision received 3 August 1999.

1991 *Mathematics Subject Classification*. Primary 26B35, 42B25, 46E30.

Mateu, Nicolau, and Orobitg's work partially supported by grant numbers PB98-0872 and PB96-1183 of the Dirección General de Enseñanza Superior and by grant number 1998 SGR00052 of the Comissió Interdepartamental de Recerca i Innovació Tecnològica.

Mattila gratefully acknowledges the hospitality of the Centre de Recerca Matemàtica at the Universitat Autònoma de Barcelona during the preparation of this work.

have the John-Nirenberg property. On the other hand, we show that for a large class of measures, the John-Nirenberg property holds.

THEOREM 1. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$. Assume that for every hyperplane L, orthogonal to one of the coordinate axes, $\mu(L) = 0$. Suppose that f is in* BMO(μ). *Then there exist constants c_1 and c_2, independent of f, so that for every $\lambda > 0$ and every cube Q with sides parallel to the coordinate axes, one has*

$$\mu\big(\{x \in Q : |f(x) - f_Q| > \lambda\}\big) \le c_1 \exp\left(\frac{-c_2\lambda}{\|f\|_*}\right)\mu(Q).$$

When the measure μ is the Lebesgue measure (or any other doubling measure) on $\mathbb{R}^n$, the John-Nirenberg theorem follows from a stopping time argument using dyadic cubes. The estimate that is needed is

$$|f_{2Q} - f_Q| \le C\|f\|_*,$$

which follows from the fact that $\mu(2Q)/\mu(Q)$ is bounded from above. When the measure μ is not doubling our approach, following an idea of Wik [W], it is based on the following covering lemma.

COVERING LEMMA. *Let μ be a positive Radon measure in $\mathbb{R}^n$ such that for every hyperplane L, orthogonal to one of the coordinate axes, $\mu(L) = 0$. Let E be a subset of $\mathbb{R}^n$, and let ρ be a real number in $(0, 1)$. Suppose that E is contained in a cube Q_0, with sides parallel to the coordinate axes, and suppose that $\mu(E) \le \rho\mu(Q_0)$. Then there exists a sequence $\{Q_j\}$ of cubes with sides parallel to the coordinate axes and contained in Q_0 such that*

(a) $\mu(Q_j \cap E) = \rho\mu(Q_j)$;
(b) *the family $\{Q_j\}$ is almost disjoint with constant $B(n)$, that is, every point of $\mathbb{R}^n$ belongs to at most $B(n)$ cubes Q_j;*
(c) $E' \subset \bigcup_j Q_j$, *where E' is the set of μ-density points of E.*

We say that x is a μ-density point of E when $\lim_{r\to 0} \mu(Q(x,r) \cap E)\mu(Q(x,r))^{-1} = 1$, where $Q(x,r)$ denotes the cube centered at x and sidelength r. The assumption on the measure μ means that, given a cube Q, with $2\mu(Q) < \mu(\mathbb{R}^n)$, there exists a cube $\tilde{Q} \supset Q$ such that $\mu(\tilde{Q}) = 2\mu(Q)$. The proof of our covering lemma uses a variant of the well-known Besicovitch covering theorem.

At first sight, the assumption on the measure in the statement of Theorem 1 seems quite restrictive, but the next result disproves this feeling.

THEOREM 2. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$. Assume that for any point $p \in \mathbb{R}^n$, $\mu(\{p\}) = 0$. Then there exists an orthonormal system $\{e_1, \dots, e_n\}$ so that for every hyperplane L with normal vector e_i $(i \in \{1, \dots, n\})$, $\mu(L) = 0$.*

As in the case of the Lebesgue measure, the John-Nirenberg theorem gives the duality $H^1(\mu)-\mathrm{BMO}(\mu)$, where $H^1(\mu)$ is a natural atomic Hardy space. Let $L^\infty_c(\mu)$ be the space of bounded functions with compact support. One can then prove the following interpolation result.

THEOREM 3. *Let μ be a nonnegative Radon measure on $\mathbb{R}^n$. Assume that for every hyperplane L, orthogonal to one of the coordinate axes, $\mu(L)=0$. Let T be a sublinear operator that is bounded from $L^\infty_c(\mu)$ to* $\mathrm{BMO}(\mu)$ *and from $H^1(\mu)$ to $L^1(\mu)$. Then T extends boundedly to every $L^p(\mu)$, $1<p<\infty$.*

As in the case of a doubling measure, the proof follows easily from a Calderón-Zygmund decomposition and from the L^p-estimates for the sharp maximal function

$$f^\#(x)=\sup\frac{1}{\mu(Q)}\int_Q|f-f_Q|\,d\mu,$$

where the supremum is taken over all cubes centered at $x\in\mathbb{R}^n$. However, as Joan Verdera pointed out to us, Theorem 3 is quite unsatisfactory. Roughly speaking, no interesting operator T maps $H^1(\mu)$ to $L^1(\mu)$, when the measure μ is not doubling. We present an example to illustrate this phenomenon.

Finally, we compare $\mathrm{BMO}(\mu)$, defined as above with cubes whose sides are parallel to the coordinate axes, with $\mathrm{BMO}(\mu)$ defined with balls: $f\in\mathrm{BMO}_b(\mu)$ if there exists $C<\infty$ such that for every ball $B\subset\mathbb{R}^n$ there is $a\in\mathbb{R}$ such that

$$\int_B|f-a|\,d\mu\le C\mu(B). \tag{2}$$

Recall that if μ is a doubling measure, the spaces $\mathrm{BMO}(\mu)$ and $\mathrm{BMO}_b(\mu)$ coincide. In our setting of nondoubling measures, the situation is quite different. Precisely, we have the following theorems.

THEOREM 4. *There exists an absolutely continuous measure μ in $\mathbb{R}^2$ and $f\in L^1(\mu)$ such that $f\in\mathrm{BMO}_b(\mu)$ but for any choice of the coordinate axes $f\notin\mathrm{BMO}(\mu)$.*

THEOREM 5. *There exists an absolutely continuous measure μ in $\mathbb{R}^2$ and $f\in L^1(\mu)$ such that $f\in\mathrm{BMO}(\mu)$ for all choices of the coordinate axes but $f\notin\mathrm{BMO}_b(\mu)$.*

In any case, the question that arises is: May the space BMO_b be a good choice for dealing with functions of bounded mean oscillation? For $1<p<\infty$, we say that $f\in\mathrm{BMO}^p_b(\mu)$ if

$$\int_B|f-a|^p\,d\mu\le C\mu(B),$$

with C, B, and a as above. It is clear that $\mathrm{BMO}_b(\mu)=\mathrm{BMO}^p_b(\mu)$ if the John-Nirenberg inequality holds. From this point of view, we realize that $\mathrm{BMO}(\mu)$ is better than $\mathrm{BMO}_b(\mu)$.

THEOREM 6. *There exists an absolutely continuous measure μ in $\mathbb{R}^2$ and $f \in L^1(\mu)$ such that $f \in \mathrm{BMO}_b(\mu)$ but $f \notin \mathrm{BMO}_b^p(\mu)$ for all $p > 1$. In particular, the John-Nirenberg inequality fails.*

This result may seem surprising. However, two elementary geometric facts, which are important in our analysis, distinguish between dealing with balls or cubes. First, a cube may be covered by a finite number of subcubes, while for balls a countable number of subballs is needed. Second, when intersecting two cubes, the sections have equal diameter, while for balls, the length of the sections may decay exponentially when the balls become nearly disjoint. The first fact provides the necessary covering properties that are used in the proof of Theorem 2. The decay mentioned in the second one gives some extra help to construct functions in BMO_b, which do not fulfill the John-Nirenberg inequality. It is worth mentioning that if instead of cubes or balls, one considers regular polygons of N sides, in the definition of BMO, the analogue of Theorem 2 holds. However, the constants c_1, c_2 depend on N.

The paper is organized as follows. Section 2 contains the proof of Theorems 1 and 2, as well as an example of a Radon measure for which the John-Nirenberg estimate does not hold. Section 3 is devoted to the Calderón-Zygmund decomposition, the L^p estimates for the sharp maximal function, and the proof of Theorem 3. Section 4 contains the proof of Theorems 4, 5, and 6. Finally, we include an appendix with a proof of the John-Nirenberg theorem on spaces of homogeneous type, because it is not easy to find in the literature a proof for the Lebesgue measure which immediately generalizes to doubling measures.

2. John-Nirenberg inequality. The main goal of this section is to prove Theorem 1; that is, the John-Nirenberg property holds for a wide class of measures. We also give an example of a measure μ and a function f, for which $f \in \mathrm{BMO}(\mu)$, but f doesn't satisfy the John-Nirenberg inequality.

In the case of the Lebesgue measure m in $\mathbb{R}^n$, the John-Nirenberg estimate follows from a stopping time argument that uses dyadic cubes. A trivial but essential fact is that for any cube Q, one has $m(2Q) \le Cm(Q)$, where C is a constant. Then if $f \in \mathrm{BMO}(m)$, it follows that

$$|f_{2Q} - f_Q| \le C,$$

which is the estimate that is needed in the stopping time argument.

When the measure μ is supported in the real line and has no atoms, one can prove Theorem 1 along the same lines. The only modification that is needed consists of replacing the usual dyadic grid by a μ-dyadic grid, which is constructed in the following way. Given an interval I, the first generation $G_1(I)$ consists of the two disjoint subintervals I_+, I_- of I satisfying $\mu(I_+) = \mu(I_-) = \mu(I)/2$. The second generation $G_2(I)$ is $G_1(I_+) \cup G_1(I_-)$. Next generations are defined recursively. One can now use the usual stopping time argument to prove the John-Nirenberg estimate for such measures.

To prove Theorem 1 for $n > 1$ we need the following Besicovitch covering theorem.

BESICOVITCH COVERING THEOREM. *Let A be a subset of $\mathbb{R}^n$, and let $\mathcal{R}$ be a family of rectangles with sides parallel to the coordinate axes, such that each point of A is the center of some rectangle of $\mathcal{R}$. Assume that A is a bounded set or that the diameters of rectangles of $\mathcal{R}$ are bounded. Moreover, suppose that the ratio of any two sidelengths of a rectangle of $\mathcal{R}$ is bounded by* 2.

Then there is a finite or countable collection of rectangles $R_j \in \mathcal{R}$ such that they cover A, and every point of $\mathbb{R}^n$ belongs to at most $B(n)$ rectangles R_i, where $B(n)$ is an integer depending only on n.

For this result, see [G]; also the proof given in [M] for balls can be easily modified.

Proof of covering lemma. For any $x \in Q_0$ and for $r > 0$ satisfying $r \leq \ell(Q_0)$, we define $\tilde{Q}(x,r)$ as the unique cube (parallel to the coordinate axes) with sidelength r, containing x, contained in Q_0 and with center y closest to x (see Figure 1).

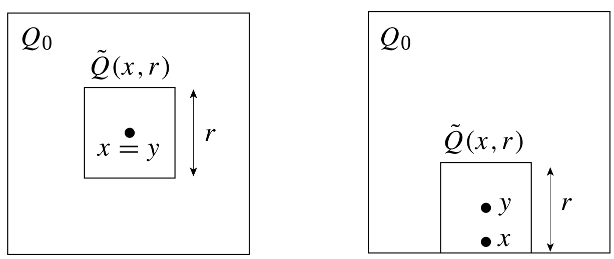

FIGURE 1

If $x \in Q_0 \cap E'$, we define the function $h_x(r) = \mu(\tilde{Q}(x,r) \cap E)/\mu(\tilde{Q}(x,r))$ for $0 < r \leq \ell(Q_0)$. Clearly, this is a continuous function satisfying $h_x(\ell(Q_0)) \leq \rho$ by hypothesis and $\lim_{r \to 0} h_x(r) = 1$, since x is a density point. Consequently, there exists a positive number r_x such that $h_x(r_x) = \rho$. Hence, for any $x \in Q_0 \cap E'$ we define $Q(x) = \tilde{Q}(x, r_x)$. Now, we could not apply the Besicovitch covering theorem, because x may not be the center of $Q(x)$. To circumvent this obstacle, for any cube $Q(x)$ we define the rectangle $R(x)$ in $\mathbb{R}^n$ as the unique rectangle in $\mathbb{R}^n$ centered on x such that $R(x) \cap Q_0 = Q(x)$. Denote by $\mathcal{R}$ this family of rectangles (see Figure 2).

It is an easy computation to check that the ratio of any two sidelengths of a rectangle in $\mathcal{R}$ is bounded by 2. So, by the Besicovitch covering theorem we have a countable

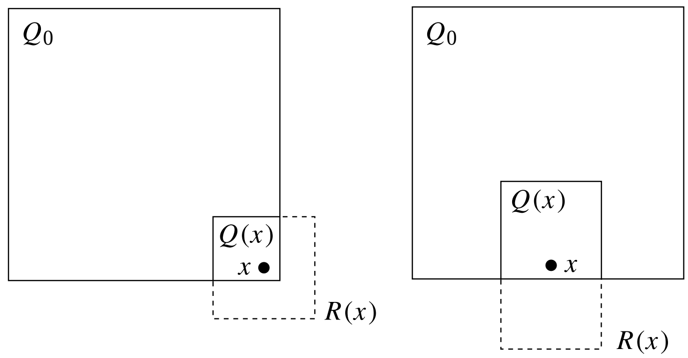

FIGURE 2

collection of rectangles $R_j \in \mathcal{R}$ such that they cover $E' \cap Q_0$, and every point of $\mathbb{R}^n$ belongs at most to $B(n)$ rectangles R_i. Then, if we take the family of cubes $\mathcal{Q} = \{Q(x) : R(x) = R_j \text{ for some } j\}$, it is clear that $\mathcal{Q}$ is a countable family of cubes satisfying properties (a), (b), and (c). □

Now, the covering lemma is one of the main ingredients we use to prove Theorem 1.

Proof of Theorem 1. We may assume that $\|f\|_* = 1$. Let Q_0 be an arbitrary cube, with sides parallel to the coordinate axes. Put, for any integer $k \geq 2$,

$$E_k = \{x \in Q_0 : f(x) - f_{Q_0} \geq k\},$$
$$S_k = \{x \in Q_0 : |f(x) - f_{Q_0}| \geq k\}.$$

Clearly, $\mu(E_k) \leq \mu(S_k) \leq \mu(Q_0)/k$. In particular, $\mu(E_k) \leq \mu(Q_0)/2$ and by the covering lemma we can cover the set E'_k, $k \geq 2$, with a sequence $\{Q_{k,j}\}$ of almost disjoint cubes with constant $B(n)$, parallel to the coordinate axes, such that

$$\mu\big(E_k \cap Q_{k,j}\big) = \frac{1}{2}\mu(Q_{k,j}).$$

Therefore,

$$\mu\big(E_k^c \cap Q_{k,j}\big) = \frac{1}{2}\mu(Q_{k,j}), \tag{3}$$

and

$$\mu(E_k) \cong \sum_j \mu\big(Q_{k,j} \cap E_k\big) \cong \sum_j \mu(Q_{k,j}). \tag{4}$$

On the other hand, for any $\ell > 0$, one has

$$1 = \|f\|_* \geq \frac{1}{2\big(\mu(Q_{k,j})\big)^2} \int_{Q_{k,j}} \left\{ \int_{Q_{k,j}} |f(y) - f(x)|\, d\mu(x) \right\} d\mu(y)$$
$$\geq \frac{1}{\big(\mu(Q_{k,j})\big)^2} \int_{E_{k+\ell} \cap Q_{k,j}} d\mu(y) \left\{ \int_{E_k^c \cap Q_{k,j}} |f(y) - f(x)|\, d\mu(x) \right\}$$
$$\geq \ell \frac{\mu\big(E_{k+\ell} \cap Q_{k,j}\big)\mu\big(E_k^c \cap Q_{k,j}\big)}{\big(\mu(Q_{k,j})\big)^2} = \ell \frac{\mu\big(E_{k+\ell} \cap Q_{k,j}\big)}{2\mu(Q_{k,j})},$$

where the last inequality comes from (3). So, one obtains $\mu(E_{k+\ell} \cap Q_{k,j}) \leq (2/\ell)\mu(Q_{k,j})$. This inequality and (4) give

$$\mu(E_{k+\ell}) \leq \sum_j \mu\big(Q_{k,j} \cap E_{k+\ell}\big) \leq \frac{2}{\ell} \sum_j \mu(Q_{k,j}) \leq \frac{C}{\ell} \mu(E_k).$$

Similar arguments can be used to obtain the same estimate for the sets $\{x \in Q_0 : f(x) - f_Q \leq k\}$, $k < 0$. The estimates combine to

$$\mu(S_{k+\ell}) \leq \frac{A}{\ell} \mu(S_k),$$

where A is an integer depending only on the dimension.

We take $\ell = 2A$ and find

$$\mu(S_{k+2A}) \leq \frac{1}{2} \mu(S_k),$$

which, for any positive integer p implies:

$$\mu\big(S_{2+2Ap}\big) \leq 2^{-p} \mu(S_2) \leq 2^{-p} \mu(Q_0),$$

from which the conclusion of the theorem follows. Observe that the constants c_1 and c_2, in the statement of Theorem 1, do not depend on μ; they only depend on n. □

Now, we give a Radon measure for which the John-Nirenberg inequality does not hold. We consider the case $n = 1$. Let $\mu = \sum_{n \geq 1} (1/2^{n^2}) \delta_{1/n}$, where $\delta_{1/n}$ is a Dirac mass in the point $1/n$, and let $f(1/n) = 2^n$. To show that f belongs to BMO(μ), it is enough to consider intervals $I = [1/N_2, 1/N_1]$, where N_1 and N_2 are positive integers and N_2 can also be infinity. Obviously,

$$\mu(I) = \sum_{n=N_1}^{N_2} \frac{1}{2^{n^2}} \cong \frac{1}{2^{N_1^2}}.$$

Therefore,

$$\frac{1}{\mu(I)}\int_I \left|f-2^{N_1}\right| d\mu \cong 2^{N_1^2} \sum_{n=N_1+1}^{N_2} \frac{1}{2^{n^2}}\left|2^n-2^{N_1}\right| \le 2^{N_1^2} \sum_{n=N_1+1}^{N_2} \frac{2^n}{2^{n^2}} \le \frac{1}{2^{N_1}}.$$

Consequently, $f \in \mathrm{BMO}(\mu)$. In order to verify that f does not satisfy the John-Nirenberg inequality, one can see that for $t = 2^N$, where N is a large positive integer and $I = [0, 1]$,

$$\mu\left\{x \in I : |f(x) - f_I| > t\right\} \cong \mu\left(\left[0, \frac{1}{N}\right]\right) = \sum_{n \ge N} \frac{1}{2^{n^2}} \cong \frac{1}{2^{N^2}}.$$

On the other hand, $2^{-tC}\mu(I)$ is of order $2^{-C\cdot 2^N}$. Hence, the John-Nirenberg inequality holds only if $1/2^{N^2} \le B2^{-C\cdot 2^N}$, for some constant B, but this inequality fails when N is big enough.

The same construction can be repeated in the plane to get a continuous example. We take a family of segments L_n, with endpoints $(1/n, 0)$ and $(1/n, 1)$, and we define a measure $\mu = \sum_{n\ge 1}(1/2^{n^2})\mathcal{H}^1|_{L_n}$, where $\mathcal{H}^1|_{L_n}$ is the 1-dimensional Hausdorff measure on L_n. Taking the function f such that $f|_{L_n} = 2^n$, one can check that $f \in \mathrm{BMO}(\mu)$, but the John-Nirenberg property is not satisfied.

However, observe that in accordance with Theorem 2 rotating the coordinate axes, the corresponding $\mathrm{BMO}(\mu)$, where μ is the above measure, has the John-Nirenberg property.

Proof of Theorem 2. We now show that if μ is a continuous Radon measure on $\mathbb{R}^n$, then we can choose the coordinate axes in such a way that $\mu(\partial Q) = 0$ for all cubes Q with sides parallel to the axes. We first give the very elementary argument in the plane.

So let μ be a continuous Radon measure on $\mathbb{R}^2$. Let $\mathcal{L}$ be the set of the lines L through the origin such that $\mu(L') > 0$ for some line L' parallel to L. We claim that $\mathcal{L}$ is at most countable. Otherwise there exist $\epsilon > 0$, $R < \infty$, and distinct lines $L_1, L_2, \ldots \in \mathcal{L}$ such that for some lines L_i' parallel to L_i and $A_i = L_i' \cap B(0, R)$, we have $\mu(A_i) \ge \epsilon$ for all i. But for $i \ne j$, $A_i \cap A_j$ is either empty or a singleton, whence $\mu(A_i \cap A_j) = 0$. Thus $\mu(B(0, R)) \ge \mu(\bigcup_i A_i) = \sum \mu(A_i) = \infty$, which is a contradiction.

Since $\mathcal{L}$ is at most countable, so is the set $\mathcal{L}^\perp$ of the orthogonal complements $L^\perp$, $L \in \mathcal{L}$. Choosing $L \notin \mathcal{L} \cup \mathcal{L}^\perp$, the lines L and $L^\perp$ give the desired coordinate axes.

To prove our claim in $\mathbb{R}^n$, it seems to be convenient (although not necessary) to use invariant measures on Grassmannians (see, e.g., [M, Chapter 3] for them). Let $G(n, m)$ be the set of all m-dimensional linear subspaces of $\mathbb{R}^n$, and let $\gamma_{n,m}$ be the unique orthogonally invariant Radon probability measure on it. We also allow $m = 0$; then $G(n, 0) = \{0\}$ and $\gamma_{n,0} = \delta_0$. For $V \in G(n, m)$ and $0 \le k < m$, we let $G(V, k)$

be the space of k-dimensional linear subspaces of V, and we let $\gamma_{V,k}$ be the natural measure on it.

Let μ be a continuous Radon measure on $\mathbb{R}^n$. Denote

$$G_m = \{V \in G(n,m) : \mu(V+x) > 0 \text{ for some } x \in \mathbb{R}^n\}.$$

We leave it to the reader to check that G_m is a Borel set.

LEMMA 1. *Let* $\gamma_{n,n-1}(G_{n-1}) = 0$.

Proof. We prove that if $0 < m < n$ and if $\gamma_{n,m}(G_m) > 0$, then $\gamma_{n,m-1}(G_{m-1}) > 0$. Thus if $\gamma_{n,n-1}(G_{n-1}) > 0$, induction gives $\gamma_{n,0}(G_0) > 0$, which means that μ has atoms and gives a contradiction.

So suppose $\gamma_{n,m}(G_m) > 0$. We have that

$$\gamma_{n,m}(G_m) = \int \gamma_{V^\perp,1}\big(\{L \in G(V^\perp,1) : L+V \in G_m\}\big)\, d\gamma_{n,m-1}V.$$

This follows from the uniqueness of $\gamma_{n,m}$ since also the right-hand side defines an orthogonally invariant Radon probability measure on $G(n,m)$. Thus the set of those $V \in G(n,m-1)$ for which

$$\gamma_{V^\perp,1}\big(\{L \in G(V^\perp,1) : L+V \in G_m\}\big) > 0 \tag{5}$$

has positive $\gamma_{n,m-1}$ measure. Hence it suffices to show that (5) implies that $V \in G_{m-1}$.

Let $V \in G(n,m-1)$ satisfy (5). For every $L \in G(V^\perp,1)$ such that $L+V \in G_m$, there is $x(L) \in \mathbb{R}^n$ such that $\mu(L+V+x(L)) > 0$. Since there are uncountably many such lines L, we must have for some $L \neq L'$ (cf. the argument for $\mathbb{R}^2$ above),

$$\mu\big((L+V+x(L)) \cap (L'+V+x(L'))\big) > 0.$$

But $(L+V+x(L)) \cap (L'+V+x(L')) \subset V+x$ for some $x \in \mathbb{R}^n$ by easy linear algebra. Consequently, $V \in G_{m-1}$, and Lemma 1 is proved. □

LEMMA 2. *If* $G \subset G(n,n-1)$ *and if* $\gamma_{n,n-1}(G) = 0$, *then there are coordinate axes in such a way that* $V \notin G$ *for every coordinate hyperplane* V.

Proof. We use induction on n. If $n = 2$, then $\gamma_{2,1}(G^\perp) = 0$ and L, $L^\perp$ will do for $L \notin G \cup G^\perp$. Suppose the lemma holds in $\mathbb{R}^{n-1}$. As above, by the uniqueness of $\gamma_{n,n-1}$, we have

$$\gamma_{n,1}\big(\{L \in G(n,1) : L^\perp \in G\}\big) = \gamma_{n,n-1}(G) = 0$$

and

$$\int \gamma_{L^\perp,n-2}\big(\{V \in G(L^\perp,n-2) : L+V \in G\}\big)\, d\gamma_{n,1}L = \gamma_{n,n-1}(G) = 0.$$

Thus, we can choose $L \in G(n,1)$ such that $L^{\perp} \notin G$ and

$$\gamma_{L^{\perp},n-2}\big(\{V \in G(L^{\perp}, n-2) : L+V \in G\}\big) = 0.$$

By our induction hypothesis, we can choose coordinate axes $L_1, \dots, L_{n-1}$ for $L^{\perp}$ such that for every coordinate $(n-2)$-plane V, we have $L+V \notin G$. Then $L_1, \dots, L_{n-1}, L$ are the required axes in $\mathbb{R}^n$. □

Combining Lemmas 1 and 2 we see that given a continuous μ there are axes such that $\mu(V) = 0$ for all hyperplanes parallel to the coordinate planes, as required. □

The last part of this section is devoted to the introduction of the predual of $\mathrm{BMO}(\mu)$: $H^1(\mu)$. In the case of the Lebesgue measure m, $\mathrm{BMO}(m)$ can be viewed very naturally as the dual of an atomic space $H^{1,\infty}(m)$ (see [J, 3.II]). In this section we claim that for the measures μ satisfying the hypothesis of Theorem 1, one can consider the atomic space $H^{1,\infty}(\mu)$, and its dual space is $\mathrm{BMO}(\mu)$.

A function a is called a p-atom, $1 < p \le \infty$, if there exists a cube Q such that

(i) $a \in L^p(\mu)$, $\|a\|_{L^p(\mu)} \le \mu(Q)^{(1/p)-1}$;

(ii) $\operatorname{spt} a \subset Q$;

(iii) $\int_Q a\, d\mu = 0$.

In the case $\mu(\mathbb{R}^n) < \infty$, the constant $1/(\mu(\mathbb{R}^n))$ is also considered an atom.

Thus, we define a Banach space $H^{1,p}(\mu)$ in the following way: $f \in H^{1,p}(\mu)$ if and only if there exist $\lambda_i \in \mathbb{R}$ and p-atoms a_i such that $\sum |\lambda_i| < \infty$ and $f = \sum_i \lambda_i a_i$. For $f \in H^{1,p}(\mu)$, we define its norm $\|f\|_{H^{1,p}(\mu)}$ to be $\inf \sum |\lambda_j|$, where the infimum is taken over all sequences $(\lambda_i)_{i\in I}$ occurring in such an atomic decomposition of f. When a is a p-atom one can check that $\|a\|_{L^1(\mu)} \le 1$, and so $H^{1,p}(\mu)$ is continuously embedded in $L^1(\mu)$. It is also an easy computation to verify that $H^{1,p_2}(\mu)$ is continuously embedded in $H^{1,p_1}(\mu)$ if $1 < p_1 \le p_2 < \infty$.

Now, we can state the following duality result.

THEOREM 7. *Let μ be a nonnegative Radon measure in $\mathbb{R}^n$. Assume that for every hyperplane L, orthogonal to one of the coordinate axes, $\mu(L) = 0$. Then $(H^{1,p}(\mu))^* = \mathrm{BMO}(\mu)$ if $1 < p \le \infty$.*

In these notes we don't prove this result because the main argument follows using the same ideas as in the case of the Lebesgue measure. For a good exposition of this result in the case of the Lebesgue measure, see [J, 3.II].

By the above theorem we have a family of Banach spaces $H^{1,p}(\mu)$, with the same dual. Hence, since they are continuously embedded, one in the other, we can conclude that they coincide, and so we can define $H^1(\mu) = H^{1,p}(\mu)$ for any p, $1 < p \le \infty$.

To finish this section we want to remark that the main ingredient in the proof of the above result is the John-Nirenberg inequality, and so Theorem 7 can be stated in a more general setting. That is, if we have a positive Radon measure μ in $\mathbb{R}^n$ (for this measure we define $\mathrm{BMO}(\mu)$, and $\mathrm{BMO}(\mu)$ satisfies the John-Nirenberg property), then we can define $H^1(\mu)$ and $(H^1(\mu))^* = \mathrm{BMO}(\mu)$.

3. Interpolation. This section is devoted to prove a Calderón-Zygmund decomposition for functions in $L^1(\mu)$ and a result on interpolation of operators. This follows from estimating the $L^p(\mu)$-norm of a function by the $L^p(\mu)$-norm of its sharp maximal function.

LEMMA 3 (Calderón-Zygmund decomposition). *Let μ be a nonnegative Radon measure in $\mathbb{R}^n$ such that $\mu(L) = 0$ for every hyperplane L orthogonal to one of the coordinate axes. Suppose we are given a function $f \in L^1(\mu)$ and a positive number λ, with $\lambda > \|f\|_1/(\mu(\mathbb{R}^n))$. There exists a decomposition of f, $f = g+b$, and a sequence of cubes $\{Q_j\}$, so that*

(i) *$|g(x)| \le C\lambda$ for μ-a.e. x;*
(ii) *$b = \sum b_j$, where each b_j is supported in Q_j, $\int b_j\, d\mu = 0$, and $\int |b_j|\, d\mu \le 4\lambda\mu(Q_j)$;*
(iii) *$\{Q_j\}$ is almost disjoint with constant $B(n)$;*
(iv) *$\sum_j \mu(Q_j) \le C/\lambda \int_{\bigcup_j Q_j} |f|\, d\mu$.*

Proof. Let Mf be the centered Hardy-Littlewood maximal function

$$Mf(x) = \sup_{r>0} \frac{1}{\mu\big(Q(x,r)\big)} \int_{Q(x,r)} |f(y)|\, d\mu(y),$$

where $Q(x,r)$ is the cube with center x and sidelength r.

For each $x \in E_\lambda = \{x : Mf(x) > \lambda\}$ we consider a cube $Q(x, r_x)$ such that $\int_{Q(x,r_x)} |f|\, d\mu > \lambda\mu(Q(x,r_x))$. Then we proceed as in the proof of covering lemma. We define

$$D(r) := \frac{1}{\mu\big(Q(x,r)\big)} \int_{Q(x,r)} |f(y)|\, d\mu(y).$$

We have $D(r_x) > \lambda$ and $\lim_{r\to\infty} D(r) = \|f\|_1/(\mu(\mathbb{R}^n)) < \lambda$. Therefore, because $D(r)$ is a continuous function on (r_x, ∞), we get a cube Q_x centered at x such that

$$\lambda = \frac{1}{\mu(Q_x)} \int_{Q_x} |f|\, d\mu. \tag{6}$$

Applying the Besicovitch covering theorem we have an almost disjoint sequence $\{Q_j\}$ of cubes such that $E_\lambda \subset \bigcup_j Q_j$ and such that (6) holds for each Q_j.

Consider functions

$$\varphi_j = \frac{\chi_{Q_j}}{\sum \chi_{Q_j}}, \qquad \frac{1}{B(n)} \le \varphi_j \le 1 \quad \text{on } Q_j,$$

and $\sum \varphi_j \equiv 1$ on $\bigcup_j Q_j$, and define

$$b_j = \big(f\varphi_j - (f\varphi_j)_{Q_j}\big)\chi_{Q_j}.$$

Clearly,

$$\int_{Q_j} |b_j|\,d\mu \le \int_{Q_j} |f|\,d\mu + \int_{Q_j} |f\varphi_j|\,d\mu \le 2\int_{Q_j} |f|\,d\mu = 2\lambda\mu(Q_j)$$

and

$$\int b_j\,d\mu = 0.$$

Finally, take $g = f - \sum_j b_j$ and $b = \sum_j b_j$.

Now, if $x \notin \bigcup Q_j$, then $g(x) = f(x)$ and the differentiation theorem gives $|g(x)| \le \lambda$ for μ-a.e. $x \notin \bigcup Q_j$. When $x \in \bigcup Q_j$,

$$g(x) = \sum_j (f\varphi_j)_{Q_j}\chi_{Q_j} \le \lambda B(n)$$

because $\{Q_j\}$ is almost disjoint and $(f\varphi_j)_{Q_j} \le \lambda$. □

THEOREM 8. *Let μ be a nonnegative Radon measure in $\mathbb{R}^n$ such that $\mu(L) = 0$ for every hyperplane L orthogonal to one of the coordinate axes.*

(a) *Assume $\mu(\mathbb{R}^n) = \infty$. Then, one has*

$$\|f\|_p \le C(p,n)\big\|f^{\#}\big\|_p, \quad 1 < p < \infty,$$

for any $f \in L^1(\mu)$.

(b) *Assume $\mu(\mathbb{R}^n) < \infty$. Then,*

$$\left\| f - \int_{\mathbb{R}^n} f\,d\mu \right\|_p \le C(p,n)\big\|f^{\#}\big\|_p, \quad 1 < p < \infty,$$

for any $f \in L^1(\mu)$.

Proof. We only prove (a) because (b) follows in a similar way. As in the case of the Lebesgue measure, one only has to prove the following good λ-inequality

$$(7) \quad \mu\big(\{x \in \mathbb{R}^n : f(x) > a\lambda,\ f^{\#}(x) < \gamma\lambda\}\big) \le m(a,\gamma,n)\mu\big(\{x \in \mathbb{R}^n : f(x) > \lambda\}\big)$$

for $\lambda > 0$ sufficiently large, where $a > 1$, $\gamma > 0$ are positive constants and where $m(a,\gamma,n) < 1$. Actually, we get

$$m(a,\gamma,n) = C(n)(a-1-2\gamma)^{-1}\gamma.$$

Let $E_\lambda = \{x \in \mathbb{R}^n : f(x) > \lambda\}$. For μ-a.e. $x \in E_\lambda$ one has

$$\lim_{r\to 0} \frac{\mu\big(Q(x,r)\cap E_\lambda\big)}{\mu\big(Q(x,r)\big)} = 1.$$

On the other hand, the Chebyshev inequality gives

$$\overline{\lim_{r\to\infty}} \frac{\mu\big(Q(x,r)\cap E_\lambda\big)}{\mu\big(Q(x,r)\big)} \le \frac{\|f\|_1}{\lambda\mu(\mathbb{R}^n)} = 0.$$

Let $\lambda > 0$. Since μ puts no mass to any hyperplane orthogonal to one of the coordinate axes, one can choose a cube $Q(x)$ centered at x such that

$$\mu\big(Q(x)\cap E_\lambda\big) = \frac{1}{4}\mu\big(Q(x)\big). \tag{8}$$

We first observe that if $Q(x)$ is centered at a point $x \in E_\lambda$ for which $f^\#(x) < \gamma\lambda$, one has

$$\frac{\lambda}{4} < f_{Q(x)} < \lambda(1+2\gamma). \tag{9}$$

Actually,

$$f_Q > \frac{1}{\mu\big(Q(x)\big)}\lambda\mu\big(Q(x)\cap E_\lambda\big) = \frac{\lambda}{4},$$

and since $Q(x)$ is centered at a point $x \in \mathbb{R}^n$, where $f^\#(x) < \gamma\lambda$, we have

$$\mu\big(\{x \in Q(x) : |f(x) - f_{Q(x)}| \ge 2\gamma\lambda\}\big) \le \frac{1}{2}\mu\big(Q(x)\big).$$

Since $\mu(Q(x)\cap E_\lambda^c) = (3/4)\mu(Q(x))$, there exists $y \in Q(x)\cap E_\lambda^c$ such that $|f(y) - f_{Q(x)}| < 2\gamma\lambda$. Since $f(y) \le \lambda$, we deduce

$$f_{Q(x)} < \lambda(1+2\gamma).$$

We apply the Besicovitch covering theorem to the family of cubes $\{Q(x)\}$, $x \in \tilde{E}_\lambda = \{x \in E_\lambda : f^\#(x) < \gamma\lambda\}$. Then we obtain an almost disjoint family of cubes $\{Q_j\}$ with the following properties

$$\mu\big(Q_j\cap E_\lambda\big) = \frac{1}{4}\mu(Q_j), \qquad \tilde{E}_\lambda \subset \bigcup_j Q_j.$$

So the estimate

$$\mu\big(Q_j\cap E_{a\lambda}\big) \le (a-1-2\gamma)^{-1}\gamma\mu(Q_j) \tag{10}$$

finishes the proof.

Using again that Q_j is centered at a point $x \in \mathbb{R}^n$, where $f^{\#}(x) < \gamma\lambda$, we have

$$\int_{Q_j} \left| f - f_{Q_j} \right| d\mu < \gamma\lambda\mu(Q_j).$$

On $E_{a\lambda}$ one has $f > a\lambda$. Thus, (9) gives

$$\int_{Q_j} \left| f - f_{Q_j} \right| d\mu \geq \mu\big(Q_j \cap E_{a\lambda}\big)\lambda(a - 1 - 2\gamma).$$

Then

$$\mu\big(Q_j \cap E_{a\lambda}\big)\lambda(a - 1 - 2\gamma) < \gamma\lambda\mu(Q_j),$$

which gives (10). □

As in the case of the Lebesgue measure, the L^p-boundedness of the sharp function gives a result on interpolation of operators.

Proof of Theorem 3. We first assume $\mu(\mathbb{R}^n) = \infty$. The proof follows closely the arguments in [J, p. 43]. Since functions in $L_c^\infty(\mu)$ with mean zero are dense in $L^p(\mu)$, $1 < p < \infty$, one only has to prove

$$\|Tf\|_p \leq C\|f\|_p$$

for such functions. In fact, we show

$$\left\|(Tf)^{\#}\right\|_p \leq C\|f\|_p, \quad 1 < p < \infty, \tag{11}$$

and apply the previous theorem. Observe that the corresponding hypothesis holds because $f \in L_c^\infty(\mu)$ with mean zero implies $f \in H^1(\mu)$; hence $Tf \in L^1(\mu)$.

By the Marcinkiewicz interpolation theorem (see [S]), the strong inequality (11) follows from the weak estimate

$$\mu\big(\big\{x \in \mathbb{R}^n : (Tf)^{\#}(x) > \lambda\big\}\big) \leq \frac{C\|f\|_p^p}{\lambda^p}, \quad 1 < p < \infty, \tag{12}$$

where $\lambda > 0$.

Let $\{Q_j\}$ be the collection of almost disjoint cubes associated with the Calderón-Zygmund decomposition of $|f|^p$ at the value λ^p. So, we write

$$f = b + g = \sum_j \big(f\varphi_j - (f\varphi_j)_{Q_j}\big)\chi_{Q_j} + g,$$

such that

$$\|g\|_\infty \leq C(n)\lambda.$$

Set

$$b_j = \big(f\varphi_j - (f\varphi_j)_{Q_j}\big)\chi_{Q_j};$$

then $\sup(b_j) \subset Q_j$, b_j has mean zero and

$$\int_{Q_j} |b_j|^p \le C(n)\lambda^p \mu(Q_j).$$

Remember

$$\sum \mu(Q_j) \le \frac{\|f\|_p^p}{\lambda^p}.$$

So, one has

$$\|b\|_{H^1(\mu)} \le C(n)\lambda \sum \mu(Q_j) \le C(n)\lambda^{1-p}\|f\|_p^p.$$

Since T is bounded from $L_c^\infty(\mu)$ to BMO(μ), the function $(Tg)^\#$ is bounded by a multiple of λ. Hence, if c_0 is a sufficiently large constant, we have

$$\big\{x : (Tf)^\#(x) > 2c_0\lambda\big\} \subseteq \big\{x : (Tb)^\#(x) > c_0\lambda\big\}.$$

Now,

$$\mu\big(\big\{x : (Tb)^\#(x) > c_0\lambda\big\}\big) \le \mu\left(\left\{x : M(Tb)(x) > \frac{1}{2}c_0\lambda\right\}\right) \le C\frac{\|Tb\|_1}{\lambda},$$

where M is the Hardy-Littlewood maximal function. The boundedness of the operator T from $H^1(\mu)$ to $L^1(\mu)$ gives

$$\|Tb\|_1 \le C\|b\|_{H^1(\mu)},$$

which gives (12) and finishes the proof.

When $\mu(\mathbb{R}^n) < \infty$, the proof follows the same lines using that constant functions are in $H^1(\mu)$. □

Given a function f on $\mathbb{R}$ denote by $T_\mu f$, the Hilbert transform of f with respect to μ,

$$T_\mu f(x) = \text{p.v.} \int \frac{f(y)}{x-y}\,d\mu(y).$$

We give an example, due to Verdera [V], of a measure μ on $\mathbb{R}$ for which we know that the operator T_μ is bounded on $L^p(\mu)$, $1 < p < \infty$, but it is not bounded from $H^1(\mu)$ to $L^1(\mu)$. Before proceeding to define the measure μ and the function $h \in H^1(\mu)$, we make some computations.

Consider $\epsilon > 0$ sufficiently small, and let $I = [\epsilon, 2\epsilon]$ and $J = [\sqrt{\epsilon}, \sqrt{\epsilon}+\epsilon]$. Write $f = (1/\epsilon)[\chi_J - \chi_I]$, and let ν be the Lebesgue measure restricted on $[-1,0]\cup I \cup J$. A simple calculation gives, when $x \in [-1,0]$,

$$T_\nu f(x) = \frac{1}{\epsilon}\big(\log(2\epsilon - x) - \log(\epsilon - x) - \log\big(\sqrt{\epsilon}+\epsilon - x\big) + \log\big(\sqrt{\epsilon} - x\big)\big) \ge 0,$$

and then we have

$$\int |T_\nu f|\,d\nu \geq \int_{-1}^{0} |T_\nu f(x)|\,dx \geq C|\log\epsilon|.$$

Observe that function f belongs to $H^1(\nu)$ with norm bounded by 2, but with respect to the Lebesgue measure f is an atom with norm of order $\epsilon^{-1/2}$. Now, one repeats this construction at different scales.

Let (ϵ_k) be a sequence of positive numbers tending to zero and verifying $\sqrt{\epsilon_{k+1}}+\epsilon_{k+1}<\epsilon_k$ (and so $\sum_{k\geq 1} k^{-2}|\log\epsilon_k| = \infty$). Define intervals

$$I_k = [\epsilon_k, 2\epsilon_k] \qquad \text{and} \qquad J_k = \left[\sqrt{\epsilon_k}, \sqrt{\epsilon_k}+\epsilon_k\right].$$

Let μ be the Lebesgue measure restricted on $[-1,0]\bigcup_{k=2}^{\infty}(I_k\cup J_k)$. Clearly, μ does not satisfy any doubling condition, and easily, T_μ is bounded in $L^2(\mu)$ because $d\mu = g\,dm$, where $g^2 = g$. Observe that the first condition on the sequence (ϵ_k) means that J_{k+1} is at the left-hand side of I_k. So, the function $a_k = \epsilon_k^{-1}(\chi_{I_k} - \chi_{J_k})$ is a μ-atom. Define

$$h = \sum_{k=2}^{\infty} \frac{1}{k^2} a_k.$$

Since $\sum k^{-2} < +\infty$ and a_k are atoms, one has $h \in H^1(\mu)$. On the other hand,

$$\int |T_\mu h|\,d\mu \geq \int_{-1}^{0} |T_\mu h(x)|\,dx = \sum_{k=2}^{\infty} \frac{1}{k^2} \int_{-1}^{0} T_\mu a_k(x)\,dx \geq C \sum_{k=2}^{\infty} \frac{1}{k^2} |\log\epsilon_k| = +\infty.$$

That is, $T_\mu h \notin L^1(\mu)$.

4. BMO with balls and cubes. Theorems 4, 5, and 6 are proved with a similar construction. To simplify slightly, we do not construct μ as an absolutely continuous measure but as a sum of weighted length measures on some circles. Since the oscillation of f on two neighboring circles is at most 1, it is clear from the proof that by replacing μ with a sum of weighted Lebesgue measures on very narrow annuli, we get the same conclusions.

We choose nonincreasing sequences (ε_i) and (λ_i), $i = 1, 2, \dots$, of positive numbers such that for all i,

$$\varepsilon_1 = \frac{1}{2}, \qquad \varepsilon_{i+1} \leq \frac{\varepsilon_i}{10}, \tag{13}$$

$$\lambda_1 = \frac{1}{2}, \qquad 0 < \lambda_i \leq 2^{-i}, \qquad \lambda_{i+1} \leq \lambda_i, \tag{14}$$

$$\sqrt{\varepsilon_{i+1}}\, 2^{-i-1} \leq \sqrt{\varepsilon_i}\, \lambda_i. \tag{15}$$

Set

$$S_i = \left\{x \in \mathbb{R}^2 : |x| = 1 - \varepsilon_i\right\},$$

$$T_i = \left\{x \in \mathbb{R}^2 : |x| = 1 + \varepsilon_i\right\},$$

$$\mu_S = \sum_i \lambda_i \mathcal{H}^1 \mid S_i,$$

$$\mu_T = \sum_i 2^{-i} \mathcal{H}^1 \mid T_i,$$

$$\mu = \mu_S + \mu_T,$$

$$f = \sum_i i \chi_{S_i \cup T_i}.$$

Clearly, f belongs to $L^1(\mu)$ with

$$\int f \, d\mu \leq 8\pi \sum_{i=1}^{\infty} i 2^{-i} = c_0.$$

We now fix a disc B with center x and radius r. If $B \cap S_i \neq \emptyset$ for some i, we let i_0 be the smallest such i. Similarly, j_0 is the smallest j such that $B \cap T_j \neq \emptyset$ if such a j exists.

LEMMA 4. *If $B \cap T_j \neq \emptyset$ for some j, then*

$$\int_B |f - j_0| \, d\mu_T \leq c_0 \mu_T(B).$$

Proof. As $B \cap T_{j_0} \neq \emptyset$, we see by simple geometry that

$$\mathcal{H}^1(B \cap T_j) \leq 5\mathcal{H}^1(B \cap T_{j_0+1}) \quad \text{for } j > j_0$$

(since T_j and T_{j_0+1} are much closer to each other than to T_{j_0}). Hence,

$$\begin{aligned} \int_B |f - j_0| \, d\mu_T &= \sum_{j > j_0} (j - j_0) 2^{-j} \mathcal{H}^1(B \cap T_j) \\ &\leq 5\mathcal{H}^1(B \cap T_{j_0+1}) \sum_{j > j_0} (j - j_0) 2^{-j} \\ &\leq c_0 2^{-j_0 - 1} \mathcal{H}^1(B \cap T_{j_0+1}) = c_0 \mu(B \cap T_{j_0+1}) \leq c_0 \mu_T(B). \qquad \square \end{aligned}$$

LEMMA 5. *We have $f \in \mathrm{BMO}_b(\mu)$. More precisely, (2) holds with an absolute constant c.*

Proof. Let B be as above. We consider four cases.

Case 1: $|x| \le 1/2$. If $B \cap S_2 = \emptyset$, f is constant on $B \cap \operatorname{spt}\mu$. Otherwise, $r > 1/3$, and using trivial geometry and the assumptions $\varepsilon_1 = \lambda_1 = 1/2$ and $\varepsilon_2 \le \varepsilon_1/10$, we have $\mathcal{H}^1(B \cap S_1) \ge 2/5$ and

$$\mu(B) \ge \lambda_1 \mathcal{H}^1(B \cap S_1) > \frac{1}{5}.$$

Thus,

$$\frac{1}{\mu(B)} \int_B f\, d\mu \le 5 \int f\, d\mu \le 5c_0,$$

and (2) holds with $a = 0$, $c = 5c_0$.

From now on we assume that $|x| > 1/2$ and, by Lemma 4, that $B \cap S_i \ne \emptyset$ for some i.

Case 2: either $B \subset \{y : |y| \le 1\}$ or $j_0 > i_0+2$. We may assume that $B \cap S_{i_0+2} \ne \emptyset$, since otherwise $f = i_0$ or i_0+1 on $B \cap \operatorname{spt}\mu$. As B meets S_{i_0} and S_{i_0+2}, and it does not meet T_{i_0+2}, one sees by simple geometry (see Figure 3) that

$$\mu(B) \ge \lambda_{i_0+1} \mathcal{H}^1(B \cap S_{i_0+1}) \ge \lambda_{i_0+1} \sqrt{r\varepsilon_{i_0+1}},$$

$$\mathcal{H}^1(B \cap S_i) \le 2\sqrt{r\varepsilon_{i_0+2}} \quad \text{for } i > i_0+1,$$

$$\mathcal{H}^1(B \cap T_i) \le 2\sqrt{r\varepsilon_{i_0+2}} \quad \text{for } i > i_0+1.$$

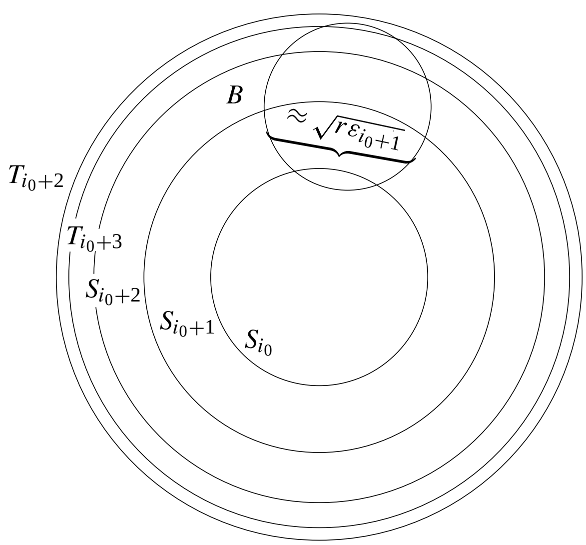

FIGURE 3. Case 2

Thus, also using (15),

$$\begin{aligned}\int_B |f-i_0|\,d\mu &\le \sum_{i>i_0}(i-i_0)\lambda_i \mathcal{H}^1(B\cap S_i)+\sum_{i>i_0+2}(i-i_0)2^{-i}\mathcal{H}^1(B\cap T_i)\\ &\le 3\mu(B)+4\sqrt{r\varepsilon_{i_0+2}}\sum_{i>i_0+2}(i-i_0)2^{-i}\\ &\le 3\mu(B)+c_0\sqrt{r\varepsilon_{i_0+2}}\,2^{-i_0-2}\\ &\le 3\mu(B)+c_0\sqrt{r\varepsilon_{i_0+1}}\,\lambda_{i_0+1}\le(3+c_0)\mu(B).\end{aligned}$$

Hence, (2) holds in this case with $c=3+c_0$. In the last two cases, we also assume that $B\cap T_j\neq\emptyset$ for some j.

Case 3: $j_0=i_0$, i_0+1, or i_0+2. Since $\varepsilon_{j_0+1}<\varepsilon_{j_0}/10$ and B meets both S_{j_0} and T_{j_0}, one easily sees that for $i>j_0$,

$$\mathcal{H}^1(B\cap S_i)\le 2\mathcal{H}^1(B\cap T_{j_0+1}).$$

Thus,

$$\begin{aligned}\int_B |f-j_0|\,d\mu_S &\le 2\mu(B\cap S_{j_0-2})+\mu(B\cap S_{j_0-1})+\sum_{i>j_0}(i-j_0)\lambda_i\mathcal{H}^1(B\cap S_i)\\ &\le 3\mu(B)+2\mathcal{H}^1(B\cap T_{j_0+1})\sum_{i>j_0}(i-j_0)2^{-i}\\ &\le 3\mu(B)+c_0 2^{-j_0-1}\mathcal{H}^1(B\cap T_{j_0+1})\\ &\le(3+c_0)\mu(B).\end{aligned}$$

Combining this with Lemma 4, we have (2) with $c=3+c_0$ in Case 3.

Our final case is the following.

Case 4: $j_0<i_0$. Then for $i>j_0$,

$$\mathcal{H}^1(B\cap S_i)\le\mathcal{H}^1(B\cap T_i)\qquad\text{and}\qquad\mathcal{H}^1(B\cap T_i)\le 5\mathcal{H}^1(B\cap T_{j_0+1}).$$

Thus,

$$\begin{aligned}\int_B |f-j_0|\,d\mu &\le \sum_{i>j_0}(i-j_0)(\lambda_i+2^{-i})\mathcal{H}^1(B\cap T_i)\\ &\le 2\sum_{i>j_0}(i-j_0)2^{-i}\mathcal{H}^1(B\cap T_i)\le 2c_0\mu(B).\end{aligned}$$

Hence, the proof of Lemma 5 is complete. □

To prove Theorem 4 we choose sequences (i_k) and (j_k) of positive integers and the sequences (ε_i) and (λ_i) in such a way that $i_k + k < j_k < i_{k+1}$,

$$\varepsilon_{i_k}^2 < \varepsilon_{i_k+k}, \tag{16}$$

(17)

$$\lambda_1 = \frac{1}{2}, \qquad \lambda_i = \lambda_{i_k} \quad \text{for } i = i_k, \ldots, i_k + k, \qquad \lambda_{i+1} \leq \frac{1}{2}\lambda_i \quad \text{for } i = i_k + k, \ldots, j_k,$$

for all i and k. We leave to the reader as an exercise to check that this choice, keeping also (13)–(15), is possible. Moreover, once the other numbers at the kth step are chosen, we can take j_k as large as we wish. The choice of j_k is determined in the proof of the following lemma.

LEMMA 6. *Under the conditions (13)–(17),* $f \notin \mathrm{BMO}(\mu)$ *for any choice of the coordinate axes.*

Proof. Let Q_k be the square with center on the x_1-axes, sidelength $2\varepsilon_{i_k}$, and the right-hand vertices on S^1 (see Figure 4).

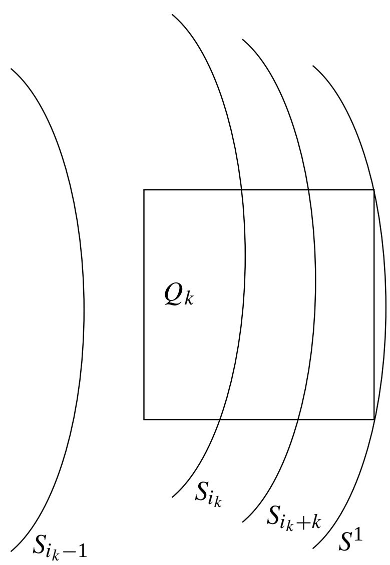

FIGURE 4. Proof of Lemma 6

By (16) for $i = i_k, \ldots, i_k + k$, $Q_k \cap S_i$ does not meet the vertical sides of Q_k, whence it is an arc with $\varepsilon_{i_k} \leq \mathcal{H}^1(Q_k \cap S_i) \leq 3\varepsilon_{i_k}$. Thus, by (17),

$$\mu(Q_k \cap S_i) = \lambda_{i_k} \mathcal{H}^1(Q_k \cap S_i) \geq \lambda_{i_k} \varepsilon_{i_k} \quad \text{for } i = i_k, \ldots, i_k + k.$$

We choose j_k such that

$$\mathcal{H}^1\big(Q_k \cap S_{i+1}\big) \leq \frac{1}{2} \mathcal{H}^1(Q_k \cap S_i) \quad \text{for } i \geq j_k.$$

By (13) and elementary geometry, this is possible. Then by (17),

$$\mu\big(Q_k \cap S_{i+1}\big) \leq \frac{1}{2} \mu(Q_k \cap S_i) \quad \text{for } i \geq i_k + k.$$

Thus,

$$\begin{aligned} \mu(Q_k) &= \sum_{i=i_k}^{i_k+k} \mu(Q_k \cap S_i) + \sum_{i > i_k + k} \mu(Q_k \cap S_i) \\ &\leq 2 \sum_{i=i_k}^{i_k+k} \mu(Q_k \cap S_i) = 2\lambda_{i_k} \sum_{i=i_k}^{i_k+k} \mathcal{H}^1(Q_k \cap S_i) \\ &\leq 12 \lambda_{i_k} k \varepsilon_{i_k}. \end{aligned}$$

For any number a, there are at least $k/3$ values $i \in \{i_k, \ldots, i_k + k\}$ such that $|f(x) - a| \geq k/3$ for $x \in S_i$. Thus

$$\frac{1}{\mu(Q_k)} \int_{Q_k} |f - a| \, d\mu \geq \frac{1}{12 \lambda_{i_k} k \varepsilon_{i_k}} \frac{k}{3} \frac{k}{3} \lambda_{i_k} \varepsilon_{i_k} = \frac{1}{108} k.$$

Hence, $f \notin \mathrm{BMO}(\mu)$ and Lemma 6 is proved; consequently Theorem 4 is also proved. □

Proof of Theorem 6. The proof follows much the same lines as the one given for Theorem 4, but in the present case both the measure and the function are not constant on circles. Let $m \geq 9$ be an integer. Let

$$\lambda_1 = 4^{-m}, \qquad \lambda_2 = \cdots = \lambda_m = m^{-2} 4^{-m}. \tag{18}$$

Let $\lambda > 0$, and choose $\varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_m$ such that (13) and (15) are satisfied. We choose λ much smaller than λ_2, and then we choose the sequence (ε_i), very quickly decreasing. Set S_i and T_i as before. Divide the unit circle S^1 into disjoint consecutive arcs $I_1, \ldots, I_{8^m}$ of length $2\pi 8^{-m}$. For each j divide I_j into disjoint consecutive arcs $I_{j,0}, \ldots, I_{j,2m}$ of length $\mathcal{H}^1(I_{j,i}) = 2\pi 8^{-m}/(2m+1) = \ell_m$. Set, with $tA = \{tx : x \in A\}$,

$$I_{j,i}^k = (1-\varepsilon_k) I_{j,i} \subset S_k, \qquad J_{j,i}^k = (1+\varepsilon_k) J_{j,i} \subset T_k,$$

$$S_{k,j} = \bigcup_{i=0}^{2m} I_{j,i}^k, \qquad T_{k,j} = \bigcup_{i=0}^{2m} J_{j,i}^k,$$

$$\mu_S = \sum_{k=1}^{m} \sum_{j=1}^{8^m} \left(\lambda_k \mathcal{H}^1 \mid I_{j,m}^k + \lambda \sum_{i \neq m} \mathcal{H}^1 \mid I_{j,i}^k \right),$$

$$\mu_T = \sum_{k=1}^{m} \sum_{j=1}^{8^m} \left(\sum_{i=0}^{m} 2^{-i-k} \mathcal{H}^1 \mid J_{j,i}^k + \sum_{i=m+1}^{2m} 2^{i-2m-k} \mathcal{H}^1 \mid J_{j,i}^k \right),$$

$$\mu = \mu_S + \mu_T, \qquad \text{and}$$

$$f = \sum_{k=1}^{m} \sum_{j=1}^{8^m} \left(\sum_{i=m-k}^{m} (i-m+k) \chi_{I_{j,i}^k \cup J_{j,i}^k} + \sum_{i=m+1}^{m+k} (m+k-i) \chi_{I_{j,i}^k \cup J_{j,i}^k} \right).$$

Therefore, easy computations give

$$\mu(T_{k,j}) \simeq 2^{-k} \ell_m \qquad \text{and} \qquad \|\mu_T\| \simeq 8^m \ell_m \simeq m^{-1} \simeq \mu(T_1),$$

$$\|\mu_S\| \simeq \frac{4^{-m}}{m} \big(1 + \lambda 4^m m^2\big) \simeq \frac{4^{-m}}{m} \simeq \mu(S_1), \quad \text{whenever } \lambda < 4^{-m} m^{-2},$$

$$\int_{T_{k,j}} f \, d\mu \simeq \ell_m 2^{-m} \qquad \text{and} \qquad \int f \, d\mu_T \simeq 2^{-m},$$

$$\int f \, d\mu_S \simeq \frac{4^{-m}}{m} \big(1 + \lambda 4^m m^3\big) \simeq \frac{4^{-m}}{m}, \quad \text{whenever } \lambda < 4^{-m} m^{-3}.$$

Then, clearly $f \in L^1(\mu)$ with $\int f \, d\mu \le c_0 \|\mu\|$, where c_0 is as before.

Note that $f = 0$ on the extreme arcs of $T_{k,j}$ and that it oscillates linearly to k in the middle. Since the measure decays exponentially in the middle, f is again nicely in $\mathrm{BMO}_b(\mu_T)$, as the next lemma shows. Let B be a disc with center x.

LEMMA 7. *There is $a \in \mathbb{R}$ such that*

$$\int_B |f-a| \, d\mu_T \le C \mu_T(B).$$

Moreover, if for some k, $B \cap T_{k,j} \neq \emptyset$ for at least two indices j, then we can take $a = 0$.

Proof. Suppose that $B \cap T_{k,j} \neq \emptyset$, and suppose that the indices i for which $B \cap J_{j,i}^k \neq \emptyset$ form a sequence $i_1, i_1+1, \dots, i_2$ such that $i_1 = 0$ or $i_1 = 1$ or $i_2 = 2m$ or $i_2 = 2m-1$ or form two sequences $i_1, i_1+1, \dots, i_2$ and $i_3, i_3+1, \dots, i_4$ such that $i_1 = 0$ or $i_1 = 1$ and $i_4 = 2m$ or $i_4 = 2m-1$. We say then that B intersects $T_{k,j}$ noncentrally. Otherwise, we say that B intersects $T_{k,j}$ centrally. We also say that B intersects T_k centrally or noncentrally if it intersects some $T_{k,j}$ centrally or, respectively, noncentrally.

We claim that if B intersects $T_{k,j}$ noncentrally, then

$$\int_{B\cap T_{k,j}} f\,d\mu \le c_0\mu\big(B\cap T_{k,j}\big). \tag{19}$$

Suppose, for example, that there is only one sequence $i_1,\dots,i_2$ as above with $i_1 = 0$ or $i_1 = 1$. The other cases are similar. If $i_2 = 0$, 1, or 2, then $f = 0$, 1, or 2 on $B\cap T_{k,j}$, and (19) is clear.

Otherwise, $J^k_{j,2} \subset B$, whence

$$\mu\big(B\cap T_{k,j}\big) \ge 2^{-2-k}\mathcal{H}^1\big(J^k_{j,2}\big) \ge 2^{-2-k}\ell_m$$

and

$$\int_{B\cap T_{k,j}} f\,d\mu \le \int_{T_{k,j}} f\,d\mu \simeq \ell_m 2^{-m}.$$

Hence, (19) follows.

Assume now that for some k, $B\cap T_{k,j} \ne \emptyset$ for at least two values of j. We claim that

$$\int_B f\,d\mu_T \le 16c_0\mu_T(B). \tag{20}$$

Let $k_1, k_1+1,\dots,k_2$ be those k for which $B\cap T_k \ne \emptyset$.

We examine three cases separately. First, suppose B intersects T_{k_1} centrally. Then $B\cap T_{k_1} = B\cap T_{k_1,j}$ for some j and

$$\int_{B\cap T_{k_1}} f\,d\mu \le \int_{T_{k_1,j}} f\,d\mu \simeq \ell_m 2^{-m}. \tag{21}$$

We see by simple geometry that if the ε_i's decrease sufficiently quickly, our assumption (that for some k, $B\cap T_{k,j} \ne \emptyset$ for at least two j) implies that

$$\mu\big(B\cap T_{k_1+1}\big) \ge 2^{-2-k_1}\ell_m. \tag{22}$$

Moreover, $k_2 = m$ and B intersects the remaining T_k noncentrally. Combining (19), (21), and (22), we get (20) for this first case. On the other hand, if B intersects both T_{k_1} and T_{k_2} noncentrally, then B intersects every T_k noncentrally and again (20) follows from (19). Finally, when B intersects T_{k_1} noncentrally and T_{k_2} centrally, our assumption implies $\mu(B\cap T_{k_1}) \ge 2^{-k_1}\ell_m$. Then B intersects T_k centrally for $k = k_3,\dots,k_2$, and as in (21),

$$\sum_{k_3}^{k_2}\int_{B\cap T_k} f\,d\mu \le (m-k_1)2^{-m}\ell_m \le 2^{-k_1}\ell_m \le \mu_T(B).$$

Combining this with (19), we also obtain (20).

To finish the proof of Lemma 7, suppose that for every k there is at most one j such that $B \cap T_{k,j} \neq \emptyset$. Then the values of k for which $B \cap T_k \neq \emptyset$ form a sequence $k_1, k_1+1, \ldots, k_2$, and the index j for which $B \cap T_{k,j} \neq \emptyset$ is the same for all $k_1 \leq k \leq k_2$. Moreover, the indices i for which $B \cap J_{j,i}^{k_1+1} \neq \emptyset$ (if there are any) form a sequence $i_1, i_1+1, \ldots, i_2$. If the sequence (ε_i) decreases sufficiently quickly, $B \cap J_{j,i}^{k} = \emptyset$ for $k > k_1$, $i_1 < i_1 - 1$, and $i > i_2 + 1$. Thus $\mu_T(B) \leq 4\mu(B \cap (T_{k_1,j} \cup T_{k_1+1,j}))$. Then, by similar easy estimates as before,

$$\int_B |f-a| \, d\mu_T \leq C \int_{B \cap (T_{k_1,j} \cup T_{k_1+1,j})} |f-a| \, d\mu \leq C^2 \mu_T(B),$$

where $a = \min\{f(x) : x \in B \cap (T_{k_1,j} \cup T_{k_1+1,j})\}$. This completes the proof of Lemma 7. □

We want to show again that (2) holds for some a with an absolute constant c; that is, $f \in \mathrm{BMO}_b(\mu)$.

Suppose first that $|x| \leq 1/2$. We may assume that $B \cap S_2 \neq \emptyset$. Then B contains at least 8^{m-1} arcs $I_{j,m}^1$ and so

$$\mu(B) \geq 8^{m-1} \lambda_1 \ell_m / 2 = 2^m \ell_m / 16 \simeq 4^{-m} m^{-1}.$$

We have

$$\int_B f \, d\mu_S \leq \int f \, d\mu_S \simeq 4^{-m} m^{-1} \lesssim \mu(B)$$

provided we choose $\lambda \leq \ell_m / (m^2 2\pi)$. If $B \cap T_{k,j} \neq \emptyset$ for at most one j for every k, then we get

$$\int_B f \, d\mu_T \leq \sum_{k=1}^{m} \int_{T_{k,j}} f \, d\mu \simeq m 2^{-m} \ell_m \leq 16 m 4^{-m} \mu(B) \leq \mu(B)$$

since $m \geq 9$.

If $B \cap T_{k,j} \neq \emptyset$ for at least two indices j, for some k, we have by Lemma 7,

$$\int_B f \, d\mu_T \leq C\mu(B).$$

Combining these inequalities we obtain

$$\int_B f \, d\mu \leq C\mu(B)$$

in case $|x| \leq 1/2$.

Assume then that $|x| > 1/2$. As previously, we study different cases. By Lemma 7 we may assume that $B \cap S_k \neq \emptyset$ for some k. The case where this happens only for $k = 1$

is trivial, so suppose that $B \cap S_k$ for some $k > 1$. If $B \cap S_2 = B \cap T_4 = \emptyset$, the diameter of B is at most $2\varepsilon_2$. Choosing $2\varepsilon_2 < \ell_m$, there are two pairs (i_1, j_1) and (i_2, j_2) such that for every k, $B \cap (I^k_{j,i} \cup J^k_{j,i}) = \emptyset$ unless $(i, j) = (i_1, j_1)$ or $(i, j) = (i_2, j_2)$. Then we can use almost identical arguments as in the proof of Theorem 4, because in this scale both the measure and the function are almost constant on circles.

Suppose then that $B \cap T_4 \neq \emptyset$. Then, choosing ε_4 sufficiently small for all $k \geq 5$,

$$\mu\big(B \cap (S_k \cup T_k)\big) \leq 2^{6-k}\mu(B \cap T_5),$$

whence, as $f \leq k$ on $S_k \cup T_k$,

$$\int_B f\,d\mu \leq 4\mu(B) + \sum_{k=5}^{m} k2^{6-k}\mu(B \cap T_5) \leq (4+8)\mu(B) = 12\mu(B).$$

Finally, we are left with the case where $B \cap S_2 \neq \emptyset$ and $B \cap T_4 = \emptyset$. If B intersects only S_2 and S_3, it is trivial. Therefore, assume $B \cap S_4 \neq \emptyset$. Let r be the radius of B. We then have

$$\mu(B \cap S_3) \geq \lambda\mathcal{H}^1(B \cap S_3) \geq \lambda\sqrt{r\varepsilon_3}$$

and for $k \geq 4$,

$$\mu\big(B \cap (S_k \cup T_k)\big) \leq 2^{-k}\mathcal{H}^1\big(B \cap (S_k \cup T_k)\big) \leq 2^{2-k}\mathcal{H}^1(B \cap S_4) \leq 2^{3-k}\sqrt{r2\varepsilon_4}.$$

Choosing $\varepsilon_4 \leq \lambda^2\varepsilon_3$, we conclude

$$\begin{aligned}\int f\,d\mu &\leq 3\mu(B) + \int_{B\setminus S_2\cup S_3} f\,d\mu\\ &\leq 3\mu(B) + \sum_{k=4}^{m} k2^{4-k}\sqrt{r\varepsilon_4}\\ &\leq 3\mu(B) + 6\sqrt{r\varepsilon_4}\\ &\leq 3\mu(B) + 6\lambda\sqrt{r\varepsilon_3} \leq 9\mu(B).\end{aligned}$$

This completes the proof that the function f belongs to $\mathrm{BMO}_b(\mu)$ with the norm independent of m. We point out that if B is a disc not contained in $\{x : |x| < 3\}$, then

$$\int_B f\,d\mu \leq C\mu(B). \tag{23}$$

Clearly, when $\mu(B \cap S) > 0$, we have $\mu(B) \geq \mu_T(B) \geq \|\mu\|/100$, and thus (23) holds. If $\mu(B \cap S) = 0$, then either $B \cap T_2 = \emptyset$ or $B \cap T_{1,j} \neq \emptyset$ for at least two indices j. In both cases we get (23) from Lemma 7.

To finish the proof of Theorem 6, we choose discs $B_1, B_2, \ldots$ such that the discs $3B_k$ are disjoint. Let $\{M_k\}$ be an increasing sequence of integer numbers (for instance,

$M_k = 9+k$). Taking $m = M_k$, we use the above construction in each B_k appropriately scaled. Then we get in $3B_k$ a measure μ_k and $f_k \in \mathrm{BMO}(\mu_k)$ with uniformly bounded BMO_b-norms. Set

$$\mu = \sum_k \mu_k, \qquad f = \sum_k f_k.$$

Let B be any disc. If $B \subset 3B_k$ for some k, then $\mu_{|B} = \mu_{k_{|B}}$, and (2) holds. Otherwise, B is not contained in $3B_k$ for any k. Then by (23),

$$\int_B f\, d\mu = \sum_k \int_B f_k\, d\mu_k \le C \sum_k \mu_k(B) = C\mu(B).$$

So $f \in \mathrm{BMO}_b(\mu)$. Denote by r_k the radius of B_k, and note that $\mu_k(B_k) \simeq r_k 4^{-M_k} M_k^{-1}$. Then for any $a \in \mathbb{R}$ and all k, we have

$$\mu\left(\left\{x \in B_k : |f(x)-a| > \frac{M_k}{3}\right\}\right) \ge \frac{M_k}{3} \cdot \frac{4^{-M_k}}{M_k^2} \cdot \ell_{M_k} 8^{M_k} \ge r_k \frac{4^{-M_k}}{3M_k^2} \simeq \frac{1}{M_k}\mu(B_k).$$

This implies trivially (by the Chebyshev inequality) that $f \notin \mathrm{BMO}_b^p(\mu)$ for all $p > 1$. Thus, finally, the proof of Theorem 6 is complete. □

Proof of Theorem 5. Let $m > 3$ be an integer; we have $\varepsilon_1, \dots, \varepsilon_m$ such that $\varepsilon_1 < 1/2$ and $0 < \varepsilon_{i+1} < \varepsilon_i/10$ for $1 \le i < m$. Let $S_i = \{x : |x| = 1-\varepsilon_i\}$ and $T_i = \{x : |x| = 1+\varepsilon_i\}$ as before. We choose the sequence (ε_i) so quickly decreasing that the following hold:

$$\mathcal{H}^1\big(Q \cap S(r)\big) \le 2^{-m}\mathcal{H}^1\big(Q \cap S_{i+1}\big) \tag{24}$$

for $1-\varepsilon_{i+2} \le r \le 1+\varepsilon_{i+2}$, $1 \le i \le m-2$, and for any square Q such that $Q \cap S_i \ne \emptyset$ and $Q \cap T_{i+2} = \emptyset$;

$$\mathcal{H}^1\big(Q \cap S(r)\big) \le 2\mathcal{H}^1\big(Q \cap T_{i+1}\big) \tag{25}$$

for $1-\varepsilon_{i+1} \le r \le 1$, $1 \le i \le m-1$, and for any square Q such that $Q \cap T_i \ne \emptyset$. Here $S(r)$ denotes $\{x : |x| = r\}$.

We define μ and f as in the proof of Theorem 4 with

$$\lambda_1 = \cdots = \lambda_m = \lambda = 2^{-m}.$$

Then $f = i$ on $S_i \cup T_i$.

First, we observe that the $\mathrm{BMO}_b(\mu)$-norm of f is at least $m/9$, since for any $a \in \mathbb{R}$ and $D = \{x : |x| \le 1\}$,

$$\frac{3}{m}\int_D |f-a|\, d\mu \ge \mu\big(\{x \in D : |f(x)-a| > m/3\}\big) \ge \mu(D)/3.$$

Second, we claim that (1) holds for any square Q with an absolute constant C. For this we can follow the argument in the proof of Theorem 4. Lemma 4 and its proof hold with B replaced by Q as they stand; i_0 and j_0 are defined in the same way with Q in place of B. We need not worry about where the center of Q lies, so our first case is the analogue of Case 2: $Q \subset D$ or $j_0 > i_0 + 2$. In this case, (24) yields

$$\mathcal{H}^1\big(Q \cap (S_i \cup T_i)\big) \le 2\lambda \mathcal{H}^1(B \cap S_{i_0+1}) \le 2\mu(B)$$

for $i > i_0 + 1$, and the inequalities in Case 2 can be repeated. In case $j_0 = i_0, i_0 + 1$, or $i_0 + 2$ (see Case 3), we have by (25),

$$\mathcal{H}^1(Q \cap S_i) \le 2\mathcal{H}^1(Q \cap T_{j_0+1})$$

for $i > j_0$, and the same argument works again. Finally, if $j_0 < i_0$, the proof runs as it is with B replaced by Q.

To complete the proof of Theorem 5, we use the same method as at the end of the proof of Theorem 6, defining $f = \sum_m f_m$ and $\mu = \sum_m \mu_m$. Then $f \in \mathrm{BMO}(\mu)$ (for all choices of the coordinate axes), but $f \notin \mathrm{BMO}_b(\mu)$. □

Another variant of functions of bounded mean oscillation is the following one. Suppose that in the definition of the space $\mathrm{BMO}(\mu)$ we only consider cubes (with sides parallel to the axes) centered at the support of the measure μ. We denote this new space of functions by $\mathrm{BMO}_c(\mu)$. Obviously, $\mathrm{BMO}(\mu)$ is contained in $\mathrm{BMO}_c(\mu)$, and as the next example shows, this inclusion can be strict, even for the restriction of the Lebesgue measure to a cube.

Example. The idea to construct our example is simple, but again the explanation becomes a little tedious. Take the square $Q_0 = [0,1] \times [-1/2, 1/2]$, and let μ be the planar Lebesgue measure restricted in Q_0. Now, we consider a collection of squares which are dyadic with respect to Q_0. For each positive integer k and for each $j = 1, 2, \dots, 2^{k-1}$, we define squares

$$Q^+_{k,j} = \left[0, \frac{1}{2^k}\right] \times \left[\frac{j-1}{2^k}, \frac{j}{2^k}\right],$$

$$Q^-_{k,j} = \left[0, \frac{1}{2^k}\right] \times \left[\frac{-j}{2^k}, \frac{1-j}{2^k}\right].$$

Let φ be a Lipschitz function satisfying

(i) $0 \le \varphi \le 1$;

(ii)

$$\varphi(z) = 1 \quad \text{if } z \in \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right],$$

$$\varphi(z) = 0 \quad \text{if } z \notin \left[-\frac{3}{4}, \frac{3}{4}\right] \times \left[-\frac{3}{4}, \frac{3}{4}\right];$$

(iii) $\|\nabla\varphi\| \leq 4$.

Then define

$$\varphi^{+}_{k,j}(z) = \varphi\bigl(2^k\bigl(z - c^{+}_{k,j}\bigr)\bigr) \qquad \text{and} \qquad \varphi^{-}_{k,j}(z) = \varphi\bigl(2^k\bigl(z - c^{-}_{k,j}\bigr)\bigr),$$

where $c^{+}_{k,j}$ and $c^{-}_{k,j}$ are the centers of $Q^{+}_{k,j}$ and $Q^{-}_{k,j}$. Thus, $\varphi^{+}_{k,j}$ is equal to 1 on $Q^{+}_{k,j}$, has support contained in $(3/2)Q^{+}_{k,j}$, and $\|\nabla\varphi^{+}_{k,j}\| \leq 8\cdot 2^k$. The function $\varphi^{-}_{k,j}$ satisfies the analogous properties.

Write

$$b = \sum_{k=1}^{\infty}\sum_{j=1}^{2^{k-1}}\bigl(\varphi^{+}_{k,j} - \varphi^{-}_{k,j}\bigr).$$

When Q is any square contained in Q_0, a standard computation (or an application of [GJ, Lemma 2.1]) gives

$$\frac{1}{\mu(Q)}\int_Q |b - b_Q|\,d\mu \leq C.$$

If Q is any square with center lying in Q_0, then there is a square $P \subset Q_0$ so that $Q \cap Q_0 \subset P$ and $\mu(Q) \geq (1/4)\mu(P)$. So,

$$\int_Q |b - b_P|\,d\mu \leq \int_P |b - b_P|\,d\mu \leq C\mu(P) \leq C\mu(Q).$$

Consequently, $b \in \mathrm{BMO}_c(\mu)$. On the other hand, let Q_k be a square such that $Q_k \cap Q_0 = [0, 1/2^k] \times [-1/2, 1/2]$. Then $\mu(Q_k) = 2^{-k}$. By symmetry of b, $\int_{Q_k} b = 0$. Observe that

$$\int_{Q_k} |b - b_{Q_k}|\,d\mu = \int_{Q_k} |b| \geq \sum_{j=k}^{\infty} \frac{j}{2^{j+1}} \geq \frac{k}{2}2^{-k} = \frac{k}{2}\mu(Q_k).$$

Therefore, $b \notin \mathrm{BMO}(\mu)$.

Now, we describe an example that shows that the John-Nirenberg inequality is false for $\mathrm{BMO}_c(\mu)$, even for absolutely continuous measures μ and cubes centered at points of the support of the measure. For simplicity, our measure again contains pieces on line segments, but obviously it can be fattened without destroying the desired properties.

Let $m > 1$ be an integer. Set

$$I_j = \bigl\{(x, y) : (j-1)/m < x \leq j/m,\ y = 1\bigr\}, \qquad j = 1, \dots, m,$$

$$J_j = \bigl\{(x, y) : (j-1)/m < x \leq j/m,\ y = 1 + 1/m\bigr\}, \quad j = 1, \dots, m,$$

$$K = \bigl\{(x, y) : 0 \leq x \leq 1,\ y = 0\bigr\},$$

$$I = \bigcup_{j=1}^{m} I_j, \qquad J = \bigcup_{j=1}^{m} J_j.$$

Then consider the measure

$$\mu = 2^{-m}\mathcal{H}^1 \mid K + m^{-1}2^{-m}\mathcal{H}^1 \mid I + \sum_{j=1}^{m} 2^{-j}\mathcal{H}^1 \mid I_j$$

and the function

$$f = \sum_{j=1}^{m} j\,\chi_{I_j \cup J_j}$$

(see Figure 5).

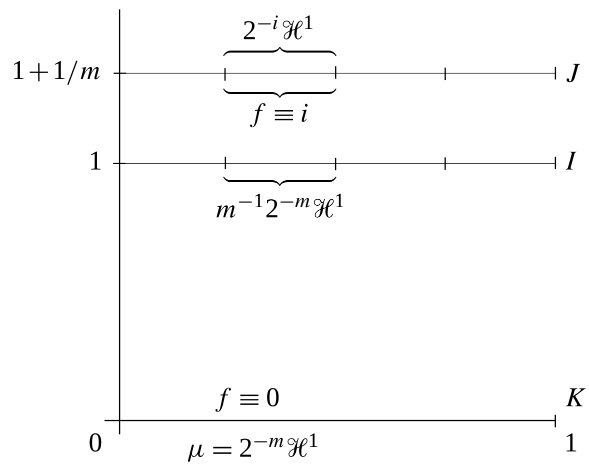

FIGURE 5. The function f and the measure μ, when $m = 4$

It is easy to check that for every square Q with sides parallel to the coordinate axes and with center in $I \cup J \cup K$, there is $a = a(Q) \in \mathbb{R}$ such that

$$\int_Q |f - a|\, d\mu \le C\mu(Q). \tag{26}$$

Actually, if the center of Q is in K, one may take $a = 0$ because

$$\int_I f\, d\mu \le C\mu(I) \qquad \text{and} \qquad \int_J f\, d\mu \le C\mu(J).$$

If the center of Q is in $I \cup J$, one may assume its sidelength $\ell(Q) \ge 1/m$. Since $\mu(I_j) \le \mu(J_j)$, one has

$$\int_Q |f - a|\, d\mu \le 3 \int_{Q \cap J} |f - a|\, d\mu.$$

Now, since the measure μ decays exponentially in J, the function $f \mid J$ looks like a logarithm. Hence, if $a = \min\{j : J_j \cap Q \neq \emptyset\}$, one has

$$\int_{Q\cap J} |f - a|\, d\mu \leq C\mu(Q \cap J).$$

So, (26) is proved. Notice also that if one has a cube Q that contains either K, I, or J, then (26) holds with $a = 0$.

Taking

$$Q = \big\{(x, y) : |x| \leq 1 + 1/2m,\ |y| \leq 1 + 1/2m\big\},$$

we have that for any $a \in \mathbb{R}$,

$$\mu\big(\{x \in Q : |f(x) - a| > m/3\}\big) \geq \frac{2^{-m}}{3m} \geq \frac{\mu(Q)}{4m}.$$

As in the proofs of Theorems 5 and 6, we now apply this construction to a sequence of squares. Consider a collection $\{Q_n : n = 1, 2, \dots\}$ of disjoint squares whose left side is in the y-axis and that satisfies

$$\operatorname{dist}(Q_n, Q_m) \geq 2\max\big\{\ell(Q_n), \ell(Q_m)\big\}, \quad n \neq m.$$

Let μ_n, f_n be the measure and the function given by the construction above in the square Q_n with $m = n$. Set

$$\mu = \sum \mu_n, \qquad f = \sum f_n.$$

One has $f \in \mathrm{BMO}_c(\mu)$. To see this, observe that if a cube Q centered at a point in Q_k intersects some other Q_j, $j \neq k$, then $Q_k \subset Q$ and, moreover, Q contains the K or J piece of the cube Q_j. So, as remarked above, one has

$$\int_Q f_j\, d\mu_j \leq C\mu_j(Q)$$

and adding up,

$$\int_Q f\, d\mu \leq C\mu(Q).$$

So, $f \in \mathrm{BMO}_c(\mu)$. On the other hand, as before, for any $a \in \mathbb{R}$, one has

$$\mu\big\{x \in Q_n : |f(x) - a| > n/3\big\} \geq \frac{\mu(Q_n)}{4n},$$

and the John-Nirenberg inequality fails.

Appendix: John-Nirenberg theorem on spaces of homogeneous type

Let (X, d, μ) be a space of homogeneous type in the sense of Coifman and Weiss [CW, p. 587]. Thus, the quasidistance d satisfies

$$d(x, y) \leq K\big(d(x, z) + d(z, y)\big),$$

and the positive Borel measure μ is doubling

$$\mu\big(B(x,2r)\big) \le D\mu\big(B(x,r)\big),$$

where $B(x,r) = \{y \in X : d(x,y) < r\}$ is the ball centered at x and radius $r > 0$.

Remark. There exists a constant α such that if $x \in B(a,R)$ and $0 < \rho \le 12KR/5$, then $B(x,\rho) \subset B(a,\alpha R)$. (Take $\alpha = K((12K/5)+1)$.)

Our goal is to check the following.

THEOREM A. *There exist two positive constants β and b such that for any $f \in$ BMO(X) and for any ball $S \subset X$, one has*

$$\mu\big(\{x \in S : |f - f_S| > \lambda\}\big) \le \beta \exp\left\{-\frac{b\lambda}{K\|f\|_*}\right\}\mu(S), \quad \textit{for all } \lambda > 0. \tag{27}$$

Proof. We guess that the proof we present here is implicit in the work of Coifman and Weiss [CW, p. 594, footnote]. However, since they didn't write it explicitly, there has been some confusion in the literature.

We follow the standard stopping time argument; that is, we assume that λ is large enough and fix some λ_1. Then we study the sets $\{x \in S : |f(x) - f_S| \le \lambda_1\}$, $\{x \in S : |f(x) - f_S| \le 2\lambda_1\}$, up to $\{x \in S : |f(x) - f_S| \le m\lambda_1 \simeq \lambda\}$.

In showing (27), we assume $\|f\|_* \le 1$ and fix $S = B(a,R)$. We define a maximal operator associated to S (if we replace S by another ball, then the maximal operator changes),

$$M_S f(x) = \sup\left\{\frac{1}{\mu(B)}\int_B |f(y) - f_S|\,d\mu(y) : B \text{ ball, } x \in B,\ B \subset B(a,\alpha R)\right\}.$$

Using a Vitali-type covering lemma, one can prove that

$$\mu\big(\{x : M_S f(x) > t\}\big) \le \frac{A}{t}\mu(S),$$

where A is a constant that only depends on K and D but not on S.

Take $\lambda_0 > A$. Consider the open set $U = \{x : M_S f(x) > \lambda_0\}$. We have $\mu(U \cap S) \le (A/\lambda_0)\mu(S) < \mu(S)$, and therefore $S \cap U^c \ne \emptyset$.

Define $r(x) = (1/(5K))\,\mathrm{dist}(x,U^c)$. If $x,\ y \in S$, then $d(x,y) \le 2KR$. Since $U^c \cap S \ne \emptyset$, if $x \in S$, we have $r(x) \le 2KR/(5K) = 2R/5$.

Clearly,

$$U \cap S \subset \bigcup_{x \in U \cap S} B\big(x, r(x)\big) \subset U.$$

Again by a Vitali-type covering lemma (e.g., see [CW, Theorem 3.1]), we can select a finite or countable sequence of disjoint balls $\{B(x_j, r_j)\}$ such that $r_j = r_j(x)$ and

$$U \cap S \subset \bigcup_j B\big(x_j, 4Kr_j\big) \subset U.$$

On the other hand,

$$B\big(x_j, 6Kr_j\big) \cap U^c \neq \emptyset$$

and

$$B\big(x_j, 6Kr_j\big) \subset B(a, \alpha R) \quad \text{because } 6Kr_j \leq \frac{12KR}{5}.$$

Thus, we get

$$\frac{1}{\mu\big(B\big(x_j, 6Kr_j\big)\big)} \int_{B(x_j, 6Kr_j)} |f - f_S|\, d\mu \leq \lambda_0,$$

and consequently, if we write $S_j^{(1)} = B(x_j, 4Kr_j)$, we obtain

$$|f_S - f_{S_j}| \leq \frac{1}{\mu(S_j)} \int_{S_j} |f - f_S| \leq c_0 \lambda_0 := \lambda_1,$$

because μ is a doubling measure.

By the differentiation theorem,

$$|f(x) - f_S| \leq \lambda_0 \quad \text{for } \mu\text{-a.e. } x \in S \setminus \bigcup_j S_j^{(1)}.$$

Moreover,

$$\sum_j \mu\big(S_j^{(1)}\big) \leq C \sum_j \mu\big(B\big(x_j, r_j\big)\big) \leq C\mu(U) \leq \frac{CA}{\lambda_0}\mu(S) = \frac{A'}{\lambda_0}\mu(S).$$

Now, we do the same construction for each $S_j^{(1)}$. Again

$$\Big| f(x) - f_{S_j^{(1)}} \Big| \leq \lambda_0 \quad \text{for } \mu\text{-a.e. } x \in S_j^{(1)} \setminus \bigcup_i S_i^{(2)},$$

and therefore for these points

$$|f(x) - f_S| \leq \Big| f(x) - f_{S_j^{(1)}} \Big| + \Big| f_{S_j^{(1)}} - f_S \Big| \leq \lambda_0 + c_0\lambda_0 \leq 2c_0\lambda_0 := 2\lambda_1.$$

It is clear (taking $\lambda_0 = 2A'$) that

$$\mu\left(\bigcup_i S_i^{(2)} \right) \leq \sum_j \frac{A'}{\lambda_0} \mu\big(S_j^{(1)}\big) \leq \left(\frac{A'}{\lambda_0} \right)^2 \mu(S) = 2^{-2}\mu(S).$$

By continuing this process, we would finish the proof. $\square$

Recently, in [B] and [MP], other correct proofs of the John-Nirenberg inequality for homogeneous-type spaces have been presented.

References

[B] S. M. Buckley, *Inequalities of John-Nirenberg type in doubling spaces,* preprint.

[CW] R. R. Coifman and G. Weiss, *Extensions of Hardy spaces and their use in analysis*, Bull. Amer. Math. Soc. **83** (1977), 569–645.

[GJ] J. B. Garnett and P. Jones, *The distance in BMO to* L^∞, Ann. of Math. (2) **108** (1978), 373–393.

[G] M. de Guzmán, *Differentiation of Integrals in* $\mathbb{R}^n$, Lecture Notes in Math. **481**, Springer, Berlin, 1975.

[J] J.-L. Journé, *Calderón-Zygmund Operators, Pseudodifferential Operators and the Cauchy Integral of Calderón*, Lecture Notes in Math. **994**, Springer, Berlin, 1983.

[MP] P. MacManus and C. Perez, *Trudinger inequalities without derivatives*, preprint.

[M] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability,* Cambridge Stud. Adv. Math. **44**, Cambridge Univ. Press, Cambridge, 1995.

[NTV] F. Nazarov, S. Treil, and A. Volberg, *Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces,* Internat. Math. Res. Notices **1997**, 703–726.

[S] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions,* Princeton Math. Ser. **30**, Princeton Univ. Press, Princeton, 1970.

[T] X. Tolsa, L^2*-boundedness of the Cauchy integral operator for continuous measures*, Duke Math. J. **98** (1999), 269–304.

[V] J. Verdera, *On the* $T(1)$*-theorem for the Cauchy integral*, to appear in Ark. Mat.

[W] I. Wik, *On John and Nirenberg's theorem,* Ark. Mat. **28** (1990), 193–200.

Mateu, Nicolau, and Orobitg: Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain; mateu@mat.uab.es; artur@mat.uab.es; orobitg@mat.uab.es

Mattila: Department of Mathematics, University of Jyväskylä, P.O. Box 35, SF-40351 Jyväskylä, Finland; pmattila@math.jyu.fi