Regularity of the Hardy–Littlewood
maximal operator on block decreasing functions

by

J. M. Aldaz (Madrid) and F. J. Pérez Lázaro (Logroño)

Abstract. We study the Hardy–Littlewood maximal operator defined via an unconditional norm, acting on block decreasing functions. We show that the uncentered maximal operator maps block decreasing functions of special bounded variation to functions with integrable distributional derivatives, thus improving their regularity. In the special case of the maximal operator defined by the ℓ_∞-norm, that is, by averaging over cubes, the result extends to block decreasing functions of bounded variation, not necessarily special.

1. Introduction. The usefulness of the Hardy–Littlewood maximal function M stems basically from two facts:

1) It is larger than the given function, since $|f| \le Mf$ a.e., but it is not too large, since $\|Mf\|_p \le c_p\|f\|_p$ for $1 < p \le \infty$, while on L^1, M satisfies a weak type $(1,1)$ inequality.
2) It is more regular than the original function: If f is measurable, then Mf is lower semicontinuous.

The fact that Mf controls f and its averages over balls (by definition), together with its L^p boundedness, leads to its frequent use in chains of inequalities, while its lower semicontinuity allows one to decompose its level sets using dyadic cubes. This is the basis of the often applied Calderón–Zygmund decomposition: Utilize Mf as a proxy for f, splitting the open set $\{Mf > t\}$ into suitable disjoint cubes. This might be impossible to do directly with $\{f > t\}$, since in principle this set is merely measurable.

Regarding derivatives, the study of the regularity properties of the Hardy–Littlewood maximal function is much more recent. It was initiated by Juha Kinnunen [Ki], who proved that the centered maximal operator is bounded on the Sobolev spaces $W^{1,p}(\mathbb{R}^d)$ for $1 < p \le \infty$. Since then, a good

[253]

deal of work has been done within this line of research (cf. for instance [KiLi], [HaOn], [KiSa], [Lu], [Bu], [Ko1], [Ko2], [Ta]). The overall emerging pattern, concerning regularity, seems to be that the worse f is, the greater the improvement of Mf when compared to f. In fact, if the functions are "good", there may not be any improvement at all. For instance, the maximal function of a C^1 function need not be C^1, while the maximal function of a Lipschitz or an α-Hölder function will be Lipschitz or α-Hölder, but in general no better than that, though constants will be lowered (so there is some quantitative improvement; cf. [ACP]).

In the present paper, only the uncentered maximal function will be considered, since it has better regularity properties than its centered relative. A model example of this fact is the following: Let f be the characteristic function of the unit interval in the real line. Then both f and the centered maximal function of f are discontinuous at 0 and 1, while Mf is Lipschitz on $\mathbb{R}$ with constant 1.

Here, the dimension d will always be at least two. The one-dimensional case was studied in [AlPe1]; there we showed that given an arbitrary interval $I \subset \mathbb{R}$, if $f : I \to \mathbb{R}$ is of bounded variation and Df denotes its distributional derivative, then Mf is absolutely continuous and $\|DMf\|_{L^1(I)} \leq |Df|(I)$, where $|Df|$ is the total variation of Df (cf. [AlPe1, Theorem 2.5]). Hence, M improves the regularity of BV functions, so, just as in the case of the Calderón–Zygmund decomposition, Mf can be used as a proxy for f, with the function DMf replacing the measure Df. Along these lines, a Landau type inequality is presented in [AlPe1, Theorem 5.1]. Of course, having a function as derivative, instead of a singular measure, makes it possible to consider $\|DMf\|_{L^p}$ for $p > 1$. In turn, this suggests the possibility of obtaining inequalities of Gagliardo–Nirenberg–Sobolev for functions less regular than those in the Sobolev classes.

Thus, it is interesting to try to find higher-dimensional versions of [AlPe1, Theorem 2.5]. In [AlPe2, Theorem 2.19 and Remark 2.20], we showed that the local maximal function M_R (where the radii of balls are bounded above by R) maps $\mathrm{BV}(\mathbb{R}^d)$ boundedly into $L^1(\mathbb{R}^d)$, with constant of the order of $\log R$. In fact, even the local, strong maximal function is bounded from $\mathrm{BV}(\mathbb{R}^d)$ into $L^1(\mathbb{R}^d)$ (with constant of the order of $\log^d R$). However, the derivative of the local, strong maximal function is not always comparable to $\|f\|_{\mathrm{BV}}$ (cf. [AlPe2, Theorem 2.21]), so we have unboundedness of this operator on BV.

Regarding the maximal operator, two questions remain open. First, whether it regularizes functions in BV, so if $f \in \mathrm{BV}$, then Mf is ACL (absolutely continuous on lines, the natural generalization of absolute continuity to $d > 1$), and second, whether the size of DMf is not too large, i.e., there exists a constant c such that $|DMf|(\mathbb{R}^d) \leq c\|f\|_{\mathrm{BV}}$. At this point both questions

seem to be intractable. We mention, after recalling that $W^{1,1}(\mathbb{R}^d) \subset \mathrm{BV}(\mathbb{R}^d)$, a related and simpler question from [HaOn, Question 1]: Is the operator $f \mapsto |\nabla Mf|$ bounded from the Sobolev space $W^{1,1}(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$?

The results presented here are obtained by restricting ourselves to a smaller class of functions: the block decreasing (or unconditional decreasing) functions of bounded variation (which, in particular, contain the $W^{1,1}(\mathbb{R}^d)$ block decreasing functions). For these functions Hajłasz and Onninen's question has a positive answer.

In general, the balls we use when defining the maximal operator are unconditional. Specializing to cubes, we obtain stronger results.

More precisely, let $f \geq 0$ be block decreasing. We show that if Mf is defined using unconditional balls, then the variation of f controls the variation of Mf (Theorem 7). If f has finite variation, then Mf is continuous a.e. with respect to the $(d-1)$-dimensional Hausdorff measure (no unconditionality needed here, cf. Theorem 8). Further assumptions on f lead to better results: If f is also of special bounded variation (so the derivative Df has no Cantor part), then Mf has an integrable weak gradient (cf. Theorem 11). When the maximal function is defined using cubes, Mf has a weak gradient even if Df has a nontrivial Cantor part (cf. Theorem 12). Identical results hold for the local maximal operator M_R. Using the fact that M_R maps $\mathrm{BV}(\mathbb{R}^d)$ boundedly into $L^1(\mathbb{R}^d)$, we obtain boundedness results for M_R from the nonnegative block decreasing functions into the functions of bounded variation (cf. Corollary 13).

2. Definitions and results

DEFINITION 1. A function $f : \mathbb{R}^d \to [-\infty, \infty]$ is *unconditional* if for all $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ we have $f(x_1, \ldots, x_d) = f(|x_1|, \ldots, |x_d|)$.

DEFINITION 2. Let $f : \mathbb{R}^d \to [-\infty, \infty]$ be unconditional. Then f is *block decreasing* if the restriction of f to the nonnegative cone $[0, \infty)^d$ is decreasing in each variable.

REMARK 3. Observe that being unconditional depends on the system of coordinates chosen for $\mathbb{R}^d$. A rotation, for instance, may destroy this property. The term "block decreasing" comes from the statistical literature, while in functional analysis "unconditional decreasing" is used instead. Radial functions with respect to unconditional norms in $\mathbb{R}^d$ (for instance, the ℓ_p-norms, $1 \leq p \leq \infty$) are block decreasing. But in general, a block decreasing function may have nonconvex level sets, in which case it is not radial with respect to any norm. On the other hand, a norm ν may fail to be unconditional; then a radial function with respect to ν is not block decreasing.

Given a norm μ in $\mathbb{R}^d$, we denote by $B_\mu(y, \delta) := \{x \in \mathbb{R}^d : \mu(x - y) \leq \delta\}$ the closed μ-ball centered at $y \in \mathbb{R}^d$ and of radius $\delta > 0$. Absolute values around a set denote its k-dimensional Lebesgue measure. While the dimension k is not indicated in $|A|$, it will usually be clear from the context. When doubts may arise, we explicitly state what k is.

DEFINITION 4. Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Then for all $x \in \mathbb{R}^d$, the *uncentered maximal function* $M_\mu f$ is defined by

$$(2.1) \qquad M_\mu f(x) = \sup_{\{y \in \mathbb{R}^d,\, \delta > 0\,:\, x \in B_\mu(y, \delta)\}} \frac{1}{|B_\mu(y, \delta)|} \int_{B_\mu(y, \delta)} |f(u)|\, du.$$

Let $R > 0$ be fixed. The *local* maximal function $M_{R,\mu} f$ is defined by imposing an extra condition on the radius of balls: $\delta \leq R$. Apart from that, the definition is identical to (2.1).

We write M_p instead of M_μ in the special case where $1 \leq p \leq \infty$ and μ is an ℓ_p-norm, i.e., given by $\|x\|_p := (|x_1|^p + |x_2|^p + \cdots + |x_d|^p)^{1/p}$ when $1 \leq p < \infty$, and by $\|x\|_\infty := \max_{1 \leq i \leq d}\{|x_1|, \ldots, |x_d|\}$. Likewise, we write B_p instead of B_μ for balls.

REMARK 5. Since all norms on $\mathbb{R}^d$ are equivalent, maximal functions defined using different norms are always pointwise comparable. However, comparability yields no information about regularity properties, or the size of derivatives, if they exist in some appropriate sense.

From now on, we assume that *all* functions appearing in this paper are *locally integrable*, including functions of the form $M_\mu f$. It might happen that for some $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, $M_\mu f \equiv \infty$. But then $M_\mu f$ is constant, and thus its variation is zero under any reasonable notion of variation. So we exclude this trivial case from any further consideration. It follows that $M_\mu f$ is finite almost everywhere. But in general, something else is needed to have local integrability of $M_\mu f$. For the type of functions studied in this paper, that is, for block decreasing functions of bounded variation, the local integrability of $M_\mu f$ is easy to check (cf. Lemma 14 below); but some auxiliary results are stated in greater generality, and for these we *assume* local integrability of $M_\mu f$ from the start.

Let $\Omega \subset \mathbb{R}^d$ be open. The following definition is taken from [AFP, p. 119].

DEFINITION 6. For $f \in L^1_{\mathrm{loc}}(\Omega)$, the *variation* $V(f, \Omega)$ of f in Ω is given by

$$V(f, \Omega) := \sup\Big\{ \int_\Omega f \operatorname{div} \phi\, dx : \phi \in [C^1_c(\Omega)]^d,\ \|\phi\|_\infty \leq 1 \Big\}.$$

Suppose $V(f, \Omega) < \infty$, i.e., f is of finite variation. If additionally $f \in L^1(\Omega)$, we write $f \in \mathrm{BV}(\Omega)$, where BV stands for bounded variation. Integration by parts shows that if f is continuously differentiable in Ω, then

$V(f, \Omega) = \int_\Omega |\nabla f| \, dx$. By Proposition 3.6, p. 120 of [AFP], $V(f, \Omega) < \infty$ if and only if there exists an $\mathbb{R}^d$-valued Radon measure $Df = (D_1 f, \ldots, D_d f)$ on Ω such that

$$\int_\Omega f \operatorname{div} \phi \, dx = -\int_\Omega \phi \, dDf \quad \forall \phi \in [C_c^1(\Omega)]^d.$$

That is, the distributional derivative is representable by a Radon measure Df on Ω with total variation $|Df|(\Omega) < \infty$. Furthermore, $|Df|(\Omega) = V(f, \Omega)$. The norm of $f \in \mathrm{BV}(\Omega)$ is defined by $\|f\|_{\mathrm{BV}(\Omega)} := \|f\|_{L^1(\Omega)} + |Df|(\Omega)$. Note that $W^{1,1}(\Omega) \subset \mathrm{BV}(\Omega)$, and the Sobolev norm on $W^{1,1}(\Omega)$ is simply the restriction to the latter space of the BV norm. Note also that a function $f \geq 0$ of bounded variation on $\mathbb{R}^d$ need not be bounded (provided $d \geq 2$, as we always assume in this paper); well known examples exist in $W^{1,1}(\mathbb{R}^d)$. However, if f is also block decreasing, then the hypothesis $V(f, \mathbb{R}^d) < \infty$ entails that the *precise representative* f^* of f (defined by taking the limsup in the Lebesgue Differentiation Theorem) must be finite except perhaps on a negligible $(d-1)$-dimensional Hausdorff measurable set. Since f is block decreasing, either $f^*(0) = \infty$ or f^* is bounded. But we must allow the possibility that $f^*(0) = \infty$, so we will consider functions with values in $[0, \infty]$. In general we do not assume that f is integrable.

The first theorem of the paper states that the variation of $M_\mu f$ is controlled by the variation of f, and the same happens with the local maximal function $M_{R,\mu} f$, with constant independent of R.

THEOREM 7. *Let $f : \mathbb{R}^d \to [0, \infty]$ be a block decreasing function and let μ be an unconditional norm in $\mathbb{R}^d$. Then $V(M_\mu f, \mathbb{R}^d) \leq c(\mu, d) V(f, \mathbb{R}^d)$, and, with the same constant $c(\mu, d)$, $V(M_{R,\mu} f, \mathbb{R}^d) \leq c(\mu, d) V(f, \mathbb{R}^d)$ for every $R > 0$.*

The next theorem states that the maximal function of a block decreasing function of finite variation is continuous, except perhaps on a negligible $(d-1)$-dimensional Hausdorff measurable set. The same happens with the local maximal function. Observe that unconditionality of the norm is not assumed here.

THEOREM 8. *Let $f : \mathbb{R}^d \to [0, \infty]$ be a block decreasing function such that $V(f, \mathbb{R}^d) < \infty$. Let μ be a norm in $\mathbb{R}^d$, and let $R > 0$. Then $M_\mu f$ and $M_{R,\mu} f$ are continuous a.e. with respect to $\mathcal{H}^{d-1}$.*

It is well known (see [AFP, pp. 184–186]) that if $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and $V(f, \mathbb{R}^d) < \infty$, then the distributional derivative Df of f can be decomposed into three parts,

$$Df = D^a f + D^j f + D^c f,$$

where $D^a f$ is absolutely continuous, $D^j f$ is the *jump part* of Df, its restriction to the jump set of f (to be defined below, cf. Definition 24), and $D^c f$ is the *Cantor part* of the measure, the singular part of Df that lives on the set where f is approximately continuous (cf. [AFP, p. 160] for the definition of approximate continuity). The functions $f : \mathbb{R}^d \to [-\infty, \infty]$ of bounded variation for which $D^c f = 0$ are called functions of *special bounded variation*, and denoted by $\mathrm{SBV}(\mathbb{R}^d)$. If both $D^c f$ and $D^j f$ vanish, then f is in the Sobolev space $W^{1,1}(\mathbb{R}^d)$, and $Df(A) = \int_A \nabla f$ for every measurable set A.

EXAMPLE 9. In order to illustrate some notions that have already been defined and others that will appear later on, and also to explain the terminology, consider the following simple example. Let f be the characteristic function of the unit square $[-1/2, 1/2]^2 \subset \mathbb{R}^2$. Then f is a block decreasing, BV function, with $|Df| = |D^j f|$ and $|Df|(\mathbb{R}^2) = 4$, the length of the boundary of the square. This boundary is also the jump set of f (cf. Definition 24 below). And $|Df| = |D^j f|$ is just the linear Lebesgue measure on the jump set. Since $D^c f = 0$, f is actually a SBV function. The *centered* maximal function of f has the same jump set as f, though the jumps are smaller, and the uncentered maximal function $M_\infty f$ associated to cubes has empty jump set. For a general norm μ, we know from the preceding theorem that the jump set of $M_\mu f$ has linear measure at most zero. In fact, it is easy to see that the jump set of $M_\mu f$ contains, at most, the four corners of the square.

The jump set of a function is obviously disjoint from the set of its continuity points. From the preceding theorem, together with the fact that if $E \subset \mathbb{R}^d$ has $\mathcal{H}^{d-1}$-measure zero then $D^j f(E) = 0$ (cf. [AFP, formula (3.90), p. 184]), we obtain the following corollary. It says that essentially (with respect to $\mathcal{H}^{d-1}$) $M_\mu f$ has no jumps.

COROLLARY 10. *Let f be a locally integrable, nonnegative block decreasing function such that $V(f, \mathbb{R}^d) < \infty$. Let μ be a norm in $\mathbb{R}^d$, and let $R > 0$. Then $D^j M_\mu f = 0$, and $D^j M_{R,\mu} f = 0$.*

The integrability of f is not assumed in the next result, so it deals with nonnegative block decreasing functions slightly more general than those in $\mathrm{SBV}(\mathbb{R}^d)$.

THEOREM 11. *Let $f : \mathbb{R}^d \to [0, \infty]$ be a block decreasing function with $V(f, \mathbb{R}^d) < \infty$ and $|D^c f| = 0$. Let μ be an unconditional norm in $\mathbb{R}^d$, and let $R > 0$. Then $M_\mu f$ has a weak gradient $\nabla M_\mu f$ in L^1, and there exists a constant $c(\mu, d) > 0$ such that*

$$\int_{\mathbb{R}^d} |\nabla M_\mu f(x)| \, dx \le c(\mu, d) V(f, \mathbb{R}^d).$$

The same result, with the same constant $c(\mu, d)$, holds for $M_{R,\mu} f$.

This theorem gives a positive answer to [HaOn, Question 1] in the special case of functions f with $|f|$ block decreasing. In fact, the condition $f \in W^{1,1}(\mathbb{R}^d)$ from [HaOn, Question 1] is relaxed (to $V(f, \mathbb{R}^d) < \infty$ and $|D^c f| = 0$).

If μ happens to be the ℓ_∞-norm, then the "no Cantor part" hypothesis $|D^c f| = 0$ can be dispensed with. The reason for this is that block decreasing functions are particularly well adapted to arguments using cubes, or more generally, rectangles with sides parallel to the axes. Even though the next result is stated for cubes only, it also holds for any norm defined using a fixed rectangle (with sides parallel to the axes).

THEOREM 12. *Let $f : \mathbb{R}^d \to [0, \infty]$ be a block decreasing function such that $V(f, \mathbb{R}^d) < \infty$, let $M_\infty f$ be the maximal function of f defined using cubes, i.e., ℓ_∞-balls, and let $R > 0$. Then $M_\infty f$ has a weak gradient $\nabla M_\infty f \in L^1(\mathbb{R}^d)$, and*

$$\int_{\mathbb{R}^d} |\nabla M_\infty f(x)|\, dx \leq c_d V(f, \mathbb{R}^d).$$

The same result holds for $M_{R,\infty} f$.

Denote by $\mathrm{BD}(\mathbb{R}^d)$ the cone of nonnegative block decreasing functions in $\mathrm{BV}(\mathbb{R}^d)$, measured with the BV norm. The preceding theorems, together with Theorem 2.19 and Remark 2.20 of [AlPe2], entail the following boundedness results on $\mathrm{BD}(\mathbb{R}^d)$ for the local maximal operator.

COROLLARY 13. *Given an unconditional norm μ on $\mathbb{R}^d$, the local maximal operator $M_{R,\mu}$ is bounded from $\mathrm{BD}(\mathbb{R}^d)$ to $\mathrm{BV}(\mathbb{R}^d)$. More precisely, there exists a constant $c = c(\mu, d) > 0$ such that for all $R > 0$ and all $f \in \mathrm{BD}(\mathbb{R}^d)$ we have*

$$(2.2) \qquad \|M_{R,\mu} f\|_{\mathrm{BV}(\mathbb{R}^d)} \leq c(\|f\|_{\mathrm{BV}(\mathbb{R}^d)} + \|f\|_{L^1(\mathbb{R}^d)} \log^+ R).$$

If $\mu = \|\cdot\|_\infty$, then $M_{R,\mu}$ is bounded from $\mathrm{BD}(\mathbb{R}^d)$ to $W^{1,1}(\mathbb{R}^d)$, so there exists a constant $c = c(\mu, d) > 0$ such that for all $R > 0$ and all $f \in \mathrm{BD}(\mathbb{R}^d)$,

$$(2.3) \qquad \|M_{R,\infty} f\|_{W^{1,1}(\mathbb{R}^d)} \leq c(\|f\|_{\mathrm{BV}(\mathbb{R}^d)} + \|f\|_{L^1(\mathbb{R}^d)} \log^+ R).$$

Of course, if R is fixed, then (2.2) reduces to

$$(2.4) \qquad \|M_{R,\mu} f\|_{\mathrm{BV}(\mathbb{R}^d)} \leq c\|f\|_{\mathrm{BV}(\mathbb{R}^d)},$$

though perhaps with a different c, and likewise for (2.3), in the case of cubes.

3. The maximal function of a block decreasing function. In this and the following sections, lemmas and proofs will refer exclusively to the maximal operator M, since they are exactly the same for the local operator M_R. The only exception occurs in Lemma 35, which is valid in the nonlocal case, under fewer hypotheses. It has to do with the Lipschitz behavior of Mf on some "good sets", for a locally integrable f. Since as R

becomes small, $M_R f$ looks more like f, any improvement of $M_R f$ over f will tend to disappear as $R \to 0$. Thus, the local case requires additional assumptions, and hence it is treated in a different lemma.

In this section we prove that if f is nonnegative and block decreasing, then $M_\mu f$ is block decreasing (see Lemma 19 below; of course, $M_\mu f$ is always nonnegative). First we deal with the local integrability of $M_\mu f$.

LEMMA 14. *Let* $f : \mathbb{R}^d \to [0, \infty]$ *be a block decreasing function such that* $V(f, \mathbb{R}^d) < \infty$, *and let* μ *be any norm in* $\mathbb{R}^d$. *Then* $M_\mu f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$.

Proof. By pointwise comparability of maximal functions associated to different norms, it is enough to prove the result in the ℓ_∞ case. Fix a ball $B = B_\infty(0, r)$ (centered at the origin) and note that for any $x \in B$, to estimate $M_\infty f(x)$ it suffices to average over cubes contained in B, by the block decreasing property of f. Now by Sobolev embedding for functions in $\mathrm{BV}(B_\infty(0, r))$ (cf. for instance, [AFP, Corollary 3.49, p. 152]), we have $f \in L^{d/(d-1)}(B_\infty(0, r))$. Using the boundedness of the maximal operator when $p > 1$, we get $M_\infty f \in L^{d/(d-1)}(B_\infty(0, r)) \subset L^1(B_\infty(0, r))$. Since r is arbitrary, it follows that $M_\infty f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. ∎

On $\mathbb{R}$, being unconditional is the same as being even, and being block decreasing, the same as being even and unimodal (decreasing on $(0, \infty)$). So the next lemma is trivial, and we omit the proof. We do mention that a (non-trivial) higher-dimensional version is known in the literature as Anderson's theorem (cf. [An] or [Ga, Theorem 11.1]).

LEMMA 15. *Let* $f : \mathbb{R} \to [0, \infty]$ *be block decreasing. Then for every* $\delta > 0$, *the function* $g_\delta(x) := \int_{x-\delta}^{x+\delta} f(u) \, du$ *is block decreasing.*

It follows from the definition of $V(f, \Omega)$ in terms of the distributional derivative Df that functions equal a.e. have the same variation. Since the measure Df may have a singular part, it is nevertheless useful to choose an everywhere defined representative of f.

DEFINITION 16. Let B_μ denote a generic ball defined using the norm μ. The *precise representative* f^* of f is

$$f^*(x) := \limsup_{|B_\mu| \downarrow 0, \, x \in B_\mu} \frac{1}{|B_\mu|} \int_{B_\mu} f(y) \, dy.$$

The notation does not reflect the fact that f^* depends on μ, since this will make no difference in the arguments below.

A related notion of precise representative can be obtained by taking the limsup over balls centered at x, instead of balls containing x, as we do above. With either choice of definition, it is not difficult to see that if f is block decreasing, then so is f^* (cf. Lemma 18 below).

From now on, we use the following notation. Let $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, let $i \in \{1, \dots, d\}$, and let $\{e_1, \dots, e_d\}$ be the canonical basis of $\mathbb{R}^d$. We denote by $\hat{x}_i$ the $(d-1)$-dimensional vector obtained from x by removing its ith component. That is, $\hat{x}_i = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)$. To shorten expressions, we write $x = (\hat{x}_i, x_i)$. Even though the notation may suggest otherwise, we do emphasize the fact that the order of the coordinates in $x = (\hat{x}_i, x_i)$ is unaltered. We also write $f(x) = f(\hat{x}_i, x_i)$ rather than $f(x) = f((\hat{x}_i, x_i))$.

LEMMA 17. *Let $f : \mathbb{R}^d \to [0, \infty]$ be block decreasing, let μ be an unconditional norm in $\mathbb{R}^d$, and let $\delta > 0$. Then $g(x) := \int_{B_\mu(x,\delta)} f(u)\, du$ is block decreasing.*

Proof. Since the argument is the same for each coordinate, we focus on the last one. It is enough to prove that if $x = (\hat{x}_d, x_d) \in [0, \infty)^d$ and $h \geq 0$, then $g(\hat{x}_d, x_d + h) \leq g(x)$. But this follows from Fubini's Theorem: Write

$$P := \{\hat{u}_d \in \mathbb{R}^{d-1} : \text{there exists a real number } u_d$$
$$\text{such that } (\hat{u}_d, u_d) \in B_\mu(0, \delta)\},$$

and

$$S_{\hat{u}_d} = \{t \in \mathbb{R} : u = (\hat{u}_d, t) \in B_\mu(0, \delta)\},$$

that is, $S_{\hat{u}_d}$ is the vertical section in $B_\mu(0, \delta)$ associated to $\hat{u}_d$. The assumption that μ is unconditional entails that each $S_{\hat{u}_d}$ is an interval centered at 0. Now

$$g(x) = \int_{B_\mu(0,\delta)} f(u + x)\, du = \int_P \left(\int_{S_{\hat{u}_d}} f(u + x)\, du_d \right) d\hat{u}_d$$

and

$$g(x + he_d) = \int_{B_\mu(0,\delta)} f(u + x + he_d)\, du = \int_P \left(\int_{S_{\hat{u}_d}} f(u + x + he_d)\, du_d \right) d\hat{u}_d.$$

For fixed $\hat{u}_d$ and $\hat{x}_d$ the function $f(\hat{u}_d + \hat{x}_d, \cdot)$ is block decreasing, so by Lemma 15,

$$\int_{S_{\hat{u}_d}} f(\hat{u}_d + \hat{x}_d, u_d + x_d + h)\, du_d \leq \int_{S_{\hat{u}_d}} f(\hat{u}_d + \hat{x}_d, u_d + x_d)\, du_d,$$

and now $g(\hat{x}_d, x_d + h) \leq g(x)$ follows by putting together the last three formulas. ∎

LEMMA 18. *Let μ be an unconditional norm in $\mathbb{R}^d$. If $f : \mathbb{R}^d \to [0, \infty]$ is block decreasing, then so is its precise representative f^*.*

Proof. It is enough to prove that if $x = (\hat{x}_d, x_d) \in [0, \infty)^d$ and $h > 0$, then $f^*(\hat{x}_d, x_d + h) \leq f^*(x)$. But this follows from the previous lemma, applied to any ball $B_\mu(a, \delta)$ containing $(\hat{x}_d, x_d + h)$ and with radius $\delta < \mu((h/2)e_d)$. ∎

LEMMA 19. *Let $f : \mathbb{R}^d \to [0, \infty]$ be block decreasing and let the norm μ be unconditional. Then $M_\mu f$ is a block decreasing function.*

Proof. It is enough to prove that if $x = (\hat{x}_d, x_d) \in [0, \infty)^d$ and $h \geq 0$, then $M_\mu f(\hat{x}_d, x_d + h) \leq M_\mu f(x)$, and to deduce this it suffices to show that given an arbitrary ball $B_\mu(a, \delta)$ containing $x + he_d$, we have

$$(3.1) \qquad \frac{1}{|B_\mu(a, \delta)|} \int\limits_{B_\mu(a,\delta)} f(u) \, du \leq M_\mu f(x).$$

Let us see why. Since $x + he_d \in B_\mu(a, \delta)$, by unconditionality $(\hat{x}_d, 0) \in B_\mu((\hat{a}_d, 0), \delta)$. Now if $a_d \leq x_d$, then $x_d \in [a_d, x_d + h)$, so $x \in B_\mu(a, \delta)$ by convexity of the ball. Thus, (3.1) holds in this case. And if $a_d > x_d$, then $x \in B_\mu((\hat{a}_d, x_d), \delta)$, and (3.1) follows from Lemma 17, since we just lowered the ball in the vertical direction. ∎

4. Controlling the variation of block decreasing functions. The purpose of this section is to find a quantity equivalent to the variation (cf. Definition 6), and easier to compute for block decreasing functions. Denote by $f_{\hat{x}_i}$ the one-dimensional function $f_{\hat{x}_i}(x_i) := f(x)$. We use the fact that finite variation can be characterized via the variation along the coordinate axes (cf. [AFP, p. 196], or [EvGa, §5.10]). Suppose f is C^1. Integrating pointwise the ℓ_1-norm and the ℓ_2-norm of its gradient, and using $\|\cdot\|_2 \leq \|\cdot\|_1 \leq \sqrt{d}\,\|\cdot\|_2$, we obtain

$$(4.1) \qquad \int\limits_{\mathbb{R}^d} |\nabla f(u)| \, du \leq \sum_{i=1}^d \int\limits_{\mathbb{R}^d} |D_i f(u)| \, du \leq \sqrt{d} \int\limits_{\mathbb{R}^d} |\nabla f(u)| \, du,$$

where $D_i f$ denotes the partial derivative of f with respect to x_i, i.e., the derivative of $f_{\hat{x}_i}$. Since for a continuously differentiable f we have $V(f, \mathbb{R}^d) = \int_{\mathbb{R}^d} |\nabla f(u)| \, du$, and for each fixed $\hat{x}_i$ we have $V(f_{\hat{x}_i}, \mathbb{R}^d) = \int_{\mathbb{R}} |D_i f_{\hat{x}_i}(t)| \, dt$, inequality (4.1) and an approximation argument show that

$$(4.2) \qquad V(f, \mathbb{R}^d) \leq \sum_{i=1}^d \int\limits_{\mathbb{R}^{d-1}} V(f_{\hat{x}_i}, \mathbb{R}) \, d\hat{x}_i \leq \sqrt{d}\, V(f, \mathbb{R}^d),$$

and this formula also holds when $V(f, \mathbb{R}^d) = \infty$. From (4.2), it follows that for a block decreasing function

$$(4.3) \quad V(f, \mathbb{R}^d) \leq 2 \sum_{i=1}^d \int\limits_{\mathbb{R}^{d-1}} [f_{\hat{x}_i}(0^+) - f_{\hat{x}_i}(\infty)] \, d\hat{x}_i$$

$$(4.4) \qquad = 2^d \sum_{i=1}^d \int\limits_{[0,\infty)^{d-1}} [f(\hat{x}_i, 0^+) - f(\hat{x}_i, \infty)] \, d\hat{x}_i \leq \sqrt{d}\, V(f, \mathbb{R}^d),$$

where $f(\hat{x}_i, 0^+) := \lim_{t \to 0^+} f(\hat{x}_i, t)$ and $f(\hat{x}_i, \infty) := \lim_{t \to \infty} f(\hat{x}_i, t)$.

Next we show that the value at infinity of a block decreasing function of finite variation is the same in essentially all directions. Exceptions may occur, though, if at least one coordinate remains fixed at 0.

LEMMA 20. *Let $f : \mathbb{R}^d \to [0, \infty]$ be a block decreasing function such that $V(f, \mathbb{R}^d) < \infty$. Then $\inf_{\mathbb{R}^d} f = \lim_{t\to\infty} f(t, \ldots, t)$. Furthermore, for every $x = (x_1, \ldots, x_d) \in (0, \infty)^d$, and all $i \in \{1 \ldots, d\}$,*

$$(4.5) \qquad \inf_{\mathbb{R}^d} f = \lim_{t\to\infty} f(\hat{x}_i, t).$$

Additionally,

$$(4.6) \qquad \inf_{\mathbb{R}^d} f = \lim_{t\to\infty} M_\mu f(\hat{x}_i, t).$$

Proof. Given $x \in (0, \infty)^d$, let $t = \max\{x_1, \ldots, x_d\}$. Since f is block decreasing, $f(x) \geq f(t, \ldots, t)$, so $\inf_{\mathbb{R}^d} f = \lim_{t\to\infty} f(t, \ldots, t)$, where the limit exists by monotonicity.

Next, we may assume, by symmetry, that $i = d$. Note that if (4.5) fails for a fixed $\hat{x}_d = (x_1, \ldots, x_{d-1})$, then it fails for all $\hat{y}_d \in (0, x_1] \times \cdots \times (0, x_{d-1}]$ by monotonicity in each variable. Thus, it suffices to prove (4.5) for $(x_1, \ldots, x_{d-1})$ in a full measure subset of $(0, \infty)^{d-1}$. We use induction on the dimension d. The result is obvious for $d = 1$, so we assume it holds for $d - 1$ and show that (4.5) also holds for $d \geq 2$. Consider the function of $d - 1$ variables $f_{x_1}(\cdot) := f(x_1, \cdot)$. Clearly, f_{x_1} is a block decreasing function. By (4.3) and (4.4), for almost all $x_1 > 0$ we have $V(f_{x_1}, \mathbb{R}^{d-1}) < \infty$. Using induction, we apply (4.5) to f_{x_1} and conclude that for all $x_2, \ldots, x_{d-1} > 0$,

$$\lim_{t\to\infty} f_{x_1}(t, \ldots, t, t) = \inf_{\mathbb{R}^{d-1}} f_{x_1} = \lim_{t\to\infty} f_{x_1}(x_2, \ldots, x_{d-1}, t).$$

Now if $\inf_{\mathbb{R}^d} f < \inf_{\mathbb{R}^{d-1}} f_{x_1}$, there exists an $N > 0$ such that for all $t' \geq N$, and all $w_2, \ldots, w_d > 0$,

$$f(x_1, w_2, \ldots, w_d) - f(t', t', \ldots, t') \geq \frac{\inf_{\mathbb{R}^{d-1}} f_{x_1} - \inf_{\mathbb{R}^d} f}{2}.$$

Thus, using (4.4) we derive the following contradiction:

$$\infty > V(f, \mathbb{R}^d) \geq \int_{[N,\infty)^{d-1}} [f(\hat{z}_1, 0^+) - f(\hat{z}_1, \infty)] \, d\hat{z}_1$$

$$\geq \int_{[N,\infty)^{d-1}} [f(\hat{z}_1, x_1) - f(N, \ldots, N)] \, d\hat{z}_1$$

$$\geq \frac{\inf_{\mathbb{R}^{d-1}} f_{x_1} - \inf_{\mathbb{R}^d} f}{2} \int_{[N,\infty)^{d-1}} d\hat{z}_1 = \infty.$$

Therefore $\inf_{\mathbb{R}^d} f = \inf_{\mathbb{R}^{d-1}} f_{x_1}$, and (4.5) follows. To obtain (4.6), note that the local integrability of f, together with the existence of a limit at infinity, entails that averages will approach this limit as $t \to \infty$. ∎

5. Variation of the maximal function. We are now ready to prove the first main result.

Proof of Theorem 7. It is clear that if we add a constant to a function, its variation does not change. Recalling that f is locally integrable and non-negative, it is easy to check that $M_\mu f(x) = M_\mu(f - \inf_{\mathbb{R}^d} f)(x) + \inf_{\mathbb{R}^d} f$, so for simplicity we suppose that $\inf_{\mathbb{R}^d} f = 0$. Under the assumptions and with the notation of Lemma 20, this entails that $\lim_{t \to \infty} M_\mu f(\hat{x}_i, t) = 0$ (even if $f \notin L^1(\mathbb{R}^d)$). Now by (4.3) and (4.4) it is enough to check that

$$\sum_{i=1}^d \int_{[0,\infty)^{d-1}} M_\mu f(\hat{x}_i, 0^+) \, d\hat{x}_i \le c(\mu, d) \sum_{k=1}^d \int_{[0,\infty)^{d-1}} f(\hat{x}_k, 0^+) \, d\hat{x}_k.$$

Because of the equivalence of all norms on $\mathbb{R}^d$, and in particular, between μ and $\|\cdot\|_\infty$, it suffices to consider the maximal operator defined by cubes and to prove that

$$(5.1) \qquad \sum_{i=1}^d \int_{[0,\infty)^{d-1}} M_\infty f(\hat{x}_i, 0^+) \, d\hat{x}_i \le c(d) \sum_{k=1}^d \int_{[0,\infty)^{d-1}} f(\hat{x}_k, 0^+) \, d\hat{x}_k.$$

Assume that $i = d$. We want to estimate

$$\int_{[0,\infty)^{d-1}} M_\infty f(y, 0^+) \, dy.$$

To this end, we divide $[0,\infty)^{d-1}$ into $d!$ suitable subsets as follows: Denote by $\mathcal{P}_n$ the set of all permutations of n elements. For each $\sigma \in \mathcal{P}_{d-1}$ we define

$$A_\sigma = \{y \in [0,\infty)^{d-1} : y_{\sigma(1)} \ge \cdots \ge y_{\sigma(d-1)}\}.$$

Then $\bigcup_{\sigma \in \mathcal{P}_{d-1}} A_\sigma = [0,\infty)^{d-1}$. By symmetry, it is enough to consider the identity permutation, the other estimates being the same. So we take

$$(5.2) \qquad A_\sigma = \{y \in [0,\infty)^{d-1} : y_1 \ge \cdots \ge y_{d-1}\}.$$

Fix $y \in A_\sigma$. To estimate $M_\infty f(y, 0)$, let $B_\infty(a, k)$ be a cube (i.e., an ℓ_∞-ball) containing $(y, 0)$. Set $p^{k,y} := (\max\{0, y_1 - k\}, \ldots, \max\{0, y_{d-1} - k\}, 0)$, and note that $0 \le p_i^{k,y} \le |a_i|$ for every $i = 1, \ldots, d-1$. Thus, by Lemma 17 we have

$$M_\infty f(y, 0) = \sup_{k>0} \frac{1}{|B_\infty(p^{k,y}, k)|} \int_{B_\infty(p^{k,y}, k)} f(u) \, du.$$

We introduce the auxiliary endpoints $y_0 := \infty$ and $y_d := 0$. With this notation, for every $k > 0$ there exists a $j \in \{0, \ldots, d-1\}$ such that

$$(5.3) \qquad y_0 \ge y_1 \ge \cdots \ge y_j \ge 4k \ge y_{j+1} \ge \cdots \ge y_{d-1} \ge y_d.$$

Note that $q := (y_1 - k, \ldots, y_j - k, 0, \ldots, 0) \in [0,\infty)^d$ satisfies $q \le p^{k,y}$ under the partial order induced by the coordinates, i.e., for all $i = 1, \ldots, d$,

$q_i \leq p_i^{k,y}$. Since $p_i^{k,y} \leq |a_i|$ also, by Lemma 17, the average of a block decreasing function over $B_\infty(a,k)$ is smaller than the average over $B_\infty(q,k)$. Let $u \in B_\infty(q,k)$ be arbitrary. By (5.3),

$$f(u) \leq f(y_1 - 2k, \ldots, y_j - 2k, u_{j+1}, \ldots, u_d) \leq f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d\right).$$

Thus,

$$(5.4) \quad \frac{1}{|B_\infty(a,k)|} \int_{B_\infty(a,k)} f(u)\, du \leq \frac{1}{|B_\infty(q,k)|} \int_{B_\infty(q,k)} f(u)\, du$$

$$\leq \frac{1}{(2k)^{d-j}} \int_{[-k,k]^{d-j}} f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d\right) du_{j+1} \cdots du_d$$

$$\leq \frac{1}{k^{d-j}} \int_{[0,k]^{d-j}} f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d\right) du_{j+1} \cdots du_d.$$

Observe that for averages of block decreasing functions over cubes centered at the origin, the smaller the radius, the greater the average. Thus, if $j \leq d-2$, the right hand side of (5.4) is bounded by

$$\left(\frac{4}{y_{j+1}}\right)^{d-j} \int_{[0, y_{j+1}/4]^{d-j}} f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d\right) du_{j+1} \cdots du_d,$$

while if $j = d-1$, then it is bounded by $f(y/2, 0)$. Hence, for every $y \in A_\sigma$,

$$Mf(y,0) \leq \max\left\{ f\left(\frac{y}{2}, 0\right), \max_{0 \leq j \leq d-2} \left(\frac{4}{y_{j+1}}\right)^{d-j} G_j(y_1, \ldots, y_j, y_{j+1}) \right\},$$

where for $1 \leq j \leq d-2$, G_j is defined via

$$G_j(y_1, \ldots, y_j, y_{j+1}) := \int_{[0, y_{j+1}/4]^{d-j}} f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d\right) du_{j+1} \cdots du_d,$$

and if $j = 0$, we just use the same formula, but integrating $f(u_1, \ldots, u_d)$ (i.e., no y_i appears as an argument of f).

Thus,

$$(5.5) \quad \int_{A_\sigma} Mf(y,0)\, dy \leq \int_{[0,\infty)^{d-1}} f\left(\frac{y}{2}, 0\right) dy$$

$$+ \sum_{j=0}^{d-2} \int_{A_\sigma} \left(\frac{4}{y_{j+1}}\right)^{d-j} G_j(y_1, \ldots, y_j, y_{j+1})\, dy.$$

Note that for $j \in \{0, \ldots, d-2\}$ (see (5.2)),

$$(5.6) \quad A_\sigma \subset \{y \in [0, \infty)^{d-1} : y_{j+1} \geq y_{j+2} \geq \cdots \geq y_{d-1}\},$$

and

$$(5.7) \qquad \int_0^{y_{j+1}} \left(\ldots \left(\int_0^{y_{d-2}} dy_{d-1} \right) \ldots \right) dy_{j+2} = \frac{y_{j+1}^{d-j-2}}{(d-j-2)!}.$$

Using Fubini's Theorem and the definition of G_j we have

$$(5.8) \qquad \int_{[0,\infty)^{j+1}} \frac{G_j(y_1, \ldots, y_j, y_{j+1})}{y_{j+1}^2} \, dy_1 \cdots dy_{j+1}$$

$$= \int_{[0,\infty)^{j+1}} \frac{1}{y_{j+1}^2} \int_{[0,y_{j+1}/4]^{d-j}} f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d \right)$$
$$du_{j+1} \cdots du_d \, dy_1 \cdots dy_{j+1}$$

$$\leq \int_{[0,\infty)^d} \left(f\left(\frac{y_1}{2}, \ldots, \frac{y_j}{2}, u_{j+1}, \ldots, u_d \right) \int_{\max\{u_{j+1}, \ldots, u_d\}}^{\infty} \frac{dy_{j+1}}{y_{j+1}^2} \right)$$
$$dy_1 \cdots dy_j \, du_{j+1} \cdots du_d$$

$$= \int_{[0,\infty)^d} \frac{f(y_1/2, \ldots, y_j/2, u_{j+1}, \ldots, u_d)}{\max\{u_{j+1}, \ldots, u_d\}} \, dy_1 \cdots dy_j \, du_{j+1} \cdots du_d.$$

The assumption $j \leq d-2$ is used in the next application of Fubini's Theorem. To prove the bound

$$(5.9) \qquad \int_{[0,\infty)^d} \frac{f(u)}{\max\{u_{j+1}, \ldots, u_d\}} \, du \leq \sum_{k=j+1}^{d} \int_{[0,\infty)^{d-1}} f(\hat{u}_k, 0) \, d\hat{u}_k,$$

split $[0, \infty)^d$ into the regions where each u_k is the maximum value. Suppose, for instance, that we are considering

$$E_d := [0, \infty)^d \cap \{\max\{u_{j+1}, \ldots, u_d\} = u_d\}.$$

Now $j + 1 \leq d - 1$, so we can select a u_k with $k \neq d$ and replace u_k with 0. Since f is block decreasing,

$$\int_{E_d} \frac{f(u)}{u_d} \, du \leq \int_{[0,\infty)^{d-1}} f(\hat{u}_k, 0) \int_0^{u_d} \frac{1}{u_d} \, du_k \, d\hat{u}_k = \int_{[0,\infty)^{d-1}} f(\hat{u}_k, 0) \, d\hat{u}_k,$$

and (5.9) follows. Next, using (5.5–5.9), we obtain

$$\int\limits_{A_\sigma} Mf(y,0)\,dy \leq 2^{d-1} \int\limits_{[0,\infty)^{d-1}} f(y,0)\,dy$$

$$+ \sum_{j=0}^{d-2} \frac{4^{d-j}}{(d-j-2)!} \int\limits_{[0,\infty)^d} \frac{f(y_1/2,\ldots,y_j/2,u_{j+1},\ldots,u_d)}{\max\{u_{j+1},\ldots,u_d\}}$$

$$dy_1 \cdots dy_j \, du_{j+1} \cdots du_d$$

$$= 2^{d-1} \int\limits_{[0,\infty)^{d-1}} f(y,0)\,dy + \sum_{j=0}^{d-2} \frac{2^{2d-j}}{(d-j-2)!} \int\limits_{[0,\infty)^d} \frac{f(u)}{\max\{u_{j+1},\ldots,u_d\}}\,du$$

$$\leq 2^{d-1} \int\limits_{[0,\infty)^{d-1}} f(y,0)\,dy + \sum_{j=0}^{d-2} \frac{2^{2d-j}}{(d-j-2)!} \sum_{k=j+1}^{d} \int\limits_{[0,\infty)^{d-1}} f(\hat{u}_k,0)\,d\hat{u}_k$$

$$\leq c'(d) \sum_{k=1}^{d} \int\limits_{[0,\infty)^{d-1}} f(\hat{u}_k,0)\,d\hat{u}_k.$$

Finally, (5.1) follows by applying the same estimate to each of the $d!$ regions A_σ and adding up. ∎

REMARK 21. Recalling inequalities (4.1) and (4.2), it is natural to define a "partial variation" for each variable x_i:

$$V_i(f,\mathbb{R}^d) := \int\limits_{\mathbb{R}^{d-1}} V(f_{\hat{x}_i},\mathbb{R})\,d\hat{x}_i.$$

We have seen that the variation of f controls the variation of $M_\mu f$ (Theorem 7). Here we show that the partial variations of f do not individually control the corresponding partial variations of $M_\mu f$, something that makes the proof of Theorem 7 harder than it would otherwise be. To see that the inequality

$$V_i(M_\mu f,\mathbb{R}^d) \leq c(\mu,d) V_i(f,\mathbb{R}^d)$$

may fail, consider the following counterexample in the case $\mu = \|\cdot\|_\infty$. Let g be a nonincreasing function on $[0,\infty)$ such that $g(0) = 1$ and $g(\infty) = 0$. Suppose also $\|g\|_1 = 1$ and $\int_0^\infty g(u)\,du/u = \infty$. For $x_1, x_2 \in \mathbb{R}$ and $m \in \mathbb{N}$ we define the block decreasing functions $f_m(x_1,x_2) = mg(m|x_1|)g(|x_2|)$. Then $V_2(f_m,\mathbb{R}^2) = 4\int_0^\infty f_m(x_1,0)\,dx_1 = 4g(0)\int_0^\infty mg(mx_1)\,dx_1 = 4g(0)\|g\|_1 = 4$. On the other hand, for $x_1 > 0$,

$$M_\infty f_m(x_1,0) \geq \frac{1}{x_1^2} \int_0^{x_1} mg(mu_1)\,du_1 \int_0^{x_1} g(u_2)\,du_2$$

$$= \frac{1}{x_1^2} \int_0^{mx_1} g(u_1)\,du_1 \int_0^{x_1} g(u_2)\,du_2 =: F_m(x_1).$$

Now $\lim_{m\to\infty} F_m(x_1) = \|g\|_1 \int_0^{x_1} g(u_2)\,du_2/x_1^2 = \int_0^{x_1} g(u_2)\,du_2/x_1^2$, so, using monotone convergence and the Fubini–Tonelli Theorem, we obtain

$$\lim_{m\to\infty} V_2(M_\infty f_m, \mathbb{R}^2) \geq \lim_{m\to\infty} \int_0^\infty F_m(x_1)\,dx_1$$

$$= \int_0^\infty \left(\int_0^{x_1} g(u_2)\,du_2 \right) \frac{dx_1}{x_1^2} = \int_0^\infty g(u_2)\,\frac{du_2}{u_2} = \infty.$$

It is easy to check that this example can be adapted to $M_\mu f$, where μ is any unconditional norm.

6. Continuity of the maximal function. In this section we prove Theorem 8, showing that if f is a block decreasing function of finite variation, then $M_\mu f$ is continuous, except perhaps on a negligible $(d-1)$-dimensional Hausdorff measurable set.

First we recall the notion of approximate continuity (see [EvGa, pp. 47 and 209]). Note that the definitions and results from [EvGa] are given in terms of euclidean balls, so we need to use the equivalence of all norms in $\mathbb{R}^d$.

DEFINITION 22. Let $f : \mathbb{R}^d \to \mathbb{R}$. We say that l is the *approximate limit* of f as $y \to x$, and write

$$\mathrm{ap}\lim_{y\to x} f(y) = l,$$

if for each $\varepsilon > 0$,

$$\lim_{r\to 0} \frac{|B_2(x,r) \cap \{|f - l| \geq \varepsilon\}|}{|B_2(x,r)|} = 0.$$

That is, if l is the approximate limit of f at x, then for all $\varepsilon > 0$ the sets $\{|f - l| \geq \varepsilon\}$ have density zero at x.

DEFINITION 23. Let $f : \mathbb{R}^d \to \mathbb{R}$. We say that $f_{\sup}(x)$ is the *approximate limsup* of f as $y \to x$ if

$$f_{\sup}(x) := \mathrm{ap}\limsup_{y\to x} f(y) = \inf\left\{ t : \lim_{r\to 0} \frac{|B_2(x,r) \cap \{f > t\}|}{r^d} = 0 \right\}.$$

Likewise, $f_{\inf}(x)$ is the *approximate liminf* of f as $y \to x$ if

$$f_{\inf}(x) := \mathrm{ap}\liminf_{y\to x} f(y) = \sup\left\{ t : \lim_{r\to 0} \frac{|B_2(x,r) \cap \{f < t\}|}{r^d} = 0 \right\}.$$

It is well known that for measurable functions the approximate limit exists a.e. [EvGa, Theorem 3, p. 47]. For locally integrable functions, this follows from the Lebesgue Differentiation Theorem. For locally integrable functions of finite variation, the approximate limsup and liminf are finite $\mathcal{H}^{d-1}$-a.e. on $\mathbb{R}^d$ (cf. [EvGa, Theorem 2, p. 211]; actually, the results from

[EvGa] are stated for BV functions, so $f \in L^1$; but it is easy to check that local integrability of f suffices to carry out the arguments).

DEFINITION 24. The *jump set* $J_f := \{f_{\inf}(x) < f_{\sup}(x)\}$ of f is the set of points where the approximate limit of f does not exist.

DEFINITION 25. Let v be a unit vector in $\mathbb{R}^d$ and let $x \in \mathbb{R}^d$. We define the half-spaces associated to x and v by

$$H_v^+ := \{y \in \mathbb{R}^d : v \cdot (y - x) \geq 0\}, \quad H_v^- := \{y \in \mathbb{R}^d : v \cdot (y - x) \leq 0\},$$

where the symbol $\cdot$ denotes the usual scalar product in $\mathbb{R}^d$.

While the notation does not make it explicit, H_v^+ and H_v^- depend on x and contain it as a boundary point. Set

$$F(x) := \frac{f_{\sup}(x) + f_{\inf}(x)}{2}.$$

From [EvGa, p. 213, Theorem 3] we get

THEOREM 26. *If $f \in L_{\mathrm{loc}}^1(\mathbb{R}^d)$ and $V(f, \mathbb{R}^d) < \infty$, then:*

(i) *for $\mathcal{H}^{d-1}$-a.e. $x \in \mathbb{R}^d - J_f$,*

$$\lim_{r \to 0^+} \frac{1}{|B_2(x,r)|} \int_{B_2(x,r)} |f(y) - F(x)|^{d/(d-1)} \, dy = 0;$$

(ii) *for $\mathcal{H}^{d-1}$-a.e. $x \in J_f$, there exists a unit vector $v \equiv v(x)$ such that*

$$\lim_{r \to 0^+} \frac{1}{|B_2(x,r)|} \int_{B_2(x,r) \cap H_v^-} |f(y) - f_{\sup}(x)|^{d/(d-1)} \, dy = 0,$$

and

$$\lim_{r \to 0^+} \frac{1}{|B_2(x,r)|} \int_{B_2(x,r) \cap H_v^+} |f(y) - f_{\inf}(x)|^{d/(d-1)} \, dy = 0.$$

In this theorem, euclidean balls can be replaced by balls defined using any other norm μ, upon noting that integrands are nonnegative, and then inserting a suitable constant. Of course, the corresponding limits are still zero. So for $\mathcal{H}^{d-1}$-almost every $x \in \mathbb{R}^d - J_f$, $F = f^*$, the precise representative of f defined using μ.

A different definition of approximate limits and related notions appears in [AFP] (using integral averages, in the line of the preceding theorem). But all such definitions coincide $\mathcal{H}^{d-1}$-a.e. with the ones given above, so it is actually immaterial which ones we use.

Since approximate limit exist for a.e. $x \in \mathbb{R}^d$, all the functions f^*, F, $f_{\inf}$ and $f_{\sup}$ represent the same equivalence class $[f]$. To study the continuity of the maximal function it will be convenient for us to use $f_{\sup}$.

LEMMA 27. *If f is block decreasing on $\mathbb{R}^d$, then $f_{\sup}$ is block decreasing.*

Proof. Fix $x_1, \ldots, x_d \geq 0$, $h > 0$, and $t > 0$. We prove that $f_{\sup}(x) \geq f_{\sup}(x + he_d)$, the argument being the same for the other coordinates. By Definition 23, it is enough to show that for all sufficiently small $r > 0$, $|B_2(x+he_d, r) \cap \{f > t\}| \leq |B_2(x, r) \cap \{f > t\}|$. Suppose $0 < r < h/2$. Given $y \in B_2(x+he_d, r) \cap \{f > t\}$, we have $y - he_d \in B_2(x, r)$, and by the choice of r, $y_d \geq |y_d - h|$, so $t < f(y) \leq f(y - he_d)$. Thus $(B_2(x+he_d, r) \cap \{f > t\}) - he_d \subset B_2(x, r) \cap \{f > t\}$, and $|B_2(x + he_d, r) \cap \{f > t\}| \leq |B_2(x, r) \cap \{f > t\}|$ follows by the translation invariance of Lebesgue measure. ∎

The next lemma states that in any half-ball resulting from the intersection of a euclidean ball $B_2(x, r)$ with a half-space having x in its boundary, there is a comparable μ-ball contained in the half-ball and containing x (as a boundary point, of course).

LEMMA 28. *Let μ be an arbitrary norm on $\mathbb{R}^d$, let $r > 0$, and let $x, v \in \mathbb{R}^d$, where $\|v\|_2 = 1$. Then there exists a constant $k_\mu > 0$ such that for every half-ball $B_2(x, r) \cap H_v^+$, we can find a center $c \in \mathbb{R}^d$ and a radius $\varrho > 0$ with*

$$x \in B_\mu(c, \varrho) \subset B_2(x, r) \cap H_v^+ \quad and \quad \frac{|B_\mu(c, \varrho)|}{|B_2(x, r)|} \geq k_\mu.$$

Proof. By a translation we may assume that $x = 0$. Let $B_\mu(0, \varrho)$ be the largest μ-ball contained in $B_2(0, r/4)$. By the convexity of $B_\mu(0, \varrho)$, it can be translated, say, to $B_\mu(c, \varrho)$, in such a way that 0 belongs to the boundary of $B_\mu(c, \varrho)$ and this ball is contained in H_v^+. Since $0 \in B_\mu(c, \varrho) \subset B_2(c, r/4) \subset B_2(0, r)$, it follows that $B_\mu(c, \varrho) \subset B_2(0, r) \cap H_v^+$. Furthermore, if $t > 0$, then $|B_\mu(c, t\varrho)| = t^d |B_\mu(c, \varrho)|$, by the scaling properties of Lebesgue measure. Using the fact that all norms in $\mathbb{R}^d$ are equivalent, we let $t > 0$ be the smallest real number such that $B_2(0, r) \subset B_\mu(0, t\varrho)$, and conclude that $|B_2(0, r)| \leq t^d |B_\mu(c, \varrho)|$. Then we take $k_\mu = t^d$. ∎

LEMMA 29. *Let $0 \leq f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, let $V(f, \mathbb{R}^d) < \infty$, and let μ be a norm. Then for $\mathcal{H}^{d-1}$-almost every $x \in \mathbb{R}^d$, we have $M_\mu f(x) \geq f_{\sup}(x)$.*

Proof. For every $r > 0$,

$$M_\mu f(x) \geq \frac{1}{|B_\mu(x, r)|} \int_{B_\mu(x,r)} f = f_{\sup}(x) + \frac{1}{|B_\mu(x, r)|} \int_{B_\mu(x,r)} (f(y) - f_{\sup}(x)) \, dy,$$

so it is enough to show that for $\mathcal{H}^{d-1}$-almost every $x \in \mathbb{R}^d$, the rightmost term in the preceding inequality tends to 0 as $r \to 0^+$. If $x \in \mathbb{R}^d \setminus J_f$, then Theorem 26(i) yields the result: First, replace μ-balls by euclidean balls (perhaps modifying some constant) and then use Jensen's inequality. And if $x \in J_f$, a similar argument, using Theorem 26(ii) and Lemma 28, yields the same result. ∎

LEMMA 30. *Let f be a block decreasing function on $\mathbb{R}^d$, and let $x \in (0, \infty)^d$. Then $f_{\sup}$ is upper semicontinuous at x.*

Proof. Suppose otherwise. Since $f_{\sup}$ is block decreasing, there is an $\varepsilon > 0$ such that for all z in the rectangle $\prod_{i=1}^d (0, x_i)$, $f_{\sup}(z) > f_{\sup}(x) + \varepsilon$. Then the density at x of the set $\{f > f_{\sup}(x) + \varepsilon\}$ is at least $1/2^d$, contradicting the definition of $f_{\sup}$. ∎

REMARK 31. Let f be block decreasing and let $x \in (0, \infty)^d$ be such that $f(x) = f_{\sup}(x)$. Arguing as in Lemma 30, we conclude that f is upper semicontinuous at x. Thus, a block decreasing function is upper semicontinuous at almost every point in $\mathbb{R}^d$.

REMARK 32. By Lemma 30, if f is block decreasing and $f_{\sup}$ is not upper semicontinuous at x, then at least one of x's coordinates must be zero. We consider these points in the next lemma.

LEMMA 33. *Let f be a block decreasing function on $\mathbb{R}^d$. Then for almost all $(x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1}$, $f_{\sup}$ is upper semicontinuous at $(x_1, \ldots, x_{d-1}, 0)$.*

Proof. Writing $\hat{x}_d = (x_1, \ldots, x_{d-1})$, $\hat{y}_d = (y_1, \ldots, y_{d-1})$, and $g(\hat{y}_d) := f_{\sup}(\hat{y}_d, 0)$, we note, first, that by Remark 31, g is upper semicontinuous for a.e. $\hat{x}_d \in \mathbb{R}^{d-1}$. Thus, it is enough to check that if g is upper semicontinuous at $\hat{x}_d$, then $f_{\sup}$ is upper semicontinuous at $(\hat{x}_d, 0)$. But this follows from the block decreasing property of $f_{\sup}$:

$$\limsup_{y \to (\hat{x}_d, 0)} f_{\sup}(y) \le \limsup_{y \to (\hat{x}_d, 0)} f_{\sup}(\hat{y}_d, 0) = \limsup_{\hat{y}_d \to \hat{x}_d} g(\hat{y}_d)$$
$$\le g(\hat{x}_d) = f_{\sup}(\hat{x}_d, 0). \blacksquare$$

Proof of Theorem 8. According to Lemma 3.4 of [AlPe1], if a locally integrable function $h \ge 0$ is upper semicontinuous at w and $h(w) \le Mh(w)$, then Mh is continuous at w. Now Lemmas 30 and 33 entail that $f_{\sup}$ is upper semicontinuous at $\mathcal{H}^{d-1}$-almost every point, while by Lemma 29, $f_{\sup}(w) \le M_\mu f_{\sup}(w)$ for $\mathcal{H}^{d-1}$-a.e. w. Since $M_\mu f = M_\mu f_{\sup}$, the result follows. ∎

REMARK 34. Note that the maximal function of a block decreasing function need not be continuous. Consider, for instance, the maximal function M_∞, defined using cubes, and let f be the characteristic function of the unit ℓ^p-quasiball in $\mathbb{R}^2$, where $0 < p \le 1$ is fixed. Then $M_\infty f$ is discontinuous at $(1, 0)$.

7. The derivative of the maximal function. Let us recall a few facts about the distributional derivative of a function of finite variation (cf. [AFP, pp. 184–186]). The measure Df vanishes on $\mathcal{H}^{d-1}$-negligible sets. Its absolutely continuous part $D^a f$ is obtained by integrating the density of Df with respect to the d-dimensional Lebesgue measure, and lives in the set $\mathcal{D}_f$

where f is approximately differentiable (cf. [AFP, Definition 3.70, p. 165] for the definition of approximate differentiability). The singular part of Df can be decomposed into a Cantor part and a jump part. The Cantor part $D^c f$ gives full measure to the set where f is approximately continuous, and the jump part $D^j f$ gives full measure to the jump set of f (a countably $\mathcal{H}^{d-1}$-rectifiable set). The measure $D^a f + D^c f$ vanishes on sets of finite (and thus of σ-finite) $\mathcal{H}^{d-1}$-measure.

Let $f \in L^1_{\text{loc}}(\mathbb{R}^d)$. From now on, we assume that $|f| = |f|^*$, the precise representative of $|f|$. In what follows, Lipschitz constants are determined by using the ℓ_2-norm.

We define

$$(7.1) \quad E_{n,k} := \left\{ x \in \mathbb{R}^n : \text{there exists a ball } B := B_\mu(c,r) \text{ with } x \in B, \right.$$

$$\left. r \geq 1/n, \; \frac{1}{|B|} \int_B |f(y)|\, dy = M_\mu f(x), \text{ and } M_\mu f(x) \leq k \right\}.$$

LEMMA 35. *Let f be a locally integrable function, and let $c_\mu > 0$ be such that for all $w \in \mathbb{R}^d$, $\|w\|_\mu \leq c_\mu \|w\|_2$. Then the restriction of $M_\mu f$ to $E_{n,k}$ is Lipschitz, with Lipschitz constant $\text{Lip}(M_\mu f) \leq c_\mu dkn$.*

Proof. Let $x, y \in E_{n,k}$. By symmetry, we may assume that $M_\mu f(x) \geq M_\mu f(y)$. Suppose $M_\mu f(x) = |B|^{-1} \int_B |f|$, where $x \in B := B_\mu(c,r)$. Now since $x \in B_\mu(c,r)$, we have $y \in B_\mu(c, r + \|x-y\|_\mu)$. Thus,

$$\frac{|M_\mu f(x) - M_\mu f(y)|}{\|x-y\|_2} = \frac{M_\mu f(x) - M_\mu f(y)}{\|x-y\|_2}$$

$$\leq \frac{M_\mu f(x) - \frac{|B_\mu(c,r)|}{|B_\mu(c,r+\|x-y\|_\mu)|} \frac{1}{|B_\mu(c,r)|} \int_{B_\mu(c,r)} |f|}{\|x-y\|_2}$$

$$\leq \frac{M_\mu f(x)}{\|x-y\|_2} \left(1 - \frac{|B_\mu(c,1)|r^d}{|B_\mu(c,1)|(r+\|x-y\|_\mu)^d} \right)$$

$$\leq \frac{c_\mu k}{\|x-y\|_\mu} \left(1 - \frac{1}{(1+\|x-y\|_\mu/r)^d} \right)$$

$$\leq \frac{c_\mu k}{\|x-y\|_\mu} \left(1 - \frac{1}{(1+n\|x-y\|_\mu)^d} \right).$$

Now $g(a) := a^{-1}(1 - (1+na)^{-d})$ is decreasing on $\{a > 0\}$, and $\lim_{a \to 0+} g(a) = dn$, so from the preceding inequality we obtain

$$\frac{|M_\mu f(x) - M_\mu f(y)|}{\|x-y\|_2} \leq c_\mu kdn. \quad \blacksquare$$

Next we deal with the local maximal function $M_{R,\mu}$. Set

(7.2) $\qquad E_{R,n} := \{x \in \mathbb{R}^n : \text{there exists a ball } B := B_\mu(c,r) \text{ with } x \in B,$

$$r \in [1/n, R], \text{ and } \frac{1}{|B|} \int_B |f(y)|\, dy = M_{R,\mu}f(x)\Big\}.$$

Of course, if $R < 1/n$, then $E_{R,n}$ is empty.

LEMMA 36. *Fix $R > 0$, and let $c_\mu := 1/|B_\mu(0,1)|$. If f is locally integrable and has finite variation, then the restriction of $M_{R,\mu}f$ to $E_{R,n}$ is Lipschitz, with Lipschitz constant $\operatorname{Lip}(M_{R,\mu}f) \leq c_\mu n^d V(f, \mathbb{R}^d)$.*

Proof. As noted above, we may assume that $R \geq 1/n$. Otherwise $E_{R,n} = \emptyset$ and there is nothing to prove. So let $x, y \in E_{R,n}$, and suppose that $M_\mu f(x) = |B|^{-1} \int_B |f| \geq M_\mu f(y)$, where $x \in B := B_\mu(c,r)$ and $1/n \leq r \leq R$. It follows from [AFP, Exercise 3.3, p. 208] that for every bounded measurable set K,

$$\int_K \frac{|f(x+h) - f(x)|}{\|h\|_2}\, dx \leq V(f, \mathbb{R}^d).$$

Thus,

$$\frac{|M_{R,\mu}f(x) - M_{R,\mu}f(y)|}{\|x - y\|_2} \leq \frac{\frac{1}{|B_\mu(c,r)|} \int_{B_\mu(c,r)} |f| - \frac{1}{|B_\mu(c+y-x,r)|} \int_{B_\mu(c+y-x,r)} |f|}{\|x - y\|_2}$$

$$\leq \frac{1}{|B_\mu(c,r)|} \int_{B_\mu(c,r)} \frac{|f(u) - f(u - x + y)|}{\|x - y\|_2} du \leq c_\mu n^d V(f, \mathbb{R}^d). \blacksquare$$

Since the exact size of the Lipschitz constants is irrelevant in the argument that follows, from now on we only consider M_μ.

The following lemma appears in [EvGa, p. 75]. We mention that in [EvGa], the same notation is used for Hausdorff measures and the outer measures they generate; in particular, the result below applies to arbitrary sets E.

LEMMA 37. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant $\operatorname{Lip}(f)$ and let $s > 0$. Then $\mathcal{H}^s(f(E)) \leq \operatorname{Lip}(f)^s \mathcal{H}^s(E)$ for all $E \subset \mathbb{R}^d$.*

Define $E := \bigcup_{n,k \in \mathbb{N}} E_{n,k}$, and note that $\mathbb{R}^d - E \subset \{M_\mu f = f^*\}$. Now it is to be expected that $DM_\mu f$ has no Cantor part on E, since $M_\mu f$ is Lipschitz on the sets $E_{n,k}$, and likewise, that $DM_\mu f$ has no Cantor part on $\{M_\mu f = f^*\}$, since by hypothesis f^* is of SBV, so $D^c f \equiv 0$. We prove that this is indeed the case, by restricting functions to lines.

LEMMA 38. *If $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ has finite variation, then $M_\mu f$ maps $\mathcal{H}^1$-negligible subsets of E into $\mathcal{H}^1$-negligible sets.*

Proof. Fix k and n, and let $N \subset \mathbb{R}^d$ be an $\mathcal{H}^1$-null subset of $\mathbb{R}^d$. By the previous lemma, $|M_\mu f(N \cap E_{n,k})| = 0$; here the absolute value signs stand for the 1-dimensional Lebesgue measure, which on the real line coincides with $\mathcal{H}^1$; while the preceding lemma refers to Lipschitz functions defined on all $\mathbb{R}^d$, one can always extend a Lipschitz function from a subset to the whole space $\mathbb{R}^d$, with the same constant by Kirszbraun's theorem, or, if one is not concerned about the constant, as in our case, by simpler extension theorems. Since a countable union of null sets is null, the result follows. ∎

We shall use a variant of the Banach–Zarecki Theorem (which states that a real-valued continuous function on a compact interval is absolutely continuous if and only if it is of bounded variation and maps null sets to null sets). As stated, the result fails for $\mathbb{R}$ even if f is bounded; for instance, the function $\sin x$ is absolutely continuous and has infinite variation. However, under the additional assumption that $f \geq 0$ is block decreasing, the variation is bounded by $2f(0)$, so the following version of the Banach–Zarecki Theorem does hold for $\mathbb{R}$.

LEMMA 39. *Let $f : \mathbb{R} \to [0, \infty)$ be a continuous, block decreasing function. Then f is absolutely continuous if and only if f maps measure zero sets to measure zero sets.*

We use the following notation to express the decomposition of a function h on $\mathbb{R}^d$ into functions $h_j(x'; t)$ defined on lines. For any $j = 1, \ldots, d$ and $x' = (x'_1, \ldots, x'_{d-1}) \in \mathbb{R}^{d-1}$ we set $h_j(x'; t) := h(x'_1, \ldots, x'_{j-1}, t, x'_j, \ldots, x'_{d-1})$. This decomposition of h leads to the corresponding disintegration result for Dh (cf. [AFP]). Regarding the Cantor part $D^c h$ of Dh, it follows from [AFP, Theorem 3.108] that it can be recovered from the Cantor parts of the derivatives of the restrictions of h to lines, and in particular, $D^c h = 0$ if and only if for almost every line parallel to the jth coordinate axis, $D^c h_j = 0$, where $j = 1, \ldots, d$. To show that on these lines the functions $h_j(x'; t)$ map Lebesgue null sets to Lebesgue null sets, we modify them so that they become continuous, and then apply the preceding lemma.

Recall that, in order to simplify notation, we are assuming that a function $f \geq 0$ is everywhere equal to its precise representative f^*.

LEMMA 40. *Let $f : L^1_{\text{loc}}(\mathbb{R}^d) \to [0, \infty]$ be a finite variation, block decreasing function with $|D^c f|(\mathbb{R}^d) = 0$. Then for $\mathcal{H}^{d-1}$-a.e. $x' \in \mathbb{R}^{d-1}$, and for every $j = 1, \ldots, d$, the function $f_j(x'; \cdot) : \mathbb{R} \to \mathbb{R}$ maps $\mathcal{H}^1$-negligible sets into $\mathcal{H}^1$-negligible sets.*

Proof. Fix $\varepsilon > 0$ and $j \in \{1, \ldots, d\}$. By (4.2), for a.e. $x' \in \mathbb{R}^{d-1}$, $f_j(x'; \cdot)$ is a function with finite variation, since $V(f, \mathbb{R}^d) < \infty$, and by [AFP, Theorem 3.108], for a.e. $x' \in \mathbb{R}^{d-1}$, $f_j(x'; \cdot)$ has null Cantor derivative. Denote by A the subset of all $x' \in \mathbb{R}^{d-1}$ for which both of the preceding conditions

hold. Note in particular that $f_j(x'; \cdot) < \infty$ on A. For each $x' \in A$, $f_j(x'; \cdot)$ is a nonnegative, real-valued function with at most a countable number of jump discontinuities. Next we modify $f_j(x'; \cdot)$ so it becomes continuous.

Suppose $\{t_n\}_{n=1}^\infty$ is a listing of the set where $f_j(x'; \cdot)$ has jumps. Since both the right and left limits $f_j(x'; t_n+)$ and $f_j(x'; t_n-)$ exist, there is $\delta_n > 0$ making the images under $f_j(x'; t)$ of the intervals $[t_n - \delta_n, t_n]$ and $[t_n, t_n + \delta_n]$ so small that $|f_j(x'; [t_n - \delta_n, t_n + \delta_n])| < \varepsilon/2^n$. Thus, $|\bigcup_n f_j(x'; [t_n - \delta_n, t_n + \delta_n])| < \varepsilon$. Actually, we want to select δ_n satisfying some additional conditions that will make it possible to obtain disjoint intervals. Using countability we assume that δ_n is chosen so $t_n - \delta_n$ and $t_n + \delta_n$ do not belong to the jump set of $f_j(x'; \cdot)$. We define inductively the sequence of disjoint intervals as follows, starting with $n_1 = 1$: Given $[t_1 - \delta_1, t_1 + \delta_1]$, let t_{n_2} be the first jump point in the list not belonging to $(t_1 - \delta_1, t_1 + \delta_1)$, if there is any (otherwise stop here). Then choose δ_{n_2}, satisfying the conditions above, so that additionally $[t_1 - \delta_1, t_1 + \delta_1] \cap [t_{n_2} - \delta_{n_2}, t_{n_2} + \delta_{n_2}] = \emptyset$. Now repeat the process, letting t_{n_3} be the first jump point not contained in the union of the preceding two intervals, and so on. Once the process stops, we define $h(t) := f_j(x'; t)$ on $\mathbb{R} - \bigcup_{n_k}[t_{n_k} - \delta_{n_k}, t_{n_k} + \delta_{n_k}]$, and on each $[t_{n_k} - \delta_{n_k}, t_{n_k} + \delta_{n_k}]$, we extend $h(t)$ affinely from $f_j(x'; t_{n_k} - \delta_{n_k})$ to $f_j(x'; t_{n_k} + \delta_{n_k})$. Now, let $N \subset \mathbb{R}$ be null. Then $|h(N)| = 0$ by Lemma 39, so $|f_j(x'; N)| \leq |h(N)| + |f_j(x'; \bigcup_{n_k}[t_{n_k} - \delta_{n_k}, t_{n_k} + \delta_{n_k}])| < \varepsilon$, and since ε is arbitrary, we conclude that $|f_j(x'; N)| = 0$. ■

Proof of Theorem 11. It is enough to show that $M_\mu f$ is absolutely continuous on lines (ACL), that is, given any coordinate axis, say x_d without loss of generality, $M_\mu f$ is absolutely continuous on almost all lines parallel to the the x_d-axis, where "almost all" refers to Lebesgue measure on the intersection of the lines with the $(d-1)$-subspace perpendicular to them. Sobolev Theory then entails that $M_\mu f \in W^{1,1}_{\text{loc}}(\mathbb{R}^d)$, so the distributional gradient $\nabla M_\mu f$ exists as a locally integrable (vector-valued) function. Furthermore, since by Theorem 7 we have

$$V(M_\mu f, \mathbb{R}^d) \leq c_{\mu,d} V(f, \mathbb{R}^d),$$

from the fact that the variation of f is finite, and under the assumption that $M_\mu f$ is ACL, we obtain

$$\int_{\mathbb{R}^d} |\nabla M_\mu f| = |DM_\mu f|(\mathbb{R}^d) = V(M_\mu f, \mathbb{R}^d) \leq c_{\mu,d} V(f, \mathbb{R}^d) < \infty,$$

so $\nabla M_\mu f \in L^1(\mathbb{R}^d)$. Next we show that $M_\mu f$ is ACL.

Recall that $E := \bigcup_{n,k} E_{n,k}$, where $E_{n,k}$ is given by (7.1). Consider the set of all lines parallel to the last coordinate x_d (of course, any other coordinate axis will do equally well). Let A be the set of all $y \in \mathbb{R}^{d-1}$ such that $f_d(y; \cdot)$ is a function of finite variation and $M_\mu f(y, t)$ is a continuous function of t. By

Theorem 8 and by (4.2), $|A^c| = 0$; here $|\cdot|$ stands for the $(d-1)$-dimensional Lebesgue measure. Given $y \in A$, denote by L_y the vertical line $\{(y, t) : t \in \mathbb{R}\}$. Let $N_y \subset L_y$ be null with respect to the 1-dimensional Lebesgue measure, denoted in what follows by $|\cdot|$. By Lemma 40, for almost all $y \in A$, $|f(N_y)| = 0$. Now $|M_\mu f(N_y \cap E)| = 0$ by Lemma 38, and $E^c \subset \{M_\mu f = f\}$, so for almost all $y \in A$, $|M_\mu f(N_y \cap E^c)| \leq |f(N_y)| = 0$. Thus, $|M_\mu f(N_y)| = 0$ for almost every $y \in \mathbb{R}^{d-1}$. ∎

Proof of Theorem 12. As in the previous argument, it is enough to show that $M_\infty f$ is ACL. We remark that the "no Cantor part of the derivative" hypothesis is not needed here, since we are using cubes and f is block decreasing.

Let $x \in (0, \infty)^d$ (the argument is the same for the other octants) and suppose, without loss of generality, that $x_1 = \min_{1 \leq i \leq d} x_i$. Since f is block decreasing, it is easy to see that in order to compute $M_\infty f(x)$, it is enough to consider cubes $B_\infty(c, r)$ of sidelength at least x_1. Thus, $[x_1, \infty)^d \subset E_{n,k}$ for $n \geq 1/x_1$ and $k \geq M_\infty f(x_1, x_1, \ldots, x_1)$ (the set $E_{n,k}$ was defined in (7.1)). So $M_\infty f$ is Lipschitz on every open set compactly contained in $(0, \infty)^d$. It follows that if L is any line parallel to the x_d-coordinate axis and $N \subset L$ is 1-null, then $M_\infty f(N \cap (0, \infty)^d)$ is 1-null. Since the same result holds for the intersection of N with any other open octant, we can conclude that $|M_\infty f(N)| = 0$, unless for some $i = 1, \ldots, d-1$, $L \subset \{x_i = 0\} \times \mathbb{R}$. But this is a null set of lines, so $M_\infty f$ is ACL by Theorem 8 and Lemma 39. ∎

References

[ACP]   J. M. Aldaz, L. Colzani and J. Pérez Lázaro, *Optimal bounds on the modulus of continuity of the uncentered Hardy–Littlewood maximal function*, preprint.

[AlPe1]   J. M. Aldaz and J. Pérez Lázaro, *Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities*, Trans. Amer. Math. Soc. 359 (2007), 2443–2461.

[AlPe2]   —, —, *Boundedness and unboundedness results for some maximal operators on functions of bounded variation*, J. Math. Anal. Appl. 337 (2008), 130–143.

[AFP]   L. Ambrosio, N. Fusco and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*, Oxford Math. Monogr., Clarendon Press, Oxford, 2000.

[An]   T. W. Anderson, *The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities*, Proc. Amer. Math. Soc. 6 (1955), 170–176.

[Bu]   S. M. Buckley, *Is the maximal function of a Lipschitz function continuous?*, Ann. Acad. Sci. Fenn. Math. 24 (1999), 519–528.

[EvGa]   L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions,* CRC Press, Boca Raton, FL, 1999.

[Ga]   R. J. Gardner, *The Brunn–Minkowski inequality*, Bull. Amer. Math. Soc. (N.S.) 39 (2002), 355–405.

[HaOn]   P. Hajłasz and J. Onninen, *On boundedness of maximal functions in Sobolev spaces*, Ann. Acad. Sci. Fenn. Math. 29 (2004), 167–176.

[Ki]   J. Kinnunen, *The Hardy–Littlewood maximal function of a Sobolev function*, Israel J. Math. 100 (1997), 117–124.

[KiLi]   J. Kinnunen and P. Lindqvist, *The derivative of the maximal function*, J. Reine Angew. Math. 503 (1998), 161–167.

[KiSa]   J. Kinnunen and E. Saksman, *Regularity of the fractional maximal function*, Bull. London Math. Soc. 34 (2003), 529–535.

[Ko1]   S. Korry, *A class of bounded operators on Sobolev spaces*, Arch. Math. (Basel) 82 (2004), 40–50.

[Ko2]   —, *Boundedness of Hardy–Littlewood maximal operator in the framework of Lizorkin–Triebel spaces*, Rev. Mat. Complut. 15 (2002), 401–416.

[Lu]   H. Luiro, *Continuity of the maximal operator in Sobolev spaces*, Proc. Amer. Math. Soc. 135 (2007), 243–251.

[Ta]   H. Tanaka, *A remark on the derivative of the one-dimensional Hardy–Littlewood maximal function*, Bull. Austral. Math. Soc. 65 (2002), 253–258.

Departamento de Matemáticas
Universidad Autónoma de Madrid
Cantoblanco
28049 Madrid, Spain
E-mail: jesus.munarriz@uam.es

Departamento de Matemáticas
y Computación
Universidad de La Rioja
Edificio J. L. Vives
Calle Luis de Ulloa s/n
26004 Logroño, Spain
E-mail: javier.perezl@unirioja.es